Investigations into the fluorescent covalent labeling of biomolecules utilizing rhodamine dyes, electrophilic leaving groups and mRNA display.

Susan D. Selaya
Virginia Commonwealth University
Investigations into the fluorescent covalent labeling of biomolecules utilizing rhodamine dyes, electrophilic leaving groups and mRNA display.

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctorate in Philosophy in Chemistry at Virginia Commonwealth University.

By

Susan Daniela Selaya
B.S. in Chemical Education, University of California San Diego, 2001
M.S.c. in Chemistry, University of California Los Angeles, 2006

Director: Mathew C.T. Hartman
Professor, Department of Chemistry
Virginia Commonwealth University
Richmond, Virginia
August, 2014
"Happy are those who spend their days in gaining knowledge,

in discovering the secrets of nature,

and in penetrating the subtleties of pure truth!"

-Abdu’l-Bahá

“Do not allow negative experiences to make you bitter.

They should make you wiser,

and with that wisdom you shall find joy.”

-Leon Brown
Acknowledgements

I’d first like to thank my husband, Edward Gregory Hawkins for his constant support, friendship and unconditional love for the 15 years we have been together. There is nothing that we cannot accomplish together, his guidance, advice and support from childcare to scientific discussions and everything in between is a blessing that I daily appreciate. I’d like to thank my son, Edward Calvin Hawkins for being the final push to help me complete this chapter in my life, his morning smiles remind me that life should be full of joy and to incorporate that mentality into each working day.

I cannot imagine a world void of the unconditional love and support I was raised with by my parents, Dr. Marcos Selaya and Mrs. Susan Selaya. They imparted in me the work ethic, kindness, moral code and fortitude required for every major achievement I have ever accomplished and for that I am forever grateful and hopeful to pass along to my own children. My sister, Dr. Annemarie Selaya and brother, Esteban Selaya are my lifelong advocates and their faith in me, support and unconditional love is a blessing I strive to reciprocate.

Currently, 27.4% of students pursuing doctoral graduate degrees in chemistry are women\(^1\), but there four times more women graduating with a PhD than they did in 1966\(^2\). I’d like to thank the women in my life that opened these doors for themselves and for continuing generations of women and reminded me that there should be no limitations to pursuing graduate degrees and careers based on gender or nationality. Thank you to my mother for taking me to her graduation ceremony so that I saw the most important role model in my life being awarded a graduate degree. Thank you to my sister whose hard work and achievements as a pediatrician has served as a guiding light along my own career path. Thank you to
Ghislaine Hawkins for serving as a mentor and introducing me to tissue culture and virology at the City of Hope. Thanks to Dr. Karen Carter, at Al Biotech for supporting my career and my decision to return to graduate school. Many thanks to Dr. Sonja Strah-Pleynet for introducing me to medicinal chemistry and for guiding me through the world of synthetic organic chemistry and teaching me the bag of tricks I will forever carry with me and pass along to new chemists.

I’d also like to thank my graduate advisor, Dr. Sarah Rutan for supporting me when having a child at the end of my graduate work seemed difficult and to Dr. Stephanie Gould for modeling the success of such a feat. These women have taught me that balance between family and science is possible and I’d like to thank these women for being such tremendous role models and for all of their advice and guidance.

I’d like to thank my mentors and bosses along the way. Dr. Murray Goodman supported and opened the doors of undergraduate research to me, Dr. Bradley Teegarden at Arena Pharmaceuticals honed my writing and presentation skills and pushed me to go to graduate school, Dr. Miguel Garcia-Garibay provided compassionate support at UCLA while I recovered from a debilitating accident, and Dr. Rob Harris at Al Biotech taught me that I can be handed anything new to solve and make it happen. Finally, I would like to thank Dr. Matthew Hartman for serving as my mentor and bringing back into my life the world of biology and integrating it with my organic background.

I’d like to thank my lab mates who have developed one of the best academic working environments I have ever had the good fortune to work in. Thank you Deboleena Mitra, Dr. Michael Dcona, Dr. Railey White, Jonathan Sheldon, David Hacker, Brittany Danzig, Patrick Dupar, Mohammed Almohaini, and Luis Garcia. I’d like to thank Mark Davis for many
wonderful scientific discussions that helped me see my chemistry in different perspectives and for Edward F. Hawkins with editorial assistance. In addition, many thanks to Dr. Patrick ShiPing Zou and Dr. Kurt Hauser for the collaboration with the astrocyte cell studies.

I’ve had the opportunity to work in five academic environments, two industrial companies and have collaborated with government projects within a CRO facility. Far too often the competitive aspect of science is found newsworthy when the reality is far less dramatic. In every lab, in every research environment I have had the opportunity to be a part of, someone helped me, to find a reagent, to show me a technique, to stop and talk about science. These small gifts of time are additive and have helped me become the scientist I am today. There are far too many people to thank individually and instead I’d like to pay it forward, with my peers, with new undergraduate researchers, as a teacher’s assistant, as a tutor, or an employee, I will continue to pay these kindnesses I have experienced forward.

Finally, for all the support and comic relief my friends and neighbors have helped provide from San Diego, Los Angeles, to Richmond, I am forever grateful. I’d also like to credit those souls whom have passed but added greatly to my life, Jarrod B. Davidson and Brianna Millard Smith.
Table of Contents

Acknowledgements ........................................................................................................................................ iii

Table of Contents ........................................................................................................................................ vi

List of Tables ............................................................................................................................................... xvi

List of Charts .............................................................................................................................................. xviii

List of Figures ............................................................................................................................................... xix

List of Molecules ......................................................................................................................................... xxii

List of Schemes ........................................................................................................................................... xxiv

List ofAbbreviations and Symbols ............................................................................................................... xxvi

Abstract ..................................................................................................................................................... xxxiii

1 Synthesis of electrophilic fluorescent dyes ............................................................................................... 1

1.1 General overview on fluorescent dyes ................................................................................................. 1

1.1.1 History of fluorescence .................................................................................................................. 1

1.1.2 What is fluorescence? .................................................................................................................... 2

1.1.3 Types of fluorophores and desired properties .............................................................................. 3

1.1.4 Designing electrophilic fluorescent analogs and their components ............................................ 5

1.2 Synthesis of rhodamine based fluorescent probes ............................................................................ 8

1.2.1 Rhodamine B introduction ............................................................................................................... 8

1.2.1.1 Background synthesis of rhodamine B analogs ....................................................................... 9
1.2.2 Sulforhodamine 101 introduction .............................................................. 12
  1.2.2.1 Background synthesis of sulforhodamine analogs ................................. 14
1.3 Results and conclusions ............................................................................. 15
  1.3.1 Rhodamine B routes of synthesis ........................................................... 15
    1.3.1.1 Rhodamine B isothiocyanate route ..................................................... 15
    1.3.1.2 Rhodamine B piperazine route ............................................................ 18
      1.3.1.2.1 Rhodamine B sulfonate esters using the piperazine moiety ............. 19
      1.3.1.2.2 Rhodamine B piperazine propanol route ..................................... 20
  1.3.2 Sulforhodamine 101 routes of synthesis ................................................. 22
    1.3.2.1 Route of synthesis for tosylate analog sulforhodamine 101 (Molecule 1-15) using PEG linker ................................................................. 22
    1.3.2.2 Route of synthesis for sulforhodamine 101 sulfonate esters utilizing click chemistry ................................................................................................. 24
      1.3.2.2.1 Sulforhodamine 101 sulfonyl chloride as the starting material for access to sulfonate esters ................................................................. 24
      1.3.2.2.2 Azide Analog of sulforhodamine 101 (Molecule 1-17) ...................... 26
      1.3.2.2.3 Tosylate analog of sulforhodamine 101 (Molecule 1-19) ............... 32
      1.3.2.2.4 Meta-ethylamido sulfonate ester analog of sulforhodamine 101 (Molecule 1-22) ................................................................. 33
      1.3.2.2.5 Sulforhodamine-(PEO)₃-Azide as a versatile intermediate ............ 36
      1.3.2.2.6 NMR studies of Molecule 1-17 ....................................................... 39
1.4 Discussion ...................................................................................................... 43
1.5 Experimental methods ..........................................................47

1.5.1 Molecule 1-1: Rhodamine B ..................................................47
1.5.1.2 Molecule 1-2: Rhodamine B Isothiocyanate .......................48
1.5.1.3 Molecule 1-3 ..................................................................49
1.5.1.4 Molecule 1-4 ..................................................................50
1.5.1.5 Molecule 1-6 ..................................................................52
1.5.1.6 Molecule 1-7 ..................................................................54
1.5.1.7 Molecule 1-8 ..................................................................55
1.5.1.8 Molecule 1-9 ..................................................................57
1.5.1.9 Molecule 1-10 .................................................................58
1.5.1.10 Molecule 1-11 .................................................................59
1.5.1.11 Molecule 1-12 Sulforhodamine 101 (free acid) ..................60
1.5.1.12 Molecule 1-13 Sulforhodamine 101 Sulfonyl Chloride (Texas Red) ..................61
1.5.1.13 Molecule 1-14 .................................................................62
1.5.1.14 Molecule 1-15 .................................................................63
1.5.1.15 Molecule 1-16 .................................................................64
1.5.1.16 Molecule 1-17 .................................................................65
1.5.1.17 Molecule 1-18 .................................................................66
1.5.1.18 Molecule 1-19 .................................................................68
1.5.1.19 Molecule 1-20 .................................................................69
1.5.1.20 Molecule 1-21 .................................................................70
1.5.1.21 Molecule 1-22 .................................................................71
Covalent labeling of biomolecules of interest using fluorescent electrophiles ........................................... 76

2.1 Introduction .................................................................................................................................................. 76

2.1.1 Labeling proteins with electrophilic small molecules ........................................................................... 76

2.1.2 Proximity effect and affinity to label proteins of interest ........................................................................ 80

2.2 Overall rationale and objective for labeling proteins .................................................................................. 83

2.2.1 Fluorescence polarization theory ........................................................................................................... 84

2.3 Results ......................................................................................................................................................... 85

2.3.1 Labeling of EF-Tu-His Tag with Nickel NTA beads experiment ............................................................. 86

2.3.2 Evaluation of binding via fluorescence polarization ................................................................................. 88

2.3.2.1 FP of rhodamine B analogs with EF-Tu proteins containing peptidic inserts .............................. 89

2.3.2.2 FP of sulforhodamine 101 analogs with proteins containing peptidic inserts ............................ 92

2.3.2.2.1 FP of sulforhodamine 101 with EF-Tu proteins containing peptidic inserts .......................... 92

2.3.2.2.2 FP of sulforhodamine 101 tosylate with EF-Tu containing peptidic inserts .............................. 93

2.3.2.2.3 FP of Molecule 1-22 with proteins containing peptidic inserts .............................................. 95

2.3.2.2.4 FP of sulforhodamine 101 with peptides and no protein ....................................................... 95

2.3.3 Labeling assays ...................................................................................................................................... 98

2.3.3.1 Labeling with sulforhodamine 101 tosylate analog, Molecule 1-19 ........................................... 101

2.3.3.2 Labeling Proteins with Molecule 1-22 .......................................................................................... 105

2.4 Discussion of labeling results ................................................................................................................... 110
2.5 Experimental details and methods

2.5.1 Synthesis of 501/512 proteins (EF-Tu and GST)

2.5.1.1 Design and synthesis of the insert

2.5.1.2 Proteins chosen for genetic modification

2.5.1.3 Protein methods and expression

2.5.1.4 Determination of protein concentrations

2.5.2 Synthesis of 501 and 512 peptides

2.5.2.1 Molecule 2-1

2.5.2.2 Molecule 2-2

2.5.2.3 501/512 Peptide Experimental

2.5.3 Fluorescent electrophiles chosen for covalent labeling studies

2.5.4 Fluorescence polarization experiments

2.5.5 General labeling assays protocol

2.5.6 Beads experiment

3 Utilizing mRNA display as a tool to find fluorescent binding peptides

3.1 Introduction in vitro selection

3.1.1 Phage display

3.1.2 Ribosome display

3.1.3 mRNA display

3.2 Strategy using mRNA display

3.2.1 Overview for the identification of short peptidic sequence with affinity for rhodamine
3.2.2 mRNA display–library design.................................................................125
3.2.3 Library selection against a small molecule ..........................................128
3.2.4 Design and choice of electrophile ..........................................................129
3.3 Results and conclusions ........................................................................131
  3.3.1 Preparation of Molecule 3-3 .................................................................131
  3.3.2 8 Rounds of mRNA peptide selection with Molecule 3-3 ..................133
  3.3.3 Stability of electrophile in the selection reaction in the presence of imidazole ....139
3.4 Discussion ...............................................................................................141
  3.4.1 Future direction ...................................................................................142
    3.4.1.1 Selecting for biotin versus the electrophile ..................................142
    3.4.1.2 Future Directions if sequencing data is obtained .......................142
3.5 Experimental strategies ........................................................................143
  3.5.1 Synthesis of electrophilic biotin analogues .........................................143
    3.5.1.1 Molecule 3-1 .............................................................................144
    3.5.1.2 Molecule 3-2 .............................................................................145
    3.5.1.3 Molecule 3-3 .............................................................................146
    3.5.1.4 Molecule 3-4 .............................................................................148
  3.5.2 Preparation of mRNA fusions ...............................................................149
    3.5.2.1 Selection ....................................................................................150
  3.5.3 Electrophile stability study- imidazole reactivity with Molecule 3-3 .......150
4 Labeling astrocytes with fluorescent probes .............................................152
  4.1 Background: selective labeling of astroglia using sulforhodamine 101 ....152
4.2 Goals and objectives for astroglial staining .......................................................... 154

4.3 Astrocytes results and conclusions ......................................................................... 155

4.3.1 Imaging of young astrocytes ............................................................................... 155

4.3.2 Imaging of mature astrocytes .............................................................................. 155

4.3.2.1 Live cell imaging ............................................................................................. 155

4.3.2.1.1 Assessment of staining time needed for live cell imaging ....................... 157

4.3.2.1.2 Assessment of re-staining properties for live cell imaging ..................... 158

4.3.2.2 Astrocyte fluorescence post fixation ............................................................... 160

4.3.2.2.1 Wide field fluorescence imaging of astrocytes post fixation ............... 160

4.3.2.2.2 Confocal Imaging of astrocytes post fixation ......................................... 160

4.3.2.3 Live Cell Imaging overlap with fixed immunohistochemical staining of mature astrocytes. ............................................................................................................ 163

4.3.2.3.1 Molecule 1-22 ......................................................................................... 163

4.3.2.3.2 Molecule 1-12 ....................................................................................... 167

4.3.3 Imaging of mixed sample of oligodendrocytes and young astrocytes ............. 169

4.4 Discussion of astrocyte binding with Molecule 1-22 ............................................ 169

4.5 Astroglia experimental procedures ......................................................................... 170

4.5.1 Preparation of fluorophores ................................................................................. 170

4.5.2 Preparation of astrocytes ..................................................................................... 170

4.5.3 Incubation of astrocytes with fluorophore .......................................................... 171

4.5.4 Qualitative assessment of astrocytes using live cell imaging and fixated cells ..... 171

4.5.5 Assessment of fixated astrocytes using confocal imaging .................................. 172
4.5.6 Immunohistochemical staining ............................................................. 172

5 Overall Conclusions/ Discussion .................................................................. 173

List of references .......................................................................................... 178

List of references .......................................................................................... 179

Appendix A: NMR and MALDI Spectra ......................................................... 192

Molecule 1-1: Rhodamine B (RhB) ................................................................. 192
Molecule 1-2: Rhodamine B Isothiocyanate Starting Material ....................... 195
Molecule 1-3 ................................................................................................. 196
Molecule 1-4 ................................................................................................. 198
Molecule 1-5: Rhodamine B Base ................................................................. 200
Molecule 1-6 ................................................................................................. 203
Molecule 1-7 ................................................................................................. 207
Molecule 1-8 ................................................................................................. 209
Molecule 1-9 ................................................................................................. 212
Molecule 1-10 .............................................................................................. 214
Molecule 1-11 .............................................................................................. 216
Molecule 1-12: Sulforhodamine 101 (SR101) .............................................. 218
Molecule 1-13: Sulforhodamine 101 Sulfanyl Chloride (Texas Red) .......... 220
Molecule 1-14 .............................................................................................. 221
Molecule 1-15 .............................................................................................. 223
Molecule 1-17 .............................................................................................. 225
Molecule 1-18 .............................................................................................. 227
Molecule 1-19........................................................................................................229
Molecule 1-20........................................................................................................231
Molecule 1-21........................................................................................................232
Molecule 1-22........................................................................................................234
Molecule 1-23........................................................................................................237
Molecule 1-24........................................................................................................239
Molecule 2-1 501 Peptide ......................................................................................241
Molecule 2-2 512 Peptide ......................................................................................242
Molecule 3-1...........................................................................................................243
Molecule 3-2...........................................................................................................245
Molecule 3-3...........................................................................................................247
Molecule 3-4...........................................................................................................248

Appendix B: mRNA display supplemental information ........................................249

DNA template ........................................................................................................249

Transcription (DNA to mRNA).............................................................................249

mRNA purification via urea gel...........................................................................250

mRNA isolation from urea gel via electroelution..................................................250

Crosslink psoralen to mRNA ..............................................................................250

Translation (mRNA to peptide fusion library)......................................................252

Preparation of working area................................................................................252

Preparation of polymix 3X....................................................................................252

Preparation of master mix....................................................................................253
Translation of mRNA fusion reaction mixture ................................................................. 256
Purification of translation mixture using Oligo dT .......................................................... 257
Column preparation ........................................................................................................ 257
Binding of Oligo dT to mRNA fusion ............................................................................ 258
Elution of mRNA fusion from Oligo dT: .......................................................................... 258
Quantification of mRNA fusion yields ............................................................................ 259
Ethanol precipitation of purified mRNA fusion with glycogen ...................................... 260
Reverse transcription (mRNA to cDNA) .......................................................................... 261
Ni-NTA purification of RT reaction .................................................................................. 262
Ni-NTA buffers for purification of RT ............................................................................ 262
Selection .......................................................................................................................... 263
Preparation of Streptavidin magnetic particles/beads ...................................................... 263
Magnetic particles for pre-clearance studies .................................................................... 264
Reaction with Molecule 3-3 ............................................................................................ 265
Removal of residual small molecules via dialysis ............................................................ 265
mRNA Fusion isolation using Streptavidin beads. ............................................................ 266
PCR amplification of cDNA from selected mRNA fusion peptides ................................. 267
Small scale optimization ................................................................................................. 267
Large scale amplification ................................................................................................. 268
Purification of PCR product ............................................................................................ 268
Vita ..................................................................................................................................... 269
List of Tables

Table 1-1 Rhodamine B molecular states and physical properties$^{66,73}$ ........................................ 10
Table 1-2 Comparison of $^1$H NMR Shifts for Isomers of Molecule 1-17........................................ 41
Table 1-3 Correlation table of COSY for Molecule 1-17 Isomer A.............................................. 43
Table 2-1 Protein mutations containing 512 or 501 aptamers....................................................... 86
Table 2-2 Rhodamine analogs tested for affinity ................................................................................ 90
Table 2-3 Table Factor mix components for competitive labeling study using Molecule 1-19 .. 103
Table 2-4 Primers for overlapping PCR.......................................................................................... 111
Table 2-5 Fluorescent electrophiles tested for binding with proteins ............................................ 116
Table 3-1 Comparison of mRNA display versus other peptide libraries........................................ 122
Table 3-2 mRNA library oligonucleotide design$^{152,153}$ .................................................................. 127
Table 3-3 Summary of Molecule 3-3 selection results using mRNA Display ................................. 136
Table B-0-1 Transcription reaction conditions .................................................................................. 249
Table B-0-2 Psoralen crosslinking reagents ...................................................................................... 251
Table B-0-3 Polymix components ..................................................................................................... 253
Table B-0-4 Components of a translation master mix ........................................................................ 255
Table B-0-5 Translation components ................................................................................................. 257
Table B-0-6 Buffer recipes for oligo dT purification ......................................................................... 258
Table B-0-7 Typical elution results of mRNA fusion from Oligo dT .................................................. 260
Table B-0-8 Reverse transcription protocol ..........................................................261
Table B-0-9 Ni-NTA buffers .........................................................................................262
Table B-0-10 Magnetic beads capture limitations required dialysis of unreacted Molecule 3-3
...........................................................................................................................................266
List of Charts

Chart 2-1 Comparison of labeling techniques ................................................................. 79
Chart 2-2 Rhodamine B (Molecule 1-1) binding via FP ...................................................... 90
Chart 2-3 Rhodamine B (Molecule 1-1) FP values .............................................................. 91
Chart 2-4 Sulforhodamine 101 analogs affinity binding testing with proteins containing peptidic inserts and peptides without proteins ................................................................ 92
Chart 2-5 FP Data sulforhodamine 101 (SR101, Molecule 1-12) with proteins containing peptidic inserts .................................................................................................................. 93
Chart 2-6 FP sulforhodamine 101 Tosylate with proteins containing peptidic inserts ......... 94
Chart 2-7 FP Binding data for Molecule 1-22 with proteins containing peptidic inserts ........ 95
Chart 2-8 FP of 512 peptide with Molecule 1-12, sulforhodamine 101, SR101 ..................... 96
Chart 2-9 FP data of 501 peptide Molecule 1-22.................................................................... 96
Chart 2-10 FP buffer study of 512 peptide with Molecule 1-12, sulforhodamine 101.......... 97
Chart 2-11 FP study of 501 Peptide with Molecule 1-12, sulforhodamine 101.................... 98
Chart 3-1 Percent yield of captured peptide fusion following 8 rounds of selection against Molecule 3-3 ..................................................................................................................... 134
Chart B-0-1 Quantification of captured biotinylated peptides from the mRNA fusion ........ 267
List of Figures

Figure 1-1 Physical properties of fluorescence.................................................................2
Figure 1-2 Targeting a protein of interest with a fluorescent probe.................................5
Figure 1-3 Design of electrophilic fluorophores ................................................................8
Figure 1-4 commercially available rhodamine B dyes .......................................................10
Figure 1-5 Amino group modification to xanthene backbone .............................................12
Figure 1-6 Structure of sulforhodamine 101 .....................................................................13
Figure 1-7 Commercially available sulforhodamine 101 analogs .....................................14
Figure 1-8 Purification of Molecule 1-3 .............................................................................16
Figure 1-9 Reaction mixture for the formation of Molecule 1-4 ........................................17
Figure 1-10 MALDI-TOF MS reaction mixture characterization of Molecule 1-6 with bis-
substituted sulfonate esters ...............................................................................................21
Figure 1-11 Sulforhodamine 101 sulfonyl chloride formation (Molecules 1-13 and 1-16) ....26
Figure 1-12 Sulforhodamine 101 azide analog purification on chromatatron (Molecules 1-17 and
1-18) ...................................................................................................................................29
Figure 1-13 HPLC purification of sulforhodamine 101 azide analogs, Molecules 1-17 and 1-18 .30
Figure 1-14 TLC resolution of Molecules 1-17 and 1-18 using methanol in dichloromethane....31
Figure 1-15: TLC resolution of Molecules 1-17 and 1-18 comparing MeOH to IPA ..............31
Figure 1-16 TLC of Molecule 1-17 isomers ......................................................................39
Figure 1-17 1H NMR of Molecule 1-17 isomers .................................................................40
Figure 1-18 2D 1H-1H COSY of TR-(PEO)3-N3 Isomer A ..................................................42
Figure 2-1 Nickel-NTA captures of proteins with sulforhodamine 101 free acid (Molecule 1-12)

........................................................................................................................................88

Figure 2-2 Preliminary labeling studies used to validate covalent labeling of fluorescent
electrophiles with electrophilic fluorophore (Molecule 1-19) and non-electrophilic
fluorophore (control, Molecule 1-12)..................................................................................100

Figure 2-3 Labeling of proteins using Molecule 1-19 shows concentration dependence ....101

Figure 2-4 Time dependent labeling of proteins using Molecule 1-19..............................102

Figure 2-5 Competition labeling using Factor Mix with Molecule 1-19..............................104

Figure 2-6 Meta-ethylamido sulfonate ester analog of sulforhodamine 101 (Molecule 1-22) and
the hydrolyzed analog (Molecule 1-24)..............................................................................106

Figure 2-7 Meta-ethylamido sulfonate ester analog of sulforhodamine 101 (Molecule 1-22)
labeling studies with 512 EF-Tu and EF-Tu .......................................................................108

Figure 2-8 Meta-ethylamido sulfonate ester analog of sulforhodamine 101 (Molecule 1-22)
labeling of cell lysate ............................................................................................................110

Figure 2-9 PCR#1 to make primers ABC for four inserts (2% Agarose Gel, Hyperladder I)....112

Figure 2-10 PCR#2 recombinant plasmid (0.8% agarose gel, Hyperladder V)..................113

Figure 2-12 In gel evaluation of reaction between fluorophore and protein........................117

Figure 3-1 mRNA display covalent linkage ..........................................................................128

Figure 3-2 Round 8i PCR amplification ................................................................................135

Figure 3-3 Controls used for mRNA peptide fusion amplification and selection...............137

Figure 3-4 PCR Optimization of concentration of electrophile exposed to mRNA fusion on
magnetic particles ..................................................................................................................138
Figure 3-5 Optimization of volume of beads needed for PCR amplification ..............................139
Figure 3-6 Imidazole testing on electrophile stability. ........................................................................................................140
Figure 3-7 Oligonucleotide sequence of the template$^{157}$ .............................................................149
Figure 4-1 Mature astroglia staining with 2 hour incubation time..........................................................157
Figure 4-2 Re-staining astroglia with fluorescent probes..............................................................................158
Figure 4-3 Confocal imaging of fixed astrocytes with Molecule 1-12 and Molecule 1-22........162
Figure 4-4 Molecule 1-22 live cell and fixed cell imaging of stained astrocytes first example...164
Figure 4-5 Molecule 1-22 Live cell and fixed cell imaging of stained astrocytes second example
........................................................................................................................................................................166
Figure 4-6 Molecule 1-12 Contrast of live cell imaging versus post immunohistochemical
staining first example ........................................................................................................................................167
Figure 4-7 Molecule 1-12 Contrast of live cell imaging versus post immunohistochemical
staining second example ........................................................................................................................................168
Figure B-0-1 Random DNA sequence ................................................................................................................249
Figure B-0-2 Psoralen crosslinking to mRNA .................................................................................................251
Figure B-0-3 Random CX12 mRNA sequence: (123 mar) .............................................................................251
Figure B-0-4 Streptavidin magnetic particles$^{180}$ .........................................................................................264
List of Molecules

Molecule 1-1 Rhodamine B (RhB).........................................................................................9
Molecule 1-2 Rhodamine B Isothiocyanate............................................................................48
Molecule 1-3..........................................................................................................................49
Molecule 1-4..........................................................................................................................50
Molecule 1-5 Rhodamine B Base$^8$ ...................................................................................51
Molecule 1-6..........................................................................................................................52
Molecule 1-7..........................................................................................................................54
Molecule 1-8..........................................................................................................................55
Molecule 1-9..........................................................................................................................57
Molecule 1-10.........................................................................................................................58
Molecule 1-11.........................................................................................................................59
Molecule 1-12 Sulforhodamine 101 Free Acid .....................................................................60
Molecule 1-13 Sulforhodamine 101 Sulfonyl Chloride (Texas Red) ....................................61
Molecule 1-14.........................................................................................................................62
Molecule 1-15.........................................................................................................................63
Molecule 1-16.........................................................................................................................64
Molecule 1-17.........................................................................................................................65
Molecule 1-18.........................................................................................................................66
<table>
<thead>
<tr>
<th>Molecule</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-19</td>
<td>68</td>
</tr>
<tr>
<td>20: 3-(ethylcarbamoyl)benzene-1-sulfonyl chloride</td>
<td>69</td>
</tr>
<tr>
<td>21: But-3-yn-1-yl-3-(ethylcarbamoyl)benzene sulfonate</td>
<td>70</td>
</tr>
<tr>
<td>22</td>
<td>71</td>
</tr>
<tr>
<td>23</td>
<td>73</td>
</tr>
<tr>
<td>24</td>
<td>74</td>
</tr>
<tr>
<td>1 501 Peptide</td>
<td>114</td>
</tr>
<tr>
<td>2 512 Peptide</td>
<td>114</td>
</tr>
<tr>
<td>3-1</td>
<td>144</td>
</tr>
<tr>
<td>3-2</td>
<td>145</td>
</tr>
<tr>
<td>3-3</td>
<td>146</td>
</tr>
<tr>
<td>3-4</td>
<td>148</td>
</tr>
</tbody>
</table>
List of Schemes

Scheme 1-1 Utilizing a small molecule to label a protein of interest (POI) ........................................6
Scheme 1-2 Rhodamine analogs as chemodosimeters .................................................................11
Scheme 1-3 Rhodamine B Route 1: Isothiocyanate synthetic scheme ...........................................15
Scheme 1-4 Route of synthesis for Molecule 1-6, a rhodamine B piperazine analog ......................18
Scheme 1-5 Route of synthesis for Molecule 1-8 using a piperazine nucleophile with bis-tosylate triethylene glycol. .........................................................................................................................19
Scheme 1-6 Route of synthesis for rhodamine B electrophiles using RhB piperazine and bis-
substituted electrophiles ...................................................................................................................20
Scheme 1-7 Route of synthesis for Molecule 1-10 ........................................................................20
Scheme 1-8 Route of synthesis for Molecule 1-11 .........................................................................22
Scheme 1-9 Route of synthesis for tosylate analog of sulforhodamine 101 using PEG linker
(Molecule 1-15) ...........................................................................................................................23
Scheme 1-10 Formation of sulforhodamine 101 sulfonyl chloride (Molecules 1-13 and 1-16) ...25
Scheme 1-11 Route of synthesis for 1-Amino-11-azido-3,6,9-trioxaundecane (Molecule 1-7) ...27
Scheme 1-12 Route of synthesis for azide analog of sulforhodamine 101 (Molecule 1-17) .......28
Scheme 1-13 Route of synthesis for Tosylate analog of sulforhodamine 101 (Molecule 1-19)
using click chemistry ....................................................................................................................33
Scheme 1-14 Route of synthesis for Molecule 1-22 .....................................................................35
Scheme 1-15 Molecule 1-17 as a useful intermediate ..........................................................36
Scheme 1-16 Biotinylated sulforhodamine 101 (Molecule 1-23).............................................38
Scheme 1-17 Molecule 1-17 as a useful intermediate for sulforhodamine 101 analogs ..........46
Scheme 2-1 Desired elements in small synthetic chemical probes.......................................78
Scheme 2-2 LDT proximity based labeling............................................................................81
Scheme 2-3 Labeling EF-Tu with a fluorescent tosylate.......................................................83
Scheme 3-1 Generalized schematic of mRNA display (151)................................................125
Scheme 3-2 Selection process utilized in mRNA display......................................................129
Scheme 3-3: Preparation of biotinylated electrophiles.........................................................132
List of Abbreviations and Symbols

10XRTBuff: 10X Reverse transcriptase buffer

T: TLC indicates not calculated as solvent front has moved too high.

Ac: Acetyl

ACN: acetonitrile, CH3CN

anhyd: anhydrous

Ar: reaction mixture under inert gas, argon

BG: background

Bn, Bzl: Benzyl

bp: boiling point, base pair

br: broad, spectral

Bu: Butyl

t-Bu: t-butyl

°C: degrees Celsius

CC: flash column chromatography

cDNA: complementary DNA

CaH: Calcium hydride, drying reagent

CD: Circular dichroic

cDNA: complementary deoxyribonucleic acid
calcd: calculated

CHCA: α-cyano-4-hydroxycinnamic acid

CHCl₃: chloroform

compd, cmpd: compound

conc: concentration

COSY: correlation spectroscopy

δ: Chemical shift in parts per million downfield from tetramethylsilane

d: density, days, doublet

DAPI: 4’6-diamidino-2-phenylindole fluorescent stain

ddH₂O: double distilled, deionized water

DIEA: diisopropylethylamine

DCM: dichloromethane

DEPT: distortion less enhancement by polarization transfer

DIEA: N,N-Diisopropylethylamine

DMAP: 4-(N,N-dimethylamino)pyridine

DMF: N,N-dimethylformamide

DMSO: dimethyl sulfoxide

DMSO-d₆: deuterated dimethyl sulfoxide

DNA: deoxyribonucleic acid

dNTP: deoxyribonucleotide

ds: double stranded

DTT: dithiothreitol
E: elution

EDC: 1-ethyl-3-(30dimethylaminopropyl)carbodiimide

EDTA: ethylenediaminetetraacetic acid

El: Electrophile

eq, eqn: equation

eq, Equiv: equivalent

Et: ethyl

EtAm: Molecule 1-22

EtOH: Ethanol

FT: flow through, Fourier transform

g: gram(s)

GFAP- glial fibrillary acidic protein

GTP: guano sine 5′-triphosphate

h: hours

HPLC: high performance liquid chromatography

Hz: hertz

I₂: iodine used for TLC

IHC: Immunohistochemical referring to immunostaining

IR: Infrared

J: coupling constant in NMR spectroscopy

k: kilo

l: liter
lit.: literature value

µ: micro

m: master mix (mRNA), multiplet (spectroscopy), meter, milli-

M: molar

M+: parent molecular ion

MALDI: matrix assisted laser desorption

Me: methyl

MeOH: Methanol

MgSO4: Magnesium sulfate, drying reagent

min: minute(s), minimum

mL: milliliter(s)

mM: millimolar (concentration)

mmol: millimole(s)

mol: mole(s)

mol. wt: molecular weight

mp: melting point

mRNA: messenger ribonucleic acid

MS: mass spectrometry

MW, mol wt: molecular weight

N: normal

nm: nanometer(s)

NMR: nuclear magnetic resonance
nu: nucleophile

Obs: observed (MALDI-TOF)

P, Prod: Product

PCR: Polymerase chain reaction

Ph: Phenyl

Pr: propyl

iPr: isopropyl

pyr: pyridine

POI: Protein of Interest

pTLC: preparatory scale thin-layer chromatography

q: quarter (spectral)

Rd: Round

rel: relative

Rf: retention factor (chromatography)

RNA: ribonucleic aid

RM: Reaction Mixture

RT: reverse transcription

rt: room temperature

rxn: Reaction

s: singlet (spectral)

SA: Streptavidin

SiO2: silica gel for chromatography
SLM: Small Labeling Molecule
SM: Starting Material
$S_{N2}$: bimolecular nucleophilic substitution
SR: sulforhodamine
SR101: sulforhodamine 101, Molecule 1-12
ss: single stranded
t: triplet (spectral), time
TCEP: Tris(2-carboxy-ethyl)phosphine hydrochloride
TEA: triethylamine
THF: tetrahydrofuran
temp: temperature
TFA: trifluoroacetic acid
TMS: tetramethylsilane (spectral)
TOF: time of flight (mass spectrometry)
TR: Texas Red/ sulforhodamine 101 sulfonyl chloride
TR Filter: Texas Red fluorescence imaging filter for microscopy
TRIS: tris(hydroxymethyl)aminomethane
TLC: thin-layer Chromatography
Ts: *para*-toluenesulfonyl (tosyl)
UV: ultra-violet
Van: Vanillin TLC Stain
vis: visible
vol: volume

v/v: volume to volume ratio, volume concentration

w: wash

wt: weight

w/v: weight to volume ratio, mass concentration

u or U: units (typically for units/µL of enzyme)
Abstract

INVESTIGATIONS INTO THE FLUORESCENT COVALENT LABELING OF BIOMOLECULES UTILIZING RHODAMINE DYES, ELECTROPHILIC LEAVING GROUPS AND MRNA DISPLAY.

By Susan Daniela Selaya

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, 2014.

Major Director: Matthew C.T. Hartman, Associate Professor of Chemistry and Massey Cancer Center

The discovery of a method by which proteins of interest can selectively be labeled with a probe of choice intracellularly is a longstanding goal in chemical biology research. Conventional labeling techniques have utilized large domain tags but despite the development of small labeling molecules there have been no short peptide sequences known to covalently label a small molecule without the aid of an enzymatic process or metal chelation. We aimed to find a sequence of nucleophilic peptides that reacted covalently and specifically with electrophilic small labeling molecules using mRNA display. Our goal was to show that an electrophilic small labeling molecule that is brought in proximal distance to a protein of interest via affinity can result in nucleophilic attack by a neighboring nucleophilic amino acid to
covalently label the protein of interest. Utilizing affinity between a small labeling molecule and a protein of interest to bring them spatially close to one another maximizes the chance that a covalent reaction can take place and provides selectivity between two components in a complex mixture. Towards this goal, we developed several electrophilic fluorescent small molecules. Covalent labeling was achieved using electrophilic bait in the form of sulfonate esters, a polyethylene oxide linker provided structural flexibility, and a fluorescent affinity tag containing a rhodamine backbone served as the potential binding site to a key peptide sequence encoded within a protein of interest. The synthetic routes to access our electrophilic rhodamine B and sulforhodamine 101 fluorophores were optimized. Key intermediates were produced and served as flexible points of modification to make various analogs of our desired electrophilic fluorophores. The affinity between proteins containing the peptide sequence and the fluorescent electrophiles were determined by fluorescence polarization. Covalent labeling was determined to be both time and concentration dependent. The expected published affinity between the peptides and fluorophore was not high enough to produce selective labeling. However, our small labeling molecules were found to be effective at labeling various proteins in vitro. In addition, our electrophilic fluorophores have been found superior to sulforhodamine 101 in live cell imaging of astrocytes.
Synthesis of electrophilic fluorescent dyes

This chapter reviews the design and synthesis of electrophilic fluorescent dyes that were prepared for the purposes of labeling biomolecules of interest.

1.1 General overview on fluorescent dyes

1.1.1 History of fluorescence

The Greek term phosphor translates to the bearer of light\(^\text{17}\), and is a term that has been used since the middle ages. Many observations of “cold light” (compounds that emit light without heating or fire) have been reported throughout history and this phenomenon is now known as incandescence. Some of these reports of glowing compounds included wood Lignum nephriticum (Monardes, 1565) impure barium sulfate in stone (Cascariolo, 1603 and Liceti, 1640), fluorspar crystals and chlorophyll solutions (Brewster, 1833). It wasn’t until 1852 when Sir George Gabriel Stokes published his studies “On the Refrangibility of Light”\(^\text{18}\) that considered the theory that the wavelengths of dispersed light might be different/longer than that of the original wavelength absorbed. Sir Stokes called this phenomenon “dispersive reflection” which he later coined fluorescence\(^\text{19}\) and his conclusions led to what is known as Stokes shift, Figure 0-1b.
Nobel laureate Antoine Henri Becquerel had also published work on phosphorescence of calcium sulfide a decade earlier and the principles behind fluorescence and phosphorescence began to take shape\textsuperscript{20}. The work by Becquerel and Stokes was described as cold light because unlike incandescence, these emissions do not require high heat or produce heat. As the field of quantum theory evolved it became evident that energy states of electrons are quantized and this provided a better understanding of luminescence. There are multiple types of luminescence and they are defined by their mode of excitation. For example, chemiluminescence is a mode of excitation caused by a chemical process such as oxidation while photoluminescence, which includes fluorescence and phosphorescence, is caused by the absorption of photons (light)\textsuperscript{21}.

1.1.2 What is fluorescence?

When an electron in a molecule is excited by the absorption of a photon to a higher energy state it can return to its ground state via multiple pathways including electron transfer, proton transfer, energy transfer, photochemical transformation\textsuperscript{22}, etc. Fluorescence emission
occur one of two ways. An electron remains in a low energy/ground state, called $S_0$ until it is excited via energy in the term of a photon into higher states $S^*$(S1, S2, etc.). The process by which this electron returns to its ground state can occur by releasing heat (radiative) or non-radiative. If the dissipation of energy occurs in the form of photons of light without intersystem crossing then fluorescence is observed. If it occurs with intersystem crossing then the photons release are typically lower in energy with a longer wavelength and called phosphorescence or undergo delayed fluorescence, Figure 0-1a.

The process by which fluorescence occurs can be an efficient or inefficient process and is a physical property for each molecule. The efficiency of this process is defined as fluorescence quantum yield and defined as the ratio between the numbers of photons emitted versus the numbers of photons released$^{23}$. For example, the quantum yield of bacteriorhodopsin at optimal temperature and pH is 0.64 meaning that for every 10 photons absorbed 6-7 are emitted$^{24}$.

1.1.3 Types of fluorophores and desired properties

Fluorophores are molecules that are fluorescent and these are typically subdivided into two classes: organic and inorganic$^{25}$. There is a plethora of organic fluorophores available commercially and through synthesis. Generally, organic fluorophores have extensive conjugation with a high level of rigidity and heteroatoms. Biological sources of fluorophores also called intrinsic fluorophores are naturally occurring and include collagen, NADH, flavins etc$^{26}$. In addition, fluorescence of inorganic chemistry in the studies of lanthanides$^{27,28}$ and quantum dots has been extensively explored recently$^{29,30}$. Small molecule fluorescent probes and their classifications, spectral characteristics, advantages and weaknesses are reviewed
extensively by Lavis and Raines in an ACS Chemical Biology review article titled Bright Ideas for Chemical Biology\textsuperscript{31}.

The desired properties of a fluorophore generally include excitation and emission in a region of the spectrum that is useful for detection (background is worse in the UV area). Many fluorophores have been designed with fluorescence above 600 nM in the near infrared region for bioanalytical purposes\textsuperscript{32} as the majority of common fluorophores have an emission maximum less than 600 nm\textsuperscript{33}. Fluorescent dyes that absorb above 600 nm in the red region are particularly desired because there is a reduction in background signal and this maximizes sensitivity for measurements\textsuperscript{32}. It is also important that the fluorophores do not self-quench, have a high quantum yield and readily absorb light. For biological purposes it is especially important that the fluorophore is non-toxic, has low non-specific binding, small in size and low in cost. Finally, photo stability is also important as the phenomena of photobleaching\textsuperscript{34} can greatly decrease the utility of certain fluorophores. For example, fluorescein is especially susceptible to photobleaching despite its other great physical properties\textsuperscript{35}.

Finally, the size of the fluorophore has lately come into question in terms of applicability in physiological conditions. Live cell imaging has been greatly enhanced by the development of selective labeling of proteins both \textit{ex vivo} and \textit{in vitro}\textsuperscript{36–38}. Genetically introduced fluorescence first became available with green fluorescent proteins for which the Nobel prize in chemistry was awarded in 2008\textsuperscript{39} and the concept has been used for a large variety of biological studies\textsuperscript{4,5,40–44}, both \textit{ex vivo} and \textit{in vitro}. A variety of other engineered fluorescent proteins have become available with spectral properties that have been modulated and are available in a range of colors. Applications of these genetically introduced fluorescent proteins have been
limited by their size and spectral range\textsuperscript{9}. To address these limitations a variety of small synthetic fluorescent probes have been developed that can be incorporated into their target biomolecule\textsuperscript{9,45} (further discussed in Ch.2 introduction, Figure 0-2).

**Figure 0-2 Targeting a protein of interest with a fluorescent probe**

1.1.4 **Designing electrophilic fluorescent analogs and their components**

One methodology used to label proteins has been to utilize a proximity effect whereby a component of a small labeling molecule (SLM) has a high affinity for a region of the protein of interest (POI), Scheme 0-1. Our goal was to show that an electrophilic SLM that is brought in proximal distance to a protein via affinity could result in nucleophilic attack by a neighboring nucleophilic amino acid to covalently label the protein of interest. Utilizing affinity between an SLM and a protein of interest (POI) to bring them spatially close to one another maximizes the chance that a covalent reaction can take place and maximizes selectivity between the SLM and the POI in in a complex mixture. Various methods have been established that utilize a proximity affect to study protein-protein interactions and have been utilized in screening assays with biotin for identification purposes\textsuperscript{46}. These concepts are elaborated on section 2.1.
To utilize the proximity effect the SLM we designed needed to contain a probe that would bind tightly to our POI. We chose to take advantage of reported binding affinities between specific peptidic aptamers and various fluorescent dyes discovered using phage display by Nolan and coworkers\textsuperscript{14,15,47}. The aptamers they found using in vitro selection were found to have an affinity toward the xanthene backbone of certain rhodamine analogs. More specifically, they found picomolar affinity between a 38-mer aptamer called TR512 and a 48-mer aptamer called TR501 with Texas Red and rhodamine red. While Texas Red contains four outer conjugated rings and rhodamine red does not, Nolan and coworkers tested binding to compare the two and found that TR512 effectively bound to both and TR501 bound only to the Texas red. They also tested various groups on the aromatic ring and did not detect a change in binding. They concluded that the xanthene backbone was the source of binding between the peptides and the rhodamine analogs. Based on this evidence we chose to test a variant of rhodamine with an electrophilic small molecule that we designed containing an excellent leaving group.

Based on the work of Cravatt and co-workers\textsuperscript{48} and Hamachi and co-workers\textsuperscript{13}, tosylates were chosen as the electrophile. Tosylate esters have been used as a model electrophile because they are stable in biological media, the corresponding leaving group is not
cytotoxic, and sulfonate esters can be tuned in terms of reactivity\textsuperscript{49}. The tosylate ester as a reactive group is typically made from a tosyl chloride in the presence of a sterically hindered non-nucleophilic base to form the sulfonate ester. While sulfonyl halides readily hydrolyze in aqueous environments, sulfonate esters are far more stable. The hydrolysis rates of two ligand directed tosylates (LDT) by the Hamachi lab were determined and they found that only 10\% of their sulfonate ester starting material was hydrolyzed after incubation at 37 °C for 48 hours\textsuperscript{50}.

Cravatt and co-workers tested a variety of electrophiles as their choice of chemical probes for activity-based protein profiling (ABPP) and this was used to profile the functionality of various enzymes in their native proteomes. While iodoacetamide and maleimide have been well characterized in terms of their biological reactivity\textsuperscript{51} Cravatt and co-workers sought to expand the variety of electrophiles that could be utilized for ABPB\textsuperscript{52}. Their application of various electrophiles including a phenyl sulfonate ester to a mouse proteome whose reactive enzymes were captured with an azide reporter tag then sequenced showed that there are disparate reactivities between the tested electrophiles with widely divergent amino acid preferences in proteomes. In addition, they concluded that the phenyl sulfonate ester is a versatile electrophile and recommended its application in ligand guided protein surface labeling\textsuperscript{53} while utilizing the other more specific electrophiles for testing cysteine functionality of specific enzymes that require this amino acid for functionality and/or post translational modification.

The electrophile that we developed for our research studies contained three main components: rhodamine B or sulforhodamine B, a flexible polyethylene oxide linker and a highly reactive sulfonate ester (Figure 0-3). In the presence of a nucleophilic amino acid, an
$S_N2$- type reaction produces a covalent bond between the reactive amino acid proximal to the insert known to have a high affinity with the fluorescent portion of the molecule. The flexible linker is important to provide spatial distance between the fluorescent portion of the molecule and the leaving group as this distance provides spatial flexibility to optimize exposure to potential neighboring nucleophiles on the protein to covalently react.

*Figure 0-3 Design of electrophilic fluorophores*

1.2 Synthesis of rhodamine based fluorescent probes

1.2.1 Rhodamine B introduction

The rhodamine family of dyes contain a number structurally related molecules observed to have high quantum yields$^{54}$ and were widely used for industrial applications including septic system tracking$^{55}$ and in the cosmetics industry$^{56}$ until they were discovered carcinogenic$^{57,58}$. In terms of biological application$^{31}$, rhodamine analogs are used in fluorescence microscopy, flow cytometry, ELISA purposes and now in an expanding field of chemodosimetry in analytical chemistry used to measure copper$^{59}$, mercury$^{60}$ and other ions in solution. Rhodamine B (Molecule 0-1) has a quantum yield of 0.65 and a fluorescence emission of 568 nM in basic
ethanol\textsuperscript{54}. The properties that make this particular class of dyes particularly attractive include low cost, high absorption coefficient, high sensitivity in the red region of the UV spectrum, low photo-bleaching\textsuperscript{61}, and high quantum yield\textsuperscript{62}.

**Molecule 0-1 Rhodamine B (RhB)**

\[
\begin{align*}
\text{Et}_2\text{N} & \quad \text{O} \\
\text{Et}_2\text{N} & \quad \text{O} \\
\text{Et}_2\text{N} & \quad + \text{NET}_2
\end{align*}
\]

**1.2.1.1 Background synthesis of rhodamine B analogs**

There are various rhodamine derivatives commercially available used for labeling and imaging purposes including FRET applications. These molecules contain various functional groups including isothiocyanates (RBTC, TRITC), sulfonyl chlorides (Texas Red), carboxytetramethylrhodamine (TAMRA), and NHS ester variants that are reactive towards a desired functional group on proteins, cell surfaces, etc. In terms of rhodamine analogs, the most useful in terms of synthetic pathways for our purposes was rhodamine B isothiocyanate\textsuperscript{63} (Molecule 1-2), Figure 0-4. Isothiocyanates readily react with amines to form thioureas typically in high yield\textsuperscript{64,65}. 
Recent variants to rhodamine dyes include Alexa Fluor®66 dyes and Dylights®67 that provide color options, enhanced brightness and decreased photobleaching.

There are key structural configurations of rhodamine B that must be considered when addressing possible routes of synthesis for analogs. All rhodamine analogs contain a xanthene backbone and this conjugated system is extended with nitrogen atoms alpha to the xanthene chromophore. The carboxylic acid of rhodamine B is protonated in acidic solvents and can be converted into a zwitterion in basic solutions that leads to a shift in both absorption and fluorescence maxima68. These changes in physical properties are also observed for the quantum yield and lifetime of these charge states, Table 0-1.

**Table 0-1 Rhodamine B molecular states and physical properties**62,69

<table>
<thead>
<tr>
<th></th>
<th>Cationic</th>
<th>Zwitterionic</th>
<th>Lactone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max Wavelength (nm)</td>
<td>553</td>
<td>543</td>
<td>311</td>
</tr>
<tr>
<td>Solvent Tested</td>
<td>Ethanol</td>
<td>Ethanol</td>
<td>Diethylether</td>
</tr>
<tr>
<td>Molar Absorptivity/10^5 (M-1 cm^-1)</td>
<td>1.17</td>
<td>1.11</td>
<td>0.16</td>
</tr>
<tr>
<td>Quantum Yield</td>
<td>0.53</td>
<td>0.7</td>
<td>0.022</td>
</tr>
</tbody>
</table>
In addition, the carboxylate of the zwitterion can react intramolecularly to form a spirolactam (Scheme 0-2) and this interruption in conjugation leads to a near complete loss of fluorescence. Utilizing the interruption in fluorescence has been extensively studied in the field of chemodosimetry as sensors for various ions, Kim and coworkers provide a review on this topic. When the acid of rhodamine B is functionalized as an amide, the nucleophilic nitrogen of the amide can intra-molecularly attack the ring system creating a lactam, Scheme 0-2. In the cyclized lactam form, the rhodamine fluorophore is not longer fluorescent and this property is utilized for chemodosimetry studies.

**Scheme 0-2 Rhodamine analogs as chemodosimeters**

Another point of modification for the xanthene backbone of rhodamine analogs includes modification of the ethylamine moieties, Figure 0-5. However, these heteroatoms play an important role in the fluorescence properties of the molecule and modification of the ethyl amines with cleavable linkers leads to a remarkable loss of fluorescence. A loss of fluorescence was not a desired goal to institute in a biological system and we considered other points of modification on rhodamine B to add our electrophilic leaving group.
To avoid loss in fluorescence most commercially available rhodamine analogs contain a modification of the carboxyphenyl ring and are typically sold as mixed isomers of rhodamine B. The carboxyphenyl group is almost perpendicular to the xanthene backbone and is not part of the chromophore system which permits modification at that site\textsuperscript{32}. Commercially, these analogs are expensive\textsuperscript{62} and while cost-effective for labeling purposes that typically require sub milligram masses, the cost is economically restrictive for the purposes of further derivatization.

### 1.2.2 Sulforhodamine 101 introduction

Sulforhodamine 101 (SR101, Molecule 1-12) is a rhodamine analog substituted with two sulfonic acids in the ortho and para positions relative to the xanthene ring system, Figure 0-6. Sulforhodamine 101 features two julolidines on the xanthene ring system that rigidify the backbone of sulforhodamine 101 and are credited for providing the intense absorption and shift to the red region of the UV.
There are various analogs of sulforhodamine 101 that are available commercially, Figure 1-7. While sulforhodamine 101 is preferable in terms of solubility to other analogs such as the isothiocyanate and the EDC/maleimide derivatives, all of these derivatives are costly, in part due to their short shelf lives and high reactivity as they readily hydrolyze in aqueous environments. In addition, their conjugation to proteins is optimized at pH 9-9.5\(^7\) that limits its scope of usefulness. In terms of selectivity, Texas Red reacts with other amino acids such as tyrosine, serine and histidine residues whose resulting structures are less stable than the desired sulfonamide. Finally, due to the high reactivity and storage difficulty, sulforhodamine 101 acid chloride is fairly expensive and is available for $10-25/mg. Despite its expense the resulting reactivity can vary from batch to batch due to hydrolysis during shipping and storage from small amounts of moisture. Sulforhodamine 101 free acid is not expensive but the analogs can be economically restrictive as intermediates for further derivatization.
The best known analog of sulforhodamine is Texas Red®, a trademark name by Molecular Probes, Inc. that is the sulfonyl chloride derivative and henceforth described as sulforhodamine 101 sulfonyl chloride. The most commonly used biological application for sulforhodamine 101 is its use as a dye for astrocytes, a type of glial cell. This is further discussed in Section 2.1.4.

1.2.2.1 Background synthesis of sulforhodamine analogs

The synthetic pathways to producing sulforhodamine analogs are similar to rhodamine B in that modifications to the xanthene core interrupt fluorescence. For example, asymmetrical sulforhodamine dyes whereby one of the julolidine rings were modified and have been developed as a quencher dyes.

Instead of a carboxylic acid ortho on the phenyl ring to the xanthene on the rhodamine B molecule, sulforhodamine 101 contains two sulfonic acids and these are not nucleophilic enough to undergo competitive cyclization to create a sultam. The sulfonic acids on sulforhodamine 101 can be activated by functional group inter-conversion to the sulfonyl
chlorides that are highly reactive with amines to form thioureas and this reaction is chemoselective for reaction with amines over alcohols to form sulfonate esters. Therefore, instead of creating two sulfonate esters which would have produced a molecule with competing electrophilic leaving groups we chose to functionalize the sulfonic acids to the sulfonyl chloride and then react it with an amine to give a stable thiourea.

1.3 Results and conclusions

1.3.1 Rhodamine B routes of synthesis

1.3.1.1 Rhodamine B isothiocyanate route

Initially, we began the route of synthesis with rhodamine B isothiocyanate (Molecule 1-2) and reacted it with (PEO)$_3$-monoamine to make the thiourea (Molecule 1-3), Scheme 1-3. The PEG linker provides a nucleophilic alcohol that reacts with $para$-toluenesulfonyl chloride to give the tosylate, Molecule 1-4.

Scheme 0-3 Rhodamine B Route 1: Isothiocyanate synthetic scheme
As a mixture of isomers, purification via column chromatography of the desired products proved difficult. Molecule 1-3 was isolated using flash chromatography as can be seen in Figure 0-8. The TLC plates in panel of Figure 1-8 show the fractions collected during flash chromatography used to purify Molecule 1-3 and the desired product was isolated as isomers in fraction 17-32. The TLC in panel b shows the starting material (Molecule 1-2) in the left lane and this is compared to the purity of the product (Molecule 1-3) in the right lane. Characterization of Molecule 1-2 using TLC and MALDI-TOF suggested to us that our product was pure and we used this material for the subsequent tosylation reaction.

**Figure 0-8 Purification of Molecule 1-3**

a) Purification of Molecule 1-3 via column chromatography.

b) Final product

Despite multiple attempts to optimize the produce the formation of the tosylate (Molecule 1-4) only trace amounts were observed and the resulting product was difficult to characterize by NMR. Monitoring the reaction by TLC indicated a disappearance of the starting
material to give a product but MALDI-TOF confirmed only a small formation of the product by mass.

**Figure 0-9 Reaction mixture for the formation of Molecule 1-4**

TLC representation of the reaction mixture to make Molecule 1-4 RbB(PEO)$_3$OTs in 1:4 MeOH:CHCl$_3$. Left TLC lane represents the RM, crude reaction mixture. Middle lane represents a co-spot of the RM with the SM. Right lane represents the starting material, Molecule 1-3.

Methods to optimize the yields included varying the reaction time, increasing the concentration of the sulfonyl chloride and testing various sterically hindered bases to push the substitution reaction forward. In addition, a different electrophile, benzenesulfonyl chloride in lieu of the $p$-toluenesulfonyl chloride, was tested at various reaction times and concentrations.

A variety of TLC conditions were tested to monitor the reaction. The Rf of the major product indicates a product and this product was isolated and characterized by both MALDI-TOF and NMR. While this material was generated four times, a variety of tosylation conditions were attempted and a pure product was never isolated. Based on these results it was decided to attempt another route of synthesis towards the goal of a rhodamine B electrophile.
1.3.1.2 Rhodamine B piperazine route

A more cost effective approach to form the rhodamine tosylate utilized a functional
group inter-conversion between the carboxylate of the carboxyphenyl ring to an amide and this
was achieved via a reactive spirolactam intermediate (spirolactam) and reaction with
piperazine\textsuperscript{80} to rhodamine B functionalized with a piperazine moiety (Molecule 1-6), Scheme
0-4.

Scheme 0-4 Route of synthesis for Molecule 1-6, a rhodamine B piperazine analog.

Molecule 1-6 was straightforward to make in moderate yields. It also permitted access
to start the route of synthesis with rhodamine B (Molecule 1-1), an inexpensive commercially
available starting material. The carboxylic acid of the phenyl ring of rhodamine B reacts in basic
conditions to give a lactone (Molecule 1-5) resulting in a non-fluorescent product. The lactone
was then reacted with trimethyl aluminum and piperazine to give Molecule 1-6. The piperazine
moiety in Molecule 1-6 is desirable because it produces a tertiary amide and the lack of a
proton on the amide prevents lactam formation. In addition, the secondary amine of the
piperazine provides a handle for further derivatization as it can be reacted with a variety of
electrophiles. Our desired goal was to produce a rhodamine B electrophile.
1.3.1.2.1 Rhodamine B sulfonate esters using the piperazine moiety

There were two approaches taken to using the piperazine amide intermediate (Molecule 1-6). The first route (Scheme 1-5) attempted to use a bis substituted electrophilic polyethylene oxide analog to give Molecule 1-8. While we were able to detect the product on MALDI-TOF, we were only able to recover trace amounts post purification via column chromatography and HPLC purification.

**Scheme 0-5 Route of synthesis for Molecule 1-8 using a piperazine nucleophile with bis-tosylate triethylene glycol.**

![Scheme 0-5](image)

In addition to the tosylate formed from (Molecule 1-7) we also tried to react it with the phenyl, nosyl, mesyl and p-methoxysulfonyl chloride analogs to make the corresponding sulfonate esters, Scheme 0-6. Similar to the tosylate results we were only able to recover trace amounts of the product leading us to believe that the tosylated PEG linkers might be unstable.
Scheme 0-6 Route of synthesis for rhodamine B electrophiles using RhB piperazine and bis-substituted electrophiles

1.3.1.2.2 Rhodamine B piperazine propanol route

The next route we tried utilized Molecule 1-6 in an S_N2 type reaction with 3-bromopropanol, Scheme 0-7. As we continued to experience difficulty with tosylation of the polyethylene oxide alcohol we decided to work with an alkyl alcohol (3-bromopropanol) and reacted it in an S_N2 type reaction with Molecule 1-6 followed by reaction with para-toluenesulfonyl chloride to give a propanol tosylate analog, Molecule 1-10.

Scheme 0-7 Route of synthesis for Molecule 1-10

Various sterically hindered bases were attempted as well as differing polar aprotic solvents including acetonitrile and N,N-dimethylformamide. Figure 0-10 shows the MALDI-TOF results of the reaction mixtures between Molecule 1-6 and the varying sulfonyl chlorides. The desired products are marked with an arrow. In most cases we discovered an unwanted byproduct with a mass of 661 that was difficult to analyze by proton NMR in terms of low
concentration and impurities present. Only trace products were observed by MALDI-MS for the 
para-nitro and para-methoxy derivatives (Figure 1-10, traces b and c).

Figure 0-10 MALDI-TOF MS reaction mixture characterization of Molecule 1-6 with bis-
substituted sulfonate esters

MALDI-TOF MS of reaction mixtures of Molecule 1-6 with Bis substituted 
sulfonate esters whereby R= to the following in each panel: a) R=OTs, b) 
R=NO2, c) R=OMe and d) R=H.

1.3.1.2.2.1 Molecule 1-11 rhodamine B benzophenone analog

Finally, we thought that it might be possible to use benzophenone as a photoaffinity 
label to covalently attach our rhodamine analog to the aptamer so we prepared a 
benzophenone analog using Molecule 1-6 and reacted with 4-bromomethyl benzophenone 
using 2,6-lutidine as the base in acetonitrile, Scheme 0-8.
We attempted to use this product to label the proteins discussed in Chapter 2 with UV light, but no labeling was achieved. Disheartened by the high reactivity of these electrophiles and the difficulty in isolation of these desired products we chose to try working with sulforhodamine 101. In addition, we were not having success using rhodamine B analogs to label our proteins and in conjunction with the difficult synthetic challenges decided to use the analog of rhodamine B with the highest affinity for the aptamer, and that fluorophore was sulforhodamine 101.

1.3.2 Sulforhodamine 101 routes of synthesis

1.3.2.1 Route of synthesis for tosylate analog sulforhodamine 101 (Molecule 1-15) using PEG linker.

We started with sulforhodamine 101 sulfonyl chloride (also known as Texas Red) commercially- available as a mixture of isomers and reacted this with (PEO)$_3$-monoamine to give the desired sulfonamide with an alcohol functional group, Molecule 1-14 in Scheme 0-9. As mentioned before the reaction is chemo-selective towards producing the sulfonamide instead of the alcohol. TLC staining using Ninhydrin and KMnO$_4$ that qualitatively stains amines and alcohols, respectively verified the product. The alcohol of molecule 1-14 was reacted in an
$S_N2$-type reaction with $p$-toluenesulfonyl chloride to give the desired tosylate. The resulting HCl salt was quenched by pyridine.

**Scheme 0-9 Route of synthesis for tosylate analog of sulforhodamine 101 using PEG linker (Molecule 1-15)**
Similar to the PEG-alcohol of the rhodamine derivatives (Molecule 1-3) we experienced difficulty with low yields and hydrolysis of the desired tosylate (Molecule 1-15). Poor resolution of the products during purification using column chromatography was attributed to mixed isomers and the similarity of retention times for the starting material and product. While the alcohol was successfully purified using HPLC chromatography on an analytical scale, the desired tosylate (Molecule 1-15) was purified by HPLC but hydrolyzed under acidic conditions.

1.3.2.2 Route of synthesis for sulforhodamine 101 sulfonate esters utilizing click chemistry

Having faced a variety of difficulties attempting to make sulfonate esters by reaction of a PEG alcohol with sulfonyl chlorides we decided to try another route to produce the desired tosylates.

1.3.2.2.1 Sulforhodamine 101 sulfonyl chloride as the starting material for access to sulfonate esters.

To begin, we needed to access a more affordable starting material and chose to make sulforhodamine 101 sulfonyl chloride following a protocol described by Titus and coworkers\textsuperscript{75} utilizing phosphorus oxychloride. Phosphorus oxychloride (POCl\textsubscript{3}) or phosphoryl chloride is a hazardous colorless liquid\textsuperscript{81} used to make aryl and alkylphosphate triesters (fire resistant hydraulic fluids\textsuperscript{82}), nucleotides and oligonucleotides in nucleic acids research\textsuperscript{83}, dopant for semiconductor grade silicon\textsuperscript{84}, as a dehydrating agent to convert primary amides to nitriles (Bischler-Napieralski reaction\textsuperscript{85}). Phosphorus oxychloride is a chlorinating agent\textsuperscript{86} and for our purposes used to convert sulfonic acids to sulfonyl chlorides.

We reacted sulforhodamine 101 free acid (SR101, Molecule 1-12) with neat phosphorus oxychloride to give the desired sulfonyl chloride, Scheme 0-10. Special care in terms of using
anhydrous conditions needed to be taken to minimize hydrolysis back to the sulfonic acid. The time of the reaction was optimized such that formation of the mono sulfonyl chloride (Molecule 1-13) was produced in decent yield while minimizing formation of the bis sulfonyl chloride (Molecule 1-16).

**Scheme 0-10 Formation of sulforhodamine 101 sulfonyl chloride (Molecules 1-13 and 1-16)**

![Scheme 0-10 Formation of sulforhodamine 101 sulfonyl chloride (Molecules 1-13 and 1-16)](image)

We tried a variety of work up conditions to minimize hydrolysis while removing unreacted/hydrolyzed sulfonic acid from the desired product. The protocol given by Titus and coworkers\(^\text{75}\) does not account for minimization of the bis-sulfonyl adduct (Molecule 1-5) versus the mono-sulfonyl adduct (Molecule 1-13) and their TLC system gives a fast migrating system that does not resolve the two adducts.

We chose to use 1:9 MeOH: DCM as the resolving system for the sulforhodamine 101 sulfonyl chloride adducts. Before the reaction is quenched by water, TLC shows the formation of both mono and bis chloro adduct formation. Once the phosphoryl oxychloride is quenched with a water extraction any unreacted sulforhodamine 101 free acid and hydrolyzed acid is removed in the aqueous layer, the TLC plate presented in Figure 0-11 show formation of both adducts. The mono-chloro adduct has a lower Rf than the bis-chloro adduct.
In addition, we avoided the recrystallization procedure after the extraction step described by Titus and coworkers as it drastically reduced the yield. Instead, we removed the chloroform following the extraction under reduced pressure and immediately lyophilized the product to remove any residual water from the extraction process. In this way we were able to isolate fair yields of sulforhodamine 101 sulfonyl chloride without an extensive purification step and moved this material forward to the next step.

1.3.2.2 Azide Analog of sulforhodamine 101 (Molecule 1-17)

Towards the goal of making a sulforhodamine 101 tosylate electrophile we decided to use an azide intermediate that would be available for reaction with an alkyne bearing a sulfonate ester. Click chemistry is a term coined by Barry Sharpless and this work in part led to a shared Nobel price in chemistry in 2001\(^7\) and refers to a [3+2] cycloaddition reaction.
between an alkyne and azide. These reactions are generally high yielding, reactive with a wide variety of alkynes and azides, are stable to biological environments, stereospecific and highly selective. These features were attractive for us in terms of adding a sulfonate ester onto our complex molecule without cleavage of the electrophile.

In order to form the azide we decided to use a flexible polyethylene oxide that might enhance solubility in biological media. To avoid formation of a sulfonate ester we needed to react the sulfonyl chloride with an amine and not an alcohol and also needed an azide on the linker. Therefore we prepared 1-amino-11-azido-3,6,9-trioxaundecane in fair yields using published procedures\(^88\) for this intermediate, Scheme 0-11.

**Scheme 0-11 Route of synthesis for 1-Amino-11-azido-3,6,9-trioxaundecane (Molecule 1-7)**

Reaction of Molecule 1-13 with Molecule 1-7 resulted in a thiourea with the desired azide functionality needed for click chemistry, Scheme 0-12. Ease of purification was highly dependent on the purity of the starting sulfonyl chloride in terms of percentage of existing bis-sulfonyl chloride adduct (Molecule 1-16) versus the mono-sulfonyl chloride adduct (Molecule 1-13). The bis-sulfonyl chloride led to a bis-azole product, (Molecule 1-18).
As a salt, the bis azide sulforhodamine analog, Molecule 1-18, was difficult to purify and the retention factor on silica gel varied significantly between different lots of silica gel from different vendors. Typically, the bis azide (1-18) would elute faster than the mono azide (1-17) but would co-elute with the mono azide as well. We attempted chromatotron purification,
standard column chromatography and prep TLC using both regular and reverse phase plates. On the chromatatron the bands appeared to separate by visual inspection but analysis of the fractions collected proved impure with the bis salt, Figure 0-12.

**Figure 0-12 Sulforhodamine 101 azide analog purification on chromatatron (Molecules 1-17 and 1-18)**

We also tried the addition of organic acids and bases in the chromatography resolving solvents to provide a counter-ion for the bis salt but neither acetic acid nor ammonium hydroxide helped to prevent the bis salt from leaching off of the silica columns used. No recrystallization methods were found to separate the isomers from the bis product successfully.

Purification of this material also proved difficult using reverse phase HPLC. While the bands that eluted from the column had consistent retention times the components of those bands proved impure, Figure 0-13. For example, four fractions (A-D) were collected for the purification of Molecules 1-17 and 1-18 using a gradient of 5-55% ACN/H₂O over 50 minutes. All fractions contained the bis-substituted azide and the purity verified by proton NMR.
Therefore, additional efforts were made to optimize the separation by varying the solvent system. It was discovered that while methanol worked on TLC, the resolution of the products were too close together to be effective for purification on silica gel, Figure 0-14.
Figure 0-14 TLC resolution of Molecules 1-17 and 1-18 using methanol in dichloromethane.

Therefore, we tried other less polar alcohols to improve the separation and found that ethanol slightly improved the separation and isopropanol was superior to ethanol. We found that in the isopropanol solution the starting materials were resolved from the desired azide, Figure 0-15.

Figure 0-15: TLC resolution of Molecules 1-17 and 1-18 comparing MeOH to IPA

Comparing solvent Systems:
10% MeOH/DCM 15% IPA/DCM

Reaction mixtures at 30 minutes spotted in duplicate on each TLC plate.
Mono (first lane) – reference of product containing mostly mono N3 by NMR
Bis (second lane) – reference of product containing mostly bis N3 by NMR
1= Reaction mixture 1 without DIEA
2= Reaction mixture with DIEA
Finally, Molecule 1-13, the starting material, had a similar retention time as the desired mono azide product (Molecule 1-17) so we decided to hydrolyze it by quenching the reaction mixture with water, and the reaction mixture was then allowed to stir for several hours, then lyophilized overnight such that column chromatography could be run the following day. In this way the mono-sulfonyl chloride (Molecule 1-13) was hydrolyzed to the sulforhodamine 101 free acid (Molecule 1-12) and its resulting retention factor was baseline and this rendered an easier purification. Optimization to separate the isomers from one another was accomplished varying the percentage of IPA and evaluation of those isomers using NMR spectroscopy is discussed in more detail in Section 1.3.2.3.

1.3.2.2.3 Tosylate analog of sulforhodamine 101 (Molecule 1-19)

Once Molecule 1-17 had been purified, the azide could be transformed into the desired triazole bearing a sulfonate ester by reaction with 3-butynyl p-toluenesulfonate, see Scheme 0-13.
The tosylate product (Molecule 1-19) was purified using RP-HPLC and taken up into DMSO at a 52.4 μM stock concentration, aliquoted into multiple tubes and this stock solution was used for protein and lysate labeling studies, see Chapter 2.

1.3.2.2.4 *Meta*-ethylamido sulfonate ester analog of sulforhodamine 101 (Molecule 1-22)

Once Molecule 1-19 was tested against various proteins and cell lysate samples we discovered that it labeled a number of proteins without the desired selectivity we were hoping
for. It was thought the reactivity of the $S_N2$-type label might be faster than the proximal rate of reactivity between the binding partners. As such, we needed to modulate the electrophilicity of the leaving group and chose a leaving group with intermediate reactivity. A novel sulfonate ester that has been used for biological labeling purposes includes a phenyl ethlamido moiety meta to a sulfonate ester$^{89}$.

The desired sulfonate ester (Scheme 0-14) was made by reacting 3-(chlorosulfonyl) benzoyl chloride with ethylamine to give Molecule 1-20. The sulfonyl chloride was then reacted with 3-butyn-1-ol to give the desired alkyne (Molecule 1-21) that could then be reacted with Molecule 1-17 to give the sulforhodamine 101 fluorophore with an ethylamido functional group meta to the sulfonate ester (Molecule 1-22) as the electrophile.
Scheme 0-14 Route of synthesis for Molecule 1-22

Molecule 1-22 was reacted with various proteins and cell lysate whose binding properties are described in chapter 2.
1.3.2.2.5 Sulforhodamine-(PEO)$_3$-Azide as a versatile intermediate

Working with the click conditions allowed us to bypass the problems with the tosylation of the alcohol on the PEG linker to form a sulfonate ester. Having recognized the azide analog of sulforhodamine 101 (Molecule 1-17) as a useful intermediate we chose to functionalize sulforhodamine 101 with other potentially useful small molecules, Scheme 0-15.

**Scheme 0-15 Molecule 1-17 as a useful intermediate**

Where $R =$

- Propargyl Biotin
- Alkyne-PEG4-maleimide
- Propargyl-N-hydroxysuccinimidy1 ester
Currently the PEG4-maleimide and N-hydroxysuccinimidy ester are in the process of purification and the biotinylated sulforhodamine 101 sulfonate ester has been prepared, Scheme 0-16.
Scheme 0-16 Biotinylated sulforhodamine 101 (Molecule 1-23)

1-17 (mix of isomers)

1-23 (mix of isomers)
1.3.2.2.6 NMR studies of Molecule 1-17

In order to differentiate the isomers separated by TLC using IPA/DCM of TR-(PEO)$_3$-N$_3$, a variety of NMR studies were employed. The isomer with the higher Rf is called Isomer A, and the isomer with a lower Rf is called Isomer B, see Figure 0-16.

Figure 0-16 TLC of Molecule 1-17 isomers

TLC representation of Molecule 1-17 isomers separated by 15%/85% IPA/DCM. The left lane represents Isomer A, with a higher Rf. The right lane represents Isomer B with a lower Rf.

To begin, the isomers were characterized and compared using $^1$H NMR spectroscopy, Figure 0-17. In addition, $^{13}$C NMR spectroscopy was used to characterize Isomer A as there was enough mass needed to obtain that data. The following proton assignments were based on theoretical ChemNMR 1H Estimation using ChemBioDraw Ultra (Trademark Perkin Elmer)$^{90}$, 2D-$^1$H COSY data, and comparison of products with starting materials via 1D $^1$H NMR. The $^1$H-NMR of isomer B shows DMF impurities. In addition, the NH of the sulfonamide peak was assigned after comparison to spectra obtained from the same sample on another 400 MHz instrument.
The isomer shown on the NMR spectra is the ortho sulfonamide isomer. Based on $^1$H NMR it is not possible to distinguish the ortho isomer from the para isomer. However, it is possible to qualitatively assess the purity of the isomers and validate that there is a difference between key protons. For example, the chemical environment for Ha is similar for both isomers and a large chemical shift was not expected. In this case, the shift was less than 0.1 ppm. However, the shift for Hb was expected to dramatically shift as it would be ortho to the sulfonamide instead of meta to the isomer and the shift observed for this proton was 0.43 ppm. A comparison of the shifts for the assigned protons is provided in Table 0-2.
Table 0-2 Comparison of \(^1\)H NMR Shifts for Isomers of Molecule 1-17

![Molecule Diagram]

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Theoretical Shift</th>
<th>Isomer B</th>
<th>Isomer A</th>
<th>Shift in ppm</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>*ortho 8.44 para 8.44</td>
<td>8.74 (s, 1H)</td>
<td>8.82 (s, 1H)</td>
<td>-0.08</td>
</tr>
<tr>
<td>b</td>
<td>*ortho 7.81 para 7.68</td>
<td>8.40 (d, J = 9.1 Hz, 1H)</td>
<td>7.97 (s, 1H)</td>
<td>0.43</td>
</tr>
<tr>
<td>c</td>
<td>*ortho 7.94 para 7.94</td>
<td>7.18 (d, J = 8.0 Hz, 1H),</td>
<td>7.16 (d, J = 7.9 Hz, 1H)</td>
<td>0.02</td>
</tr>
<tr>
<td>d,e</td>
<td>6.88</td>
<td>6.73 (s, 2H),</td>
<td>6.77 (s, 2H)</td>
<td>-0.04</td>
</tr>
<tr>
<td>f</td>
<td>7.74</td>
<td>5.50 (s, NH)</td>
<td>5.72 (d, J = 5.9 Hz, 1H)</td>
<td>-0.22</td>
</tr>
<tr>
<td>h,i,j,l</td>
<td>3.54-3.76</td>
<td>3.58 (m, 9H),</td>
<td>3.58 (m, 10H)</td>
<td>0.00</td>
</tr>
<tr>
<td>m,r,k</td>
<td>2.79,</td>
<td>3.46 (s, 14H),</td>
<td>3.37 (m, 11H)</td>
<td>0.09</td>
</tr>
<tr>
<td>g</td>
<td>3.47</td>
<td>3.16 (d, J = 1.3 Hz, 2H),</td>
<td>3.30 (d, J = 5.4 Hz, 2H)</td>
<td>0.04</td>
</tr>
<tr>
<td>p</td>
<td>3.78</td>
<td>3.02 (s, 4H),</td>
<td>2.98 (t, J = 6.3 Hz, 4H)</td>
<td>0.04</td>
</tr>
<tr>
<td>o</td>
<td>3.78</td>
<td>2.74 (m, 4H),</td>
<td>2.69 (m, 4H),</td>
<td>0.05</td>
</tr>
<tr>
<td>n</td>
<td>1.68</td>
<td>2.11 (m, 4H)</td>
<td>2.06 (m, 4H)</td>
<td>0.05</td>
</tr>
<tr>
<td>q</td>
<td>1.68</td>
<td>1.98 (m, 3H)</td>
<td>1.92 (m, 4H)</td>
<td>0.06</td>
</tr>
</tbody>
</table>

*relative to sulfonamide julolidine ring system

The two dimensional spectrum in Figure 0-18 resulted from a \(^1\)H COSY experiment in CDCl₃. \(^1\)H-CORrelated SpectroscopY (COSY) is a useful method for determining proton proton correlations between neighboring protons, typically up to four bonds apart.
Analysis of the cross peaks provided evidence of the protons that are coupled. In this way the following correlations were established, Table 0-3. The aromatic peaks were assigned relative to their splitting and correlation to other aromatic protons or lack thereof. Ha is ortho to both the sulfonamide and sulfonic acid, is a single, integrates to a single proton and does not correlate to any other protons. Hb and Hc are ortho to one another and split each other. They are differentiated from one another based on the large shift observed for Hb between the two isomers. Hd and He are chemically equivalent and integrate to 2 protons as a singlet. The NH of the sulfonamide splits into a triplet due to its neighboring methylene group and correlates to a methylene Hg. The alkyl groups of the julolidine ring system and the polyethylene oxide
linker overlapped one another. COSY provided correlation between the multiplets that integrated to 4H. For example, Hn integrates to 4H and sits between Ho and Hm, and a correlation was observed for both Ho and Hm. The same is true for Hq and both Hr and Hp.

Table 0-3 Correlation table of COSY for Molecule 1-17 Isomer A

<table>
<thead>
<tr>
<th>Correlation</th>
<th>( \text{a} ) to ( \text{X} )</th>
<th>( \text{b} ) to ( \text{c} )</th>
<th>( \text{d} ) to ( \text{e} )</th>
<th>( \text{f} ) to ( \text{g} )</th>
<th>( \text{h} ) to ( \text{i} )</th>
<th>( \text{j} ) to ( \text{k} )</th>
<th>( \text{l} ) to ( \text{m} )</th>
<th>( \text{n} ) to ( \text{o} )</th>
<th>( \text{p} ) to ( \text{q} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{a} ) to ( \text{X} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
</tr>
<tr>
<td>( \text{b} ) to ( \text{c} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
</tr>
<tr>
<td>( \text{d} ) to ( \text{e} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
</tr>
<tr>
<td>( \text{f} ) to ( \text{g} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
</tr>
<tr>
<td>( \text{h} ) to ( \text{i} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
</tr>
<tr>
<td>( \text{j} ) to ( \text{k} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
</tr>
<tr>
<td>( \text{l} ) to ( \text{k} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
</tr>
<tr>
<td>( \text{r} ) to ( \text{q} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
</tr>
<tr>
<td>( \text{m} ) to ( \text{n} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
</tr>
<tr>
<td>( \text{p} ) to ( \text{q} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
<td>( \text{O} )</td>
</tr>
</tbody>
</table>

\( \text{X} \) corresponds to a proton that does not exhibit a correlation to any other proton.
\( \text{O} \) corresponds to a proton that has a correlation and the protons for which it is coupled to is on the same line.

1.4 Discussion

Development of a rhodamine based electrophile proved difficult using rhodamine B (Molecule 1-1) as the starting material. For rhodamine B the competitive intramolecular
cyclization as a competing reaction for derivatization led to synthesis of the piperazine intermediate (Molecule 1-6) with a tertiary amide that could not cyclize. However, utilizing a PEG linker as the source of the hydroxyl functional group for tosylation led to an unstable rhodamine electrophile. Pegylated tosylates are not inherently unstable and a di-tosylated triethylene glycol that we had made was stable at room temperature for several months. Multiple attempts were made to isolate and characterize the major product that was a result of the degradation.

The successful synthesis of various rhodamine analogs was accomplished despite challenging purification problems. Many of these difficulties were overcome by optimizing the eluents used in flash column chromatography, increasing the silica gel to compound ratios (typically 1:100 or above), and adjusting the run times and gradients in HPLC separations. In addition, a number of variables were tested to optimize the yields of the products in the reaction mixtures including time, concentration of reaction components, solvents, and temperature. The formation of a sulfonate ester by reaction of an alcohol with a sulfonyl chloride expected to follow an SN$_2$-type mechanism. These types of reaction are dependent on the concentration of the nucleophile (Molecule 1-3 and Molecule 1-7) and the electrophile ($p$-toluenesulfonyl chloride). There are four major factors affecting the rate of the reaction, the first is the steric hindrance and stabilization factor of the electrophilic substrate. Second is the strength of the nucleophile, alkoxide anions are strong nucleophiles and react readily with sulfonyl chlorides. Thirdly, SN$_2$-type nucleophilic substitution reactions are optimized using polar aprotic solvents with low dielectric constants favor and as such, DMSO, DMF and acetonitrile were tested. Finally, the leaving group affects the rate of the reaction and sulfonyl
chlorides are excellent leaving groups. Based on these results we concluded that the formation of the product was not the limiting factor in isolating our rhodamine analogs. However, the stability of the product formed based on degradation with time observed on isolated products using NMR and MALDI-TOF analysis was believed to be the source of the challenges in isolating the desired sulfonate esters.

The difficulties associated with tosylation of the PEG linker on the rhodamine B analogs were also observed for the sulforhodamine 101 analogs and the same methods of optimization were attempted as described above. Thereafter, we chose to bypass formation of the tosylate via the PEG alcohol and instead attached the sulfonate ester directly via a [3+2] cycloaddition reaction. The sulfonate ester linked to sulforhodamine 101 via a triazole was found to be stable. Another complication in the route of synthesis towards our sulforhodamine 101 analogs was formation of a positively charged bis-azide sulforhodamine 101 salt that co-eluted with the desired mono-azide using HPLC and flash column chromatography. The best method to minimize these purification difficulties was to first attempt to minimize the formation of the bis-sulfonyl chloride and secondly to utilize a sterically hindered alcohol with a lower polarity than methanol with dichloromethane to optimize separation of the bis analog as well as providing a means of separation for the ortho and para- mono-substituted isomers. While methanol was compared to ethanol and isopropanol it would be interesting to also test n-butanol or 2-butanol to see if this would enhance the resolution.

Optimization of the route of synthesis towards electrophilic sulforhodamine 101 analogs was accomplished. In addition, the purification methods provide a means to separate the ortho/para isomers from one another. Commercially, the sulforhodamine 101 analogs that are
currently available for derivatization are economically restrictive and are not isomerically resolved. In addition, the azide analog prepared creates a highly flexibly intermediate that provides the means to derivatization of sulforhodamine 101 with nucleophiles, electrophiles and other useful analogs, some of the possibilities available are highlighted in Scheme 0-17.

**Scheme 0-17 Molecule 1-17 as a useful intermediate for sulforhodamine 101 analogs**
1.5 Experimental methods

MALDI-MS experiments were performed on a Micromass MALDI-TOF mass spectrometer and a CEM Voyager MALDI-TOF mass spectrometer using CHCA (10 mg/mL in 1:1:0.3 water: ACN: TFA), CHCA (12 mg/mL in 0.7:0.3:0.1 water: ACN: TFA) or DHB (10 mg/mL in 1:0.1 water: TFA). Experiments were run in positive ion mode with a linear flight path for detection. $^1$H NMR/$^{13}$C spectra were recorded on 400 MHz Avance and 300 and/or 400 MHz Varian Instruments. HPLC purification was carried out on a Shimadzu prominence system using Vydac (218TP C18 5µ column using 0.1% TFA in acetonitrile and water as the eluents and monitored at $\lambda_{\text{max}}= 586$ and 254 nm. Flash chromatography was performed using Silica gel (32-63 µ). TLC monitoring was used with visible light and short wave (254 nm) and long wave (365 nm) UV light.

1.5.1.1 Molecule 1-1: Rhodamine B

Commercially-available

Molecule 1-1: $N$-(9-(2-carboxyphenyl)-6-(diethylamino)-3$H$-xanthen-3-ylidene)$-N$-ethylethanaminium. $^1$H NMR (400 MHz, DMSO- $d_6$): $\delta$ 8.24 (dd, $J = 7.8$, 1.4 Hz, 1H), 7.85 (dtd, $J$ =
25.2, 7.5, 1.4 Hz, 2H), 7.48 (dd, J = 7.5, 1.3 Hz, 1H), 7.11 (dd, J = 9.5, 2.4 Hz, 2H), 7.06 – 6.95 (m, 4H), 3.66 (q, J = 7.1 Hz, 8H), 1.22 (t, J = 7.0 Hz, 12H). $^1$H NMR (101 MHz, DMSO): δ 166.21, 157.05, 155.00, 132.62, 130.96, 130.80, 130.31, 130.16, 114.44, 112.84, 95.84, 45.28, 30.66, 12.41. MS (MALDI-TOF) (m/z) Calculated for C$_{28}$H$_{31}$N$_{2}$O$_{3}$ [M+H$^+$]: 444.24, observed: 444.22.

1.5.1.2 *Molecule 1-2: Rhodamine B Isothiocyanate*

![Chemical structure of Rhodamine B Isothiocyanate]

Chemical Formula: C$_{29}$H$_{30}$ClN$_{3}$O$_{3}$S  
Exact Mass: 535.17  
Molecular Weight: 536.08

Commercially-available

Molecule 1-2: N-(9-(2-carboxy-4-isothiocyanatophenyl)-6-(diethylamino)-3H-xanthen-3-ylidene)-N-ethylethanaminium. $^1$H NMR (300 MHz, DMSO-$d_6$): δ 8.40 – 8.09 (m, 1H), 8.01 – 7.75 (m, 1H), 7.74 – 7.36 (m, 1H), 7.35 – 6.81 (m, 12H), 3.72 – 3.58 (m, 15H), 3.53 – 3.19 (m, 8H), 1.30 – 1.14 (m, 15H), 1.03 (td, J = 7.1, 3.8 Hz, 3H), 0.90 – 0.81 (m, 2H).
Procedure: To a pre-dried 20 mL scintillation vial with a rubber lined septa was added PEO$_3$-monoamine (111.5 mg, 0.75 mmol) followed by 6.0 mL of anhydrous DMF and 1.0 mL of anhydrous dichloromethane. This mixture was cooled to 4°C followed by the addition of the rhodamine B isothiocyanate (50.0 mg, 0.09 mmol). The reaction mixture was allowed to stir overnight under nitrogen at 4°C. Progress of the reaction was monitored via TLC using 20% methanol in chloroform. Preparative TLC was used to purify the product using the same TLC conditions to give a purple solid, Rf=0.41.

Molecule 1-3: N-(9-(2-carboxy-4-(3-(2-(2-hydroxyethoxy)ethoxy)ethyl)thioureido)phenyl)-6-(diethylamino)-3H-xanthen-3-ylidene)-N-ethylethanaminium chloride. Rf= 0.41 (20% methanol in chloroform). Purple solid. Yield: 11.1 mg, 20%. $^1$H NMR (DMSO-d$_6$, 400 MHz): $\delta$ 7.68 (d, $J = 8.2$ Hz, 5H), 7.39 (d, $J = 7.9$ Hz, 4H), 3.51 – 3.34 (m, 14H), 3.17 (d, $J = 5.2$ Hz, 1H), 2.87 (dd, $J =...
10.3, 4.3 Hz, 3H), 2.57 (d, J = 2.4 Hz, 1H), 2.41 (s, 1H), 2.38 (s, 5H). δ = . MS (MALDI-TOF) (m/z) calculated for C_{35}H_{45}N_{4}O_{6}S \ [M+H]^+: 650.32, observed: 650.32.

1.5.1.4 *Molecule 1-4*

![Chemical structure of Molecule 1-4](image)

Chemical Formula: C_{42}H_{51}N_{4}O_{8}S_{2}^+
Exact Mass: 803.31
Molecular Weight: 804.01

To a stirred solution of Molecule 1-3 (30.4 mg, 0.05 mmol) in anhydrous degassed dichloromethane (1.26 mL) was added p-toluenesulfonyl chloride (12.0 mg, 0.01 mmol), N,N-diisopropylethylamine (17.5 mg, 0.15 mmol) and 4-dimethylaminopyridine (1.7 mg, .01 mmol), stirred at room temperature overnight. The solvent was removed under vacuum and monitored via TLC and MALDI-TOF. Only trace products were detected, Rf= 0.64+0.77 (isomers) in 20% methanol in chloroform. Therefore, 2 additional equivalents of base and p-toluenesulfonyl chloride were added with no change in progress. The trace product was purified using preparatory TLC in 20% methanol in chloroform, trace product isolated had an Rf=0.73.
Molecule 1-4: \( N-(9-(2\text{-carboxy}-4-(3-(2-(2\text{-}(\text{tosyloxy})\text{ethoxy})\text{ethoxy})\text{ethyl})\text{thioureido})\text{phenyl})-6-(\text{diethylamino})-3H\text{-xanthen}-3\text{-ylidene})-N\text{-ethylethanaminium}. \) Rf = 0.73 (methanol/chloroform = 1/4). Crude \(^1\text{H} \text{NMR} \) (400 MHz, DMSO-\( \text{d}_6 \)) \( \delta \) 7.72 – 7.65 (m, 1H), 7.62 (d, \( J = 3.8 \) Hz, 2H), 7.60 (d, \( J = 3.6 \) Hz, 2H), 7.38 (d, \( J = 9.6 \) Hz, 1H), 7.15 (d, \( J = 3.2 \) Hz, 2H), 7.13 (d, \( J = 2.3 \) Hz, 2H), 7.00 (d, \( J = 4.7 \) Hz, 1H), 3.64 (s, 3H), 3.60 – 3.44 (m, 7H), 3.37 (s, 21H), 3.08 (dd, \( J = 10.8, 3.5 \) Hz, 19H), 2.54 (d, \( J = 3.3 \) Hz, 5H), 2.43 – 2.23 (m, 9H), 1.24 (s, 12H), 1.17 (dd, \( J = 8.9, 5.6 \) Hz, 33H), 1.07 (s, 4H).

Crude reaction mixture MS (MALDI-TOF) (m/z) Calculated for \( \text{C}_{42}\text{H}_{51}\text{N}_4\text{O}_8\text{S}_2 \) [M+H\(^+\)]: 804.32, observed: 804.38.

**Molecule 0-5 Rhodamine B Base\(^80\)**

\[ \text{Chemical Formula: C}_{28}\text{H}_{30}\text{N}_2\text{O}_3 \]
\[ \text{Exact Mass: 442.23} \]
\[ \text{Molecular Weight: 442.55} \]

Rhodamine B (4g, 8.35 mmol) was dissolved into 500 mL of 1M NaOH to which was added 200 mL ethyl acetate. For the first extraction, additional EtOAc was added until the extraction produced discernible layers. The aqueous later was then extracted another 4X 200 mL portions of EtOAc, until the aqueous layer ceased to be pink in color. The combined organic layers were washed with 1M NaOH that had been saturated with sodium chloride. The organic layer was
dried over anhydrous magnesium sulfate, filtered, and the solvent removed under reduced pressure to give pink foam. This pink foam was triturated with ether to give a pink solid.

Molecule 1-5: 3',6'-bis(diethylamino)-3',9a'-dihydro-3H-spiro[isobenzofuran-1,9'-xanthen]-3-one. Dark pink solid. Yield: 3.23 g, 7.3 mmol, 87%. $^1$H NMR (400 MHz, DMSO-$d_6$): $^1$H NMR (400 MHz, DMSO-$d_6$): $\delta$ 7.98 (dd, $J = 7.5$, 1.1 Hz, 1H), 7.73 (dtd, $J = 28.7$, 7.5, 1.1 Hz, 1H), 7.25 (d, $J = 7.6$ Hz, 1H), 6.52 – 6.35 (m, 6H), 3.35 (q, $J =$ 7.0 Hz, 8H), 1.09 (t, $J = 7.0$ Hz, 12H). $^{13}$C NMR (101 MHz, DMSO) $\delta$ 163.05, 158.38, 150.60, 149.70, 140.77, 129.98, 129.22, 128.99, 128.82, 128.20, 119.99, 118.19, 113.57, 108.68, 95.49, 60.36, 56.06, 52.78, 13.86. MS (MALDI-TOF) (m/z) Calculated for $C_{28}H_{30}N_2O_3$: 443.24, observed: 443.22.

1.5.1.5 Molecule 1-6

Molecule 0-6

1-6

Chemical Formula: $C_{32}H_{39}N_4O_2^+$
Exact Mass: 511.31
Molecular Weight: 511.68

CAUTION: Read and review procedures for safe use of pyrophoric liquid reagents.
A two-neck 50 ml round bottom flask fitted with a 14/20-addition funnel and reflux condenser was purged with argon three times containing rhodamine B base, Molecule 0-5 (1g, 2.26 mmol) in anhydrous dichloromethane (6 mL) to which was added trimethyl aluminum (2.26 mL, 2.0M in hexanes) via a syringe over 10 minutes. Voluminous white smoke was formed during the addition, however no exotherm was noted. The reaction was stirred for one hour and a white sticky solid appeared on the sides of the flask. Under argon, the addition funnel was quickly removed, covered with a rubber septa and the mixture was allowed to reflux for 24 hours. 0.1 N HCl was added until gas evolution desisted, (note that gas evolution was difficult to see due to the dark purple coloration). The reaction mixture was then filtered using dichloromethane to rinse and 4:1 DCM/MeOH. The filtrate was combined, concentrated, taken up into dichloromethane, then filtered and concentrated again. The solid was taken up into dilute aqueous sodium carbonate, until the acid was neutralized and the aqueous layer was saturated with sodium chloride and washed with ethyl acetate three times, this removed unreacted starting material. The aqueous layer was then acidified with concentrated HCl until pH=2. Additional sodium chloride was added until saturated and the aqueous layer was then extracted with 2:1 CH₂Cl₂: isopropanol until a faint pink color persisted in the aqueous layer. Note that the brine prepared and used to effectively separate the layers should be as saturated as possible. The organic layers were combined and then concentrated. It is necessary to remove all of the isopropanol from the organic layer as it dissolved drying agents that precipitated from the following recrystallization step. Therefore, once the organic layer had been concentrated, it was taken up into dichloromethane and then dried with anhydrous magnesium sulfate and concentrated. Recrystallization: Dissolved purple brown solid with
minimal amount of methanol to which was dripped slowly into a liter of ether. The solution was filtered to reveal a dark brown shiny solid. The filtrate was concentrated and the recrystallization repeated to yield an additional portion of the product. A third recrystallization did not produce any additional product.

Molecule 1-6: N-(6-(diethylamino)-9-(2-(piperazine-1-carbonyl)phenyl)-3H-xanthen-3-ylidene)-N-ethylethanaminium. Dark brown solid. Yield: 0.7483 g, 68% yield. $^1$H NMR (400 MHz, DMSO-d$_6$): $\delta = 10.17$ (s, 2H), 7.91 – 7.68 (m, 3H), 7.52 (dd, $J = 5.8$, 3.1 Hz, 1H), 7.19 – 7.04 (m, 4H), 6.94 (d, $J = 2.2$ Hz, 2H), 3.68 (dq, $J = 17.9$, 10.0, 7.1 Hz, 20H), 3.00 (s, 2H), 2.93 – 2.79 (m, 2H), 1.21 (t, $J = 7.0$ Hz, 12H). $^{13}$C NMR (101 MHz, DMSO) $\delta = 215.93$, 215.71, 176.81, 174.41, 158.55, 155.69, 154.59, 153.57, 130.45, 128.65, 125.87, 125.79, 118.39, 116.86, 108.09, 70.28, 31.85, 24.35, 22.23. MS (MALDI-TOF) (m/z) Calculated for C$_{32}$H$_{39}$N$_4$O$_2$ [M+H$^+$]: 512.32, observed: 512.12.

1.5.1.6 Molecule 1-7

Molecule 0-7

Triethylene glycol (1.5 g, 10 mmol) was dissolved in 30 mL of THF and triethylamine (4.0 g, 40 mmol), DMAP (13 mg, 1.1 mmol) and gassed/degassed three times with argon then cooled to 0°C. p-Toluenesulfonyl chloride (4.0 g, 21 mmol) dissolved in 20 mL of THF was added drop wise
over a period of 30 minutes and warmed to room temperature. The mixture was extracted with DCM three times and the organic layers combined, washed with 1 M HCl, brine and then dried over anhydrous sodium sulfate. The product was purified using flash column chromatography with a 1:3 mixture of EtOAc: Hexanes with an Rf=0.34, using a KMnO₄ and PMA stain to visualize the unreacted TEG and UV light at 254 nm to visualize the desired product.

Molecule 1-7: (ethane-1,2-diylbis(oxy))bis(ethane-2,1-diyl) bis(4-methylbenzenesulfonate).

White solid. Rf=0.34 in ethyl acetate. Yield: 2.67 g, 58.4%. Rf=0.34 using 25% ethyl acetate in hexanes. \( ^1H \text{NMR} \) (400 MHz, DMSO-\(d_6\)): \( \delta = 7.84 – 7.74 \) (m, 2H), 7.48 (d, \( J = 8.1 \) Hz, 2H), 4.18 – 4.04 (m, 2H), 3.59 – 3.52 (m, 2H), 3.40 (s, 3H), 2.42 (s, 3H). \( ^13C \text{NMR} \) (101 MHz, DMSO): \( \delta \) 144.85, 132.40, 130.08, 127.57, 69.89, 69.55, 67.86, 21.03.

1.5.1.7 Molecule 1-8

Molecule 0-8

Chemical Formula: \( C_{45}H_{57}N_4O_7S \)
Exact Mass: 797.39
Molecular Weight: 798.02
To a pre-dried 5 mL vial with a stir bar was added Molecule 1-7 (179.3 mg, 0.391 mmol), Molecule 1-6 (100.0 mg, 0.195 mmol) and anhydrous DMSO (1.75 mL) and triethylamine (44.2 mg, 0.342 mmol). The reaction was allowed to stir 4 days adding one additional equivalent of Molecule 1-7 after the first day. The reaction was monitored using two TLC solvent systems, 9:1 IPA: NH4OH and 1:9 MeOH: dichloromethane. The reaction mixture was washed three times with dichloromethane in brine and lyophilized to remove any residual DMSO. The product was purified by RP-HPLC over a 20-60% gradient over 45 minutes using Water: ACN with 0.1% TFA as the mobile phase. The product was lyophilized to give a dark purple solid (10.2 mg, 9.2% yield).

Molecule 1-8: Dark purple solid. Yield: 10.2 mg, 9.2% yield. 1H NMR (400 MHz, DMSO-d6): \( \delta = \) 7.95 (s, 1H), 7.93 – 7.79 (m, 1H), 7.73 (d, \( J = 8.8 \) Hz, 3H), 7.66 – 7.60 (m, 1H), 7.49 – 7.30 (m, 1H), 7.27 – 7.07 (m, 5H), 6.95 (d, \( J = 2.1 \) Hz, 2H), 3.67 (dd, \( J = 8.6, 5.7 \) Hz, 11H), 3.56 – 3.40 (m, 6H), 2.89 (s, 3H), 2.73 (s, 3H), 2.28 (s, 1H), 1.25 – 1.18 (m, 15H). 13C NMR (101 MHz, DMSO) \( \delta = \) 159.89, 157.03, 155.10, 151.17, 145.50, 137.72, 130.10, 128.12, 128.05, 127.54, 125.49, 125.44, 113.00, 95.88, 89.83, 69.97, 69.52, 67.84, 53.55, 48.57, 45.35, 41.80, 20.75, 18.03, 16.69, 12.37, 2.92. MS (MALDI-TOF) (m/z) Calculated for C_{45}H_{57}N_{4}O_{7}S^{+} [M+H^{+}]: 798.40, observed: 798.64.
To a 5 mL vial with a stir bar was added 3-bromopropanol (10.1 mg, 0.07 mmol), 0.5 mL DMF, and DIEA (16.6 mg, 0.13 mmol), followed by Molecule 1-6 (30.0 mg, 0.06 mmol). The reaction was monitored by MALDI-TOF MS using CHCA as the matrix, once the starting material had disappeared after 24 hours, the reaction mixture was taken up into 100 mL of saturated sodium bicarbonate and washed three times with 150 mL of EtOAc. The aqueous layer was then extracted with a 1:3 solution of isopropanol and dichloromethane, the organic layers combined and lyophilized after freezing to give a dark purple solid (9.1 mg, 19.5% yield).

Molecule 1-9: Dark purple solid. Yield: 9.1 mg, 19.5% yield. $^1$H NMR (400 MHz, DMSO-$d_6$): $\delta =$ 9.65 (s, 2H), 7.77 (s, 3H), 7.55 (s, 1H), 7.15 (d, $J = 9.5$ Hz, 4H), 6.95 (s, 2H), 3.73 – 3.54 (m, 12H), 3.16 – 3.06 (m, 4H), 2.54 (s, 9H), 1.39 (s, 4H), 1.35 – 1.25 (m, 25H), 1.22 (d, $J = 6.7$ Hz, 16H).

MS (MALDI-TOF) (m/z) Calculated for C$_{35}$H$_{45}$N$_4$O$_3$ [M+H$^+$]: 570.36, observed: 570.25.
1.5.1.9  **Molecule 1-10**

**Molecule 0-10**

![Chemical structure of Molecule 0-10]

To a 100 mL round bottom flask containing the lyophilized Molecule 1-9 (46.8 mg, 0.08 mmol) was added 10 mL of anhydrous THF and TEA (33 mg, 0.33 mmol) and allowed to mix under argon for 30 minutes followed by the addition of *para*-toluenesulfonyl chloride. The reaction mixture was allowed to stir overnight at room temperature. The progress of the reaction was monitored by MALDI-TOF MS and TLC using 10% Methanol in dichloromethane. The solvent of the reaction mixture was removed under vacuum and purified by column chromatography using a reverse phase column (Teledyne ISCO- C18 column, 40-60 micron average particle size) using 10%-100% acetonitrile in water. The product was a dark purple solid, (24.9 mg, 42.2% yield).

Molecule 1-10: Dark purple solid. Yield: 24.9 mg, 42.2% yield. $^1$H NMR (400 MHz, DMSO-$d_6$): $\delta$ = 7.83 (d, $J = 8.2$ Hz, 2H), 7.74 (d, $J = 8.0$ Hz, 2H), 7.42 (t, $J = 7.8$ Hz, 2H), 7.30 (t, $J = 7.3$ Hz, 2H), 4.52 (t, $J = 6.2$ Hz, 1H), 3.85 – 3.68 (m, 6H), 3.67 – 3.34 (m, 9H), 2.09 – 1.97 (m, 5H), 1.37 – 1.03
(m, 10H), 0.88 (s, 3H). MS (MALDI-TOF) (m/z) Calculated for C_{34}H_{31}N_{2}O_{5}S [M+H⁺]: 724.37, observed: 724.98.

1.5.1.10 Molecule 1-11

Molecule 0-11

To a 2 mL vial with a stir bar was added Molecule 1-6 (50.0 mg, 0.1 mmol) and dissolved in 0.5 mL acetonitrile to which was added 2,6-lutidine (11.7 mg, 0.12 mmol) mixed in 0.5 mL acetonitrile and 4-bromomethyl benzophenone (31.7 mg, 0.12 mmol) was added\textsuperscript{91}. The reaction was monitored via MALDI-TOF MS and pushed to completion with the addition of more dichloromethane. After 24 hours the reaction was dried under decreased pressure and purified via flash column chromatography using 7% methanol in dichloromethane, \( R_f = 0.31 \) revealing a purple solid, 36.8 mg, 34.8% yield.
Molecule 1-11: Dark purple solid. Yield: 36.8 mg, 34.8% yield. Rf-0.31 in 7% methanol in dichloromethane. $^1$H NMR (400 MHz, Chloroform-$d$) $\delta$ 7.75 (dd, $J = 18.3$, 7.6 Hz, 3H), 7.70 – 7.55 (m, 2H), 7.51 – 7.38 (m, 2H), 7.38 – 7.29 (m, 1H), 7.27 (s, 1H), 6.99 (d, $J = 9.4$ Hz, 1H), 6.81 (d, $J = 2.3$ Hz, 1H), 3.65 (d, $J = 7.1$ Hz, 5H), 3.58 (s, 2H), 2.38 (d, $J = 27.5$ Hz, 2H), 2.09 (s, 2H), 1.33 (t, $J = 7.1$ Hz, 6H), 1.26 (s, 6H). MS (MALDI-TOF) (m/z) Calculated for C$_{46}$H$_{49}$N$_4$O$_3$ [M+H$^+$]: 706.39, observed: 706.06.

1.5.1.11 Molecule 1-12 Sulforhodamine 101 (free acid)

Molecule 0-12 Sulforhodamine 101 Free Acid

\[ 	ext{Chemical Formula: C}_{31}	ext{H}_{30}	ext{N}_2	ext{O}_7	ext{S}_2 \]
\[ \text{Exact Mass: 606.15} \]
\[ \text{Molecular Weight: 606.71} \]

Commerciably-Available

Molecule 1-12: $^1$H NMR (400 MHz, 1:1 CDCl$_3$: DMSO-$d_6$): $\delta$ = 8.73 (s, 1H), 8.11 (d, 1H, $J = 4.0$), 7.10 (d, 1H, $J = 4.0$), 3.47-3.43 (m, 8H), 3.04 (s, 4H), 2.73 – 2.78 (m, 4H), 2.08 (s, 4H), 1.95 (s, 4H). MS (MALDI-TOF) (m/z) Calculated for C$_{31}$H$_{30}$N$_2$O$_7$S$_2$ [M+H$^+$]: 607.16, observed: 607.47.
**1.5.1.12 Molecule 1-13 Sulforhodamine 101 Sulfonyl Chloride (Texas Red)**

Molecule 0-13 Sulforhodamine 101 Sulfonyl Chloride (Texas Red)

![Chemical structures of para-sulfonyl chloride and ortho-sulfonyl chloride](image)

1-13 (mixed isomers)
Chemical Formula: C₃₁H₂₉ClN₂O₆S₂⁻
Exact Mass: 624.12
Molecular Weight: 625.16

Commercially-Available as a mix of isomers or made as follows:

In neat phosphorus (V) oxychloride (958µL, 10.25 mmol) was added sulforhodamine 101, free acid (191.4 mg, 0.32 mmol) at 0 °C under nitrogen atmosphere. After 30 minutes the reaction mixture was warmed to room temperature and allowed to stir for 12 hours. The mixture was brought to 4 °C with icy water and extracted three times with chloroform. The organic layers were combined and washed with ice cold water and dried over anhydrous sodium sulfate, filtered, concentrated and stored at -20 °C under argon to reveal a dark purple/blue solid as the crude intermediate, as a mixture of the ortho and para isomers of the monosulfonyl chloride derivatives, (154.1 mg, 78.1% yield). Rf=0.44 in 15% isopropanol in dichloromethane.

Molecule 1-13: Yield= 154.1 mg, 78.1%. Rf=0.44 in 15% isopropanol in dichloromethane. ¹H NMR (400 MHz, Chloroform-d): δ 8.94 (s, 1H), 8.65 (s, 1H), 8.02 (s, 5H), 7.24 – 7.14 (m, 1H), 6.41
(s, 2H), 4.29 – 4.05 (m, 4H), 3.53 (s, 8H), 3.05 (s, 4H), 2.72 (d, \( J = 23.8 \) Hz, 4H), 2.13 (s, 4H), 2.00 (s, 4H), 1.49 (t, \( J = 7.0 \) Hz, 2H), 1.41 – 1.22 (m, 8H).

### 1.5.1.13 Molecule 1-14

**Molecule 0-14**

[Chemical structure image]

**1-14** (mixed isomers)

Chemical Formula: \( C_{37}H_{43}N_3O_9S_2 \)

Exact Mass: 737.24

Molecular Weight: 737.88

To a pre-dried vial containing a stir bar under argon gas was added sulforhodamine 101 sulfonyl chloride (51.5 mg, 0.08 mmol) to which was added 2 mL of anhydrous dichloromethane followed by the addition of \((\text{PEO})_3\)-monoamine (18.4 mg, 0.12 mmol). The product was allowed to react overnight and followed by MALDI-TOF MS and TLC (1:3 MeOH: DCM). An extraction in water with dichloromethane removed the hydrolyzed sulforhodamine 101 free acid and the organic layer was removed under reduced pressure and the purple solid purified by RP-HPLC.

Initial attempts to purify using flash column chromatography with methanol/DCM were not successful as the retention factor for disubstituted isomer is similar to the monoisomer. The product was run on a 10-55% Water: ACN gradient with 0.1% TFA over 40 minutes and the
desired fractions were combined and lyophilized to give a dark purple solid (17.1 mg, 21.5% yield).

Molecule 1-14: Dark purple solid. Yield: 17.1 mg, 21.5% yield. $^1$H NMR (400 MHz, Chloroform-$d$) δ= 8.85 (d, $J = 10.4$ Hz, 0H), 8.77 (s, 1H), 8.36 (d, $J = 7.8$ Hz, 1H), 7.99 (d, $J = 7.9$ Hz, 0H), 7.77 – 7.34 (m, 2H), 7.20 – 6.96 (m, 2H), 6.76 (d, $J = 27.1$ Hz, 2H), 6.38 (s, 1H), 3.77 (s, 3H), 3.69 – 3.55 (m, 9H), 3.54 – 3.32 (m, 9H), 3.14 (s, 2H), 3.08 – 2.91 (m, 4H), 2.74 (s, 3H), 2.35 (d, $J = 4.9$ Hz, 2H), 2.20 – 1.91 (m, 10H), 1.03 – 0.74 (m, 13H). MS (MALDI-TOF) (m/z) Calculated for C$_{37}$H$_{43}$N$_3$O$_9$S$_2$ [M+H$^+$]: 738.25, observed: 738.44.

1.5.1.14 Molecule 1-15

To a 5 mL pre-dried vial with a stir bar was added Molecule 1-14 (10.5 mg, 0.14 mmol) with 3 mL of anhydrous dichloromethane and pyridine (2.4 mg, 0.03 mmol) followed by the addition of
para-toluenesulfonyl chloride (3.2 mg, 0.017 mmol). The reaction mixture was monitored via MALDI-TOF MS over a period of 24 hours and the solvent removed under reduced pressure then the crude mixture was taken up into a 1:1 mixture of ACN: water and purified by RP-HPLC from 10-60% Water: ACN gradient over 40 minutes. The pure fractions were combined and lyophilized to produce a purple solid (1.5 mg, 12% yield).

Molecule 1-15: Dark purple solid. Yield: 1.5 mg, 12% yield. $^1$H NMR (400 MHz, Chloroform-$d$) $\delta$ 8.79 (s, 0H), 8.39 (d, $J = 9.4$ Hz, 1H), 8.00 (d, $J = 17.7$ Hz, 1H), 7.52 (s, 1H), 7.00 (s, 1H), 6.74 (s, 0H), 3.76 – 3.68 (m, 4H), 3.63 (d, $J = 28.4$ Hz, 3H), 3.45 (d, $J = 6.9$ Hz, 2H), 3.31 (s, 0H), 3.02 (s, 1H), 2.96 (s, 0H), 2.88 (s, 0H), 2.75 (s, 1H), 2.35 (s, 1H), 2.17 (s, 0H), 2.10 (s, 2H), 2.00 (s, 2H), 1.69 (s, 1H), 1.43 – 1.33 (m, 5H), 1.32 – 1.18 (m, 25H).

1.5.1.15 Molecule 1-16

Molecule 0-16

Chemical Formula: $C_{31}H_{29}Cl_2N_2O_5S_2^+$

Exact Mass: 643.09
Molecular Weight: 644.61
In neat phosphorus (V) oxychloride (492 mg, 0.05 mmol) was added sulforhodamine 101, free acid (33.1 mg, 0.0546 mmol) at 0 °C under nitrogen atmosphere. After 30 minutes the reaction mixture was warmed to room temperature and allowed to stir for 18 hours. The mixture was brought to 4 °C with icy water and extracted three times with CHCl₃. The organic layers were combined and washed with ice cold water and dried over anhydrous sodium sulfate, filtered, concentrated and stored at -20 °C under argon to reveal a dark purple/blue solid, (14.2 mg, 40% yield). Allowing this reaction to proceed for greater than 12 hours results in a higher percentage of the bis sulfonyl chloride. This highly reactive intermediate was utilized for experiments in the formation of Molecule and 1-18 and characterized when stable as Molecule 1-18. *Rf=*0.60 in 15% isopropanol in dichloromethane.

**1.5.1.16 Molecule 1-17**

To a solution of crude Texas Red (34.1 mg, 0.055 mmol) in anhydrous dichloromethane (2 mL) was added 2-[2-(2-azidoethoxy)ethoxy]ethanamine (12.2 mg, 0.082 mmol) at 0 °C under nitrogen atmosphere. The reaction mixture was allowed to come to room temperature and
allowed to stir for 15 hours. The reaction mixture was concentrated under removed pressure and purified by silica gel column chromatography (dichloromethane: MeOH= 100:5) to yield compound Molecule 1-17 (15.8 mg, 30%) as a purple solid. TLC Rf=0.33 in 15% isopropanol in dichloromethane.

Molecule 1-17: Dark purple solid. Yield: 15.8 mg, 30% yield. Rf=0.33 in 15% isopropanol in dichloromethane. $^1$H NMR (400 MHz, Chloroform-d) $\delta$ 8.86 – 8.65 (m, 1H), 8.37 (d, J = 9.2 Hz, 0H), 8.07 – 7.90 (m, 1H), 7.18 (dd, J = 7.9, 4.7 Hz, 1H), 6.79 (s, 1H), 6.70 (s, 1H), 5.61 (s, 1H), 3.79 – 3.57 (m, 13H), 3.43 (dddd, J = 33.5, 24.8, 16.1, 5.1 Hz, 18H), 3.12 (s, 2H), 3.00 (td, J = 13.2, 6.6 Hz, 4H), 2.73 (d, J = 6.0 Hz, 3H), 2.64 (d, J = 20.3 Hz, 2H), 2.13 – 2.05 (m, 5H), 1.26 (s, 6H). MS (MALDI-TOF) (m/z) Calculated for C$_{37}$H$_{42}$N$_{6}$O$_{8}$S$_{2}$ [M+H$^+$]: 763.26, observed: 763.35.

1.5.1.17 Molecule 1-18

Molecule 0-18

![Chemical Structure](image)

1-18

Chemical Formula: C$_{43}$H$_{55}$N$_{10}$O$_{9}$S$_{2}$

Exact Mass: 919.36

Molecular Weight: 920.09

To a solution of crude Molecule 1-13 (34.1 mg, 0.055 mmol) in dry DCM (2 mL) was added 2-[2-(2-azidoethoxy)ethoxy]ethanamine (9.5 mg, 0.055 mmol) at 0 °C under nitrogen atmosphere.
The reaction mixture was allowed to come to room temperature and allowed to stir for 15 hours. The reaction mixture was concentrated under removed pressure and purified by silica gel column chromatography (DCM: MeOH= 100:5) to yield Molecule 1-18 (15.8 mg, 38% yield) as a purple solid. Isolated as a side product in reactive desired to make Molecule 1-17. TLC Rf=0.58 in 15% isopropanol in dichloromethane.

Molecule 1-18: Dark purple solid. Yield: 15.8 mg, 38% yield. Rf=0.58 in 15% isopropanol in dichloromethane. $^1$H NMR (300 MHz, Chloroform-d) $\delta$= 8.97 (s, 1H), 8.34 (s, 1H), 8.24 (d, $J = 8.0$ Hz, 1H), 7.87 (d, $J = 8.3$ Hz, 1H), 7.34 (d, $J = 7.9$ Hz, 2H), 7.05 (d, $J = 20.7$ Hz, 2H), 6.67 (s, 2H), 6.39 (s, 1H), 3.79 – 3.59 (m, 10H), 3.59 – 3.33 (m, 22H), 3.08 (d, $J = 33.6$ Hz, 7H), 2.75 (s, 3H), 1.99 (s, 8H). MS (MALDI-TOF) (m/z) Calculated for C$_{43}$H$_{55}$N$_{10}$O$_9$S$_2^+$ [M+H$^+$]: 920.37, observed: 920.38.
1.5.1.18 Molecule 1-19

Molecule 0-19

![Molecule 0-19](image)

1-19 (mix of isomers)

Chemical Formula:
\[ \text{C}_{48}\text{H}_{54}\text{N}_{6}\text{O}_{11}\text{S}_{3} \]

Exact Mass: 986.30

Molecular Weight: 987.17

To a solution of 3-butynyl p-toluenesulfonate (1.7 mg, 0.007 mmol) in DMSO (1.1 mL) and water (0.7 mL) was added the azide, Molecule 1-17 (5.2 mg, 0.007 mmol) at room temperature. To this mixture was then added in the following order: copper sulfate (0.85 mg, 0.0034 mmol), sodium ascorbate (3.38 mg, 0.017 mmol), and tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine, TBTA (9.05 mg, 0.017 mmol). The reaction was stirred overnight and concentrated under removed pressure followed by solvent removal via lyophilization. The crude solid was purified by RP-HPLC using a 10-62% Water: ACN gradient over 40 minutes to yield the desired analogue as a purple solid. All compounds but the azide were taken up into DMSO or water and thoroughly mixed and added to the reaction mixture to obtain an accurate addition of the small masses used in this reaction.  Yield=10.62 mg, 64.3% yield

Molecule 1-19: Dark purple solid. Yield: 10.6 mg, 64.3% yield. \(^1\)H NMR (400 MHz, DMSO-\(d_6\)) \(\delta\)

8.24 – 7.75 (m, 3H), 7.69 (s, 2H), 7.47 (dd, \(J = 32.9, 6.8\) Hz, 2H), 7.37 – 7.03 (m, 4H), 5.45 – 5.25
(m, 1H), 5.05 (d, J = 22.5 Hz, 3H), 3.51 (s, 7H), 2.37 – 1.88 (m, 10H), 1.59 (d, J = 32.7 Hz, 11H).

MS (MALDI-TOF) (m/z) Calculated for C_{48}H_{54}N_{6}O_{11}S_{3} [M+H]^+: 987.31, observed: 987.54.

1.5.1.19 Molecule 1-20

Molecule 0-20: 3-(ethylcarbamoyl)benzene-1-sulfonyl chloride

\[
\begin{array}{c}
\text{O} \\
\text{N} \\
\text{H} \\
\text{O=S=O} \\
\text{Cl}
\end{array}
\]

1-20

Chemical Formula: C_{9}H_{10}ClNO_{3}S

Exact Mass: 247.01

Molecular Weight: 247.70

To a pre-dried 1 mL vial containing a stir bar was added 3-chlorosulfonyl benzoyl chloride (405.0 mg, 1.694 mmol) in 1016 µL of DCM to which was drop-wise added a solution containing DIEA (275.9 mg, 2.14 mmol) and diethylamine (2M in THF, 423.5µL, 0.847 mmol) over a period of sixty minutes at 0°C under argon. The reaction mixture was then allowed to warm to room temperature and stirred for two hours. The solvent was removed under decreased pressure and purified by column chromatography using 2:3 ethyl acetate to hexanes. The reaction mixture was an oily solid. By TLC using the same solvent system there was one main spot at 0.4 and another lighter spot around 0.15 with a baseline spot for the DIEA. The oil was only slightly soluble in EtOAc so the crude product was taken up into DCM and loaded onto a flash column with a static gradient. The slightly yellow solid was dried under pressure and then stored at 0°C under nitrogen, (124.2 mg, 59.2% yield).
Molecule 1-20: 3-(ethylcarbamoyl)benzene-1-sulfonyl chloride. Yellow solid. Rf=0.4 in 2:3 ethyl acetate: hexanes. Yield: 124.2 g, 59.2%. $^1$H NMR (400 MHz, Chloroform-d) $\delta$ 8.39 (s, 1H), 8.17 (dd, $J = 21.2, 7.9$ Hz, 2H), 7.71 (t, $J = 7.9$ Hz, 1H), 6.52 (s, 1H), 3.54 (p, $J = 7.2$ Hz, 2H), 1.31 (s, 3H).

1.5.1.20 Molecule 1-21

Molecule 0-21: But-3-yn-1-yl-3-(ethylcarbamoyl)benzene sulfonate

\[
\text{\includegraphics[width=0.2\textwidth]{molecule_121.png}}
\]

Chemical Formula: $\text{C}_{13}\text{H}_{15}\text{NO}_4\text{S}$

Exact Mass: 281.07

Molecular Weight: 281.33

To a pre-dried 5 mL vial with a stir bar was added Molecule 1-20 (124.2 mg, 0.504 mmol), 1 mL of anhydrous DCM and 3-butyn-1-ol (42.3 mg, 0.60 mmol), DMAP (9.2 mg, 0.755 mmol) and DIEA (162.7 mg, 1.26 mmol). The reaction was allowed to stir under argon for 6 hours and the reaction monitored by TLC using 1:1 EtOAc: Hex. The product was detected on MALDI-TOF as well and the solvent removed under reduced pressure and purified using the same TLC conditions via flash chromatography to reveal a slightly tan solid, (88.6 mg, 65.8% yield).

Molecule 1-21: but-3-yn-1-yl 3-(ethylcarbamoyl)benzenesulfonate. Off-white solid. Yield: 88.6 mg, 65.8%. $^1$H NMR (400 MHz, Chloroform-d) $\delta$ 8.39 (s, 1H), 8.17 (dd, $J = 21.2, 7.9$ Hz, 2H), 7.71
(t, J = 7.9 Hz, 1H), 6.52 (s, 1H), 3.54 (p, J = 7.2 Hz, 2H), 1.31 (s, 3H). MS (MALDI-TOF) (m/z)

Calculated for C\textsubscript{13}H\textsubscript{15}NO\textsubscript{4}S [M+H\textsuperscript{+}]: 282.08, observed: 282.42.

1.5.1.21 Molecule 1-22

Molecule 0-22

\[
\text{Molecule 1-22 (mix of isomers)}
\]

Chemical Formula: C\textsubscript{50}H\textsubscript{57}N\textsubscript{7}O\textsubscript{12}S\textsubscript{3}
Exact Mass: 1043.32
Molecular Weight: 1044.22

To a solution of 3-ethylamido-butynyl p-toluene sulfonate (5.4 mg, 0.042 mmol) in DMSO (1.1 mL) and water (0.7 mL) was added the azide, Molecule 1-17 (16.8 mg, 0.022 mmol) at room temperature. To this mixture was then added in the following order: copper sulfate (2.8 mg, 0.0034 mmol), sodium ascorbate (10.9 mg, 0.055 mmol), and tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine, TBTA (27.5 mg, 0.55 mmol). The reaction was stirred overnight and concentrated under removed pressure then followed by solvent removal via lyophilization. The
crude solid was purified column chromatography using a two-phase gradient starting with 5% methanol solution in dichloromethane then step-wise to a 10% methanol solution in dichloromethane. The fractions collected gave two pure compounds; the first was the desired purple product (18.0 mg, 29.4% yield) and the hydrolyzed alcohol, Molecule 1-24 (11.7 mg). The two isomers of the product had an Rf= 0.27, 0.31 in 10% methanol in dichloromethane.

Molecule 1-22: Purple solid. Yield: 18.0 mg, 29.4%. The two isomers of the product had an Rf= 0.27, 0.31 in 10% methanol in dichloromethane. 1H NMR (400 MHz, DMSO-d6) δ 8.79 (s, 1H), 8.40 (d, J = 6.1 Hz, 1H), 8.28 (s, 1H), 8.20 (d, J = 7.8 Hz, 1H), 8.13 – 8.05 (m, 1H), 8.04 – 7.86 (m, 3H), 7.79 (d, J = 13.7 Hz, 1H), 7.72 (t, J = 7.8 Hz, 1H), 7.38 – 7.26 (m, 2H), 6.97 (s, 0H), 6.74 (s, 0H), 6.61 – 6.48 (m, 2H), 5.59 (s, 1H), 4.48 – 4.34 (m, 2H), 4.29 (q, J = 6.3 Hz, 2H), 3.73 (dt, J = 21.7, 5.4 Hz, 2H), 3.63 (s, 1H), 3.39 (d, J = 5.5 Hz, 12H), 3.20 (s, 3H), 3.07 – 2.85 (m, 9H), 2.77 (d, J = 5.7 Hz, 1H), 2.01 (s, 4H), 1.80 (d, J = 11.2 Hz, 4H), 1.60 (s, 3H). MS (MALDI-TOF) (m/z) Calculated for C50H57N7O12S3 [M+H+]: 1044.34, observed: 1044.14.
1.5.1.22 Molecule 1-23

Molecule 0-23

![Chemical Structure](image)

1-23 (mix of isomers)

Chemical Formula: C₅₀H₆₁N₉O₁₀S₃

Exact Mass: 1043.37

Molecular Weight: 1044.27

To a solution of propargyl biotin (8.86 mg, 0.032 mmol) in DMSO (1.3 mL) and water (0.9 mL) was added the azide, Molecule 1-17 (20.0 mg, 0.026 mmol) at room temperature. To this mixture was then added in the following order: copper sulfate (3.3 mg, 0.013 mmol), sodium ascorbate (13.0 mg, 0.066 mmol), and tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine, TBTA (34.8 mg, 0.066 mmol). The reaction was stirred overnight and concentrated under reduced pressure then followed by solvent removal via lyophilization. The crude solid was purified by RP-pTLC using 5% MeOH in DCM. The fractions collected gave the product as a purple solid, the (19.0 mg, 69.3% yield).

Molecule 1-23: Purple solid. Yield: 19.0 mg, 69.3. ¹H NMR (400 MHz, Chloroform-d) δ= 8.86 (s, 1H), 7.86 (s, 1H), 7.17 (d, J = 7.9 Hz, 2H), 6.75 (d, J = 5.3 Hz, 3H), 6.16 (s, 1H), 5.63 – 5.44 (m, 3H), 4.50 (s, 4H), 4.47 – 4.39 (m, 1H), 4.25 (dd, J = 19.5, 5.4 Hz, 2H), 3.90 (d, J = 5.1 Hz, 3H), 3.58 (s, 9H), 3.53 – 3.34 (m, 12H), 3.30 (d, J = 5.0 Hz, 3H), 3.13 – 2.91 (m, 8H), 2.83 (d, J = 5.0 Hz, 1H),
2.72 – 2.52 (m, 6H), 2.33 (d, J = 14.7 Hz, 1H), 2.05 (s, 9H), 2.00 (s, 17H). MS (MALDI-TOF) (m/z)

Calculated for C₅₀H₆₁N₉O₁₀S₃ [M+H⁺]: 1044.38, observed: 1044.15.

### 1.5.1.23 Molecule 1-24

**Molecule 0-24**

![Molecule Diagram](Image)

1-26 (mix of isomers)

Chemical Formula: C₄₁H₄₈N₆O₉S₂

Exact Mass: 832.29

Molecular Weight: 832.98

To a solution of 3-ethylamido-butynyl p-toluene sulfonate (5.4 mg, 0.042 mmol) in DMSO (1.1 mL) and water (0.7 mL) was added the azide, Molecule 1-17 (16.8 mg, 0.022 mmol) at room temperature. To this mixture was then added in the following order: copper sulfate (2.8 mg, 0.0034 mmol), sodium ascorbate (10.9 mg, 0.055 mmol), and tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine, TBTA (27.5 mg, 0.55 mmol). The reaction was stirred overnight and concentrated under removed pressure then followed by solvent removal via lyophilization. The crude solid was purified column chromatography using a two-phase gradient starting with 5% methanol solution in dichloromethane then step-wise to a 10% methanol solution in
dichloromethane. The fractions collected gave two pure compounds; the first Molecule 1-22 (18.0 mg, 29.4% yield) and the hydrolyzed alcohol, Molecule 1-24 (11.7 mg). The two isomers of the hydrolyzed product had an Rf= 0.17, 0.21 in 10% methanol in dichloromethane.

Molecule 1-24: Purple solid. Yield: 11.7 mg. $^1$H NMR (400 MHz, Chloroform-d). MS (MALDI-TOF) (m/z) Calculated for C$_{41}$H$_{48}$N$_6$O$_9$S$_2$ [M+H$^+$]: 833.30, observed: 833.11.
Covalent labeling of biomolecules of interest using fluorescent electrophiles

1.6 Introduction

This chapter reviews the methods and results of utilizing our electrophilic fluorescent molecules discussed in chapter one to covalently label proteins of interest. Very little has been accomplished in the way of combining small molecule covalent labeling ex vivo intracellularly with a highly specific short peptide interaction without the use of metal ligands.

Affinity in conjunction with a proximity effect to provide the selectivity necessary to site selectively label proteins of interest was our ultimate goal. Our research was innovative, in our opinion, because it attempted to create a new methodology by which to identify highly specific small molecule/protein interactions.

1.6.1 Labeling proteins with electrophilic small molecules

Development of an intracellularly active small molecule tag with high selectivity for biomolecules of interest is an ongoing challenge and would provide a powerful tool in cellular biology. In particular, protein-protein interactions play key roles in various human diseases but elucidation of protein interfaces is often difficult and can be disrupted by large protein tags. Therefore, site selective attachment using small molecules to specific proteins is a primary goal. Conventional labeling techniques have utilized large domain tags to study biomolecules but
there have been numerous developments in the discovery of small labeling molecules (SLM).

Despite the numerous techniques currently available, there are key limitations to each technique that fundamentally limit utility of that technique in differing cellular environments.

Desired elements of a good chemical tag include small size, specificity of labeling to the protein of choice in the presence of a complex mixture of other biomolecules, appropriate kinetics and quantitative labeling, cell permeability for intracellular labeling, minimal cytotoxicity and should be reactive in the desired cellular environment (oxidizing versus reducing)\textsuperscript{37}, Scheme 0-1. The trend towards chemical labeling using small peptide tags is due to varying restrictions of size interference by larger domain tags when studying POI. The potential to disrupt protein folding or function is increased when using larger tags, especially when studying the POI in a native environment\textsuperscript{96}. Desired elements in small synthetic probes
In general, labeling techniques have been classified by their size (domain tag versus peptide aptamer), binding properties (covalent versus non-covalent), and by the types of cellular proteins that can be labeled (membrane versus intracellular). In addition, the cellular environment (oxidizing versus reducing) of the POI dictates the technique best employed. For example, the cytosol is a reducing environment and useful for bis-arsenical labeling but single chain antibody techniques (though not covalent) have been found useful in oxidizing environments such as in the endoplasmic reticulum (ER) and cell surfaces. There have been numerous advancements in terms of development of domain tags used to label cells but there remains an inherent challenge to find highly selective recognition between a small molecule and a short amino acid sequence found within a protein of interest. As such, fewer...
techniques exist for small peptide tags\textsuperscript{99} and even fewer for those that incorporate covalent binding without enzyme mediated binding or metal based ligands. Affinity technology is limiting because built in ligands with desired reactive proteins can block active sites and may diminish protein activity in its native state\textsuperscript{96}. For example, ligand binding could have allosteric effects that result in unwanted functional changes, which may include affinity for substrate and changes in specificity.

**Chart 0-1 Comparison of labeling techniques**

<table>
<thead>
<tr>
<th>Type of Tag</th>
<th>Type of Binding</th>
<th>Mechanism</th>
<th>Type of Binding</th>
<th>Mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peptide</td>
<td>Non-covalent</td>
<td>Non-Enzymatic</td>
<td>Enzymatic</td>
<td>Covalent</td>
</tr>
<tr>
<td>Protein/Domain Tag</td>
<td>Enzymatic</td>
<td>Enzymatic</td>
<td>Non-Enzymatic</td>
<td>Enzymatic</td>
</tr>
</tbody>
</table>

Whereby EC=Extracellular, IC=Intracellular, AA= Amino acids

Overall, these new labeling techniques have become powerful tools that have enhanced our understanding of protein function in living cells. These techniques have aided in studying protein-protein interactions\textsuperscript{99}, protein association/ protein folding\textsuperscript{100–102}, small molecule sensors (Ca\textsuperscript{2+}\textsuperscript{103}, H\textsubscript{2}O\textsubscript{2} \textsuperscript{104}, pH sensors \textsuperscript{105}, protein localization, dynamics and trafficking\textsuperscript{106–112}. Our goal was to transcend the current limitations associated with larger domain tags and covalently label proteins of interest intracellularly using a small labeling molecule with the flexibility of using a probe of choice.
1.6.2 Proximity effect and affinity to label proteins of interest

One methodology used to label proteins has been to utilize a proximity effect whereby a component of a small labeling molecule (SLM) has a high affinity for a region of the protein of interest. Once the SLM is spatially proximal to the protein, a neighboring nucleophile on the protein can recognize and attack the electrophilic leaving group built into the SLM and the protein is then covalently bound to the SLM. Utilizing affinity between the two components to bring the two molecules together maximizes the chance that a covalent reaction can take place and maximizes selectivity between the SLM and the desired protein in a complex mixture.

Utilizing this proximity effect has been beautifully demonstrated by Hamachi and coworkers 50,89,113,114 called Ligand directed Tosyl (LDT) whereby they used human carbonic anhydrase II as their target and functionalized its known ligand with a sulfonate ester linked to a probe of choice (fluorophore or Biotin) and found that the result of bringing together two molecules (enzyme/ligand) whereby one carries an electrophile permitted nucleophilic attack by a neighboring amino acid and the enzyme was then covalently bound to the probe of choice, Scheme 0-2.
Utilizing a proximity effect that results in covalent bonding between a small molecule and a peptide intracellularly has been a goal that is challenged by selectivity. In one study, Fenical and coworkers built in a leaving group to Marinopyrrole A\textsuperscript{115} and reacted it with fractionated cell extracts followed by LC/MS/MS protein identification to determine which proteins interacted with their natural product. The result of their design led to a methodology that could potentially be used to screen protein target determination and affinity for a natural product. The natural product was built into the SLM as both the affinity portion of the molecule and as the leaving group. Ligation directed acyl transfer resulted in the displacement of the Marinopyrrole A and simultaneous covalent formation of the labeled protein with a fluorescent dye. Marinopyrrole A was found to selectively bind to actin and actinin. In the case of Marinopyrrole A, the source of high affinity interactions between a natural product to its corresponding protein was identified by varying the protein it interacted with and this could inherently be useful in target therapy for drug development.
Incorporation of affinity and covalent labeling in a single step without permanently blocking the active site or genetically modifying the protein of interest has been achieved by Hamachi and coworkers\textsuperscript{13}. In their studies, they took proteins with known affinities to specific ligands and modified the ligands with a tosylate ester as a leaving group, allowed the two to interact and proved that neighboring nucleophiles within the protein could covalently bind the other end of the ligand (containing the probe) whilst kicking off the ligand as the leaving group. They mixed and matched the proteins and ligands and found that the natural affinity between the complementary pair drove the chemical reaction that led to covalent binding of the protein to the probe in a highly selective manner, allowing the affinity to satisfy the target selectivity requirement and achieved this \textit{ex vivo} and in whole cell studies.

These current approaches require a full-size protein with an active binding site for covalent binding. Currently there are very few aptamers known to bind to small molecules. Those that do exist utilize metal ligands that can be toxic. For this reason, we were greatly interested in the research of Nolan and coworkers when they identified a ligand with a strong, non-covalent affinity to a peptide aptamer\textsuperscript{47} that was selected using phage display against Texas red fluorophore and applied in live cells as a highly sensitive calcium sensor. The strong ligand they identified was on the order of picomolar binding affinity on phage and nanomolar affinity for solution phase binding. What is striking about these results is that the aptamers are fairly small in size and it is remarkable that their binding affinities were so strong to the xanthene backbone of sulforhodamine 101. In addition, they found other aptamers with binding affinities to other fluorescent molecules such as Oregon green 514, rhodamine red, and fluorescein\textsuperscript{15} in addition to sulforhodamine 101. The conclusions from these studies provided
evidence that peptides could be identified that can bind to small molecule fluorophores with high affinity and future screens could lead to the identification of additional dye-binding peptides.

1.7 Overall rationale and objective for labeling proteins

Our goal was to establish a method by which an electrophilic small molecule would covalently label a protein of interest ex vivo by utilizing our highly specific ligand/protein interaction without occupying the active site of that protein. A proof of concept was tested to covalently label a model protein containing a peptide tag with a known affinity for a fluorophore, thus establishing proximity for reactivity of neighboring reactive nucleophiles within the protein, Scheme 0-3. There are two major components to these studies, the protein containing a peptidic insert with a high affinity for Texas Red and the electrophile containing a leaving group and a fluorophore.

Scheme 0-3 Labeling EF-Tu with a fluorescent tosylate

![Scheme 0-3 Labeling EF-Tu with a fluorescent tosylate](image-url)
1.7.1 Fluorescence polarization theory

Having chosen the proteins of interest containing aptamers that would bind to our fluorescent electrophiles we chose to use fluorescent polarization (FP) as a tool for determining the binding affinity between the two entities. Fluorescent polarization or fluorescence anisotropy is a technique that has been used to monitor the interaction between a small fluorescent molecules with a protein\textsuperscript{116}. When a fluorescent molecule is exposed to light it emits that light in the same direction as it absorbed. However, in solution the molecule tumbles about freely and emits light in all directions and the emitted light is not polarized. If however, the molecule interacts with another larger molecule (such as a protein) the tumbling slows down and the emission becomes polarized\textsuperscript{117}. The amount of polarization is dependent on the degree to which the tumbling fluorophore is bound to the protein and this permits a quantitative assessment of the affinity between the two molecules. Polarization ratio ($p$) is interchangeably used with emission anisotropy ($r$) whereby the intensity measurements of the emission signal are made parallel (\textit{I}||) or perpendicular to the light source (\textit{I}⊥). If all of the light is emitted parallel to the light source, then $r$ and $p$ would be equal to 1. If all of the light is emitted perpendicular to the light source then $p$ would be equal to -1 and $r$ would be equal to -0.5. The equations that relate the terms anisotropy ($r$) and polarization are given in Equation 2-1.

\textbf{Equation 0-1 Fluorescence anisotropy and polarization}

\begin{align*}
\text{Polarization ratio (p)} &= \frac{\text{I}|| - \text{I}⊥}{\text{I}|| + \text{I}⊥} \\
\text{Emission Anisotropy (r)} &= \frac{\text{I}|| - \text{I}⊥}{\text{I}|| + 2\text{I}⊥} \\
\text{Relationship between r and p:} & \quad r = \frac{2p}{(3-p)} \\
& \quad p = \frac{3r}{(2+r)}
\end{align*}
In general, this technique makes use of the unequal intensities of light emitted from a fluorophore along different axes of polarization. As the fluorophore tumbles in solution this polarization is greater it is bound to a protein that in turn correlates to the degree of binding between the two entities that can be quantitatively measured.

While fluorescence polarization is often utilized to measure the binding affinity between a small fluorophore and a large biomolecule it is not often specifically used to monitor covalent labeling between the two molecules. One aspect to note in the experimental design of fluorescence polarization is that the protein is at a high concentration (high nano to micromolar) whilst the fluorophore is kept at a constant low concentration (5-10 nM). The fluorophore is kept very low in terms of concentration because it is the entity that is being measured, and as such, the reaction is being pushed to determine what percentage of the fluorophore is interacting with a large abundance of protein. If covalent labeling is occurring at these small concentrations then the fluorophore is tumbling in space as it would in a binding situation, which is a much slower rate than if it were not interacting. We assume that the method of interaction (covalent versus binding) results in the same polarization effects.

1.8 Results

As described in Chapter 1 we developed a fluorescent electrophile containing three key components, an electrophilic sulfonate ester, a PEG linker and a fluorescent small molecule known to have binding affinity to aptamers 512 and 501. Therefore, two model proteins were chosen (known to have high expression yields) and the two proteins were mutated to contain those inserts. A summary of the proteins chosen with the peptide sequences that were
inserted are shown in Table 0-1 and the experimental methods used to produce these proteins are provided in Section 2.5, page 111.

### Table 0-1 Protein mutations containing 512 or 501 aptamers

<table>
<thead>
<tr>
<th>Name of Peptide Inserted into Protein</th>
<th>Peptide Sequence Inserted</th>
<th>Protein Expressed</th>
<th>Size (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wildtype</td>
<td></td>
<td>EF-Tu</td>
<td>43.3</td>
</tr>
<tr>
<td>501</td>
<td>GGGSKVILFEGPAGRTIWEPEASNHTGAPGSKVILFEGGPG (42AA)</td>
<td>EF-Tu</td>
<td>48.3</td>
</tr>
<tr>
<td>512WPE</td>
<td>GGGSKVILFEGPAGRTWPEISEGAPGSKVILFEGGPG (38AA)</td>
<td>EF-Tu</td>
<td>47.9</td>
</tr>
<tr>
<td>512</td>
<td>GGGSKVILFEGPAGRTWPEISEGAPGSKVILFEGGPG (38AA)</td>
<td>EF-Tu</td>
<td>47.9</td>
</tr>
</tbody>
</table>

The proteins described in Table 0-1 were incubated with our fluorescent electrophiles and the resulting binding and/or covalent labeling was established via the following methods: The labeled proteins were captured on Ni-NTA Agarose beads via their affinity tag and qualitatively assessed to determine if there was covalent binding (Section 2.3.1). Binding affinities between the proteins and fluorophores were quantitatively assessed using fluorescence polarization (Section 2.3.2). Finally, covalent labeling of the proteins was determined via purification on an SDS page gel followed by fluorescence imaging (Section 2.3.3). These series of experiments allowed us to determine if there was binding between the two components, if that binding was covalent, and it helped direct our synthetic strategies in terms of modulating the electrophilicity of our fluorophore such that we could optimize the labeling.

#### 1.8.1 Labeling of EF-Tu-His Tag with Nickel NTA beads experiment

A qualitative test to monitor the affinity of the protein containing the 512 insert was accomplished by capturing the EF-Tu and 512 EF-Tu containing His tags on Ni-NTA beads then mixed with varying concentrations of sulforhodamine 101 (Molecule 1-12, SR101), washed thoroughly and visualized using a Texas red fluorescence filter on a microscope. As seen in
Figure 0-1, the control beads that were not exposed to protein or dye did not fluoresce (Panel A-TRF). However, the beads that did not capture protein did fluoresce at higher concentrations of Molecule 1-12 (Panel B-TRF). Therefore, background staining is problematic if the concentration of the dye is too high. For those beads that captured EF-Tu and 512-EF-Tu there was background staining at the higher concentration of Molecule 1-12 (Panels D and F). However, at a lower concentration there was less binding of the dye to EF-Tu (Panel E-TRF) compared to 512 EF-Tu (Panel G-TRF). This suggested to us that there was some level of affinity between the 512 aptamer and sulforhodamine 101 (Molecule 1-12) and that using Ni-NTA beads provides a technique to qualitatively assess this affinity. This data provided early evidence that led us to believe that the binding between 512 aptamer and various fluorophores was worth further pursuit.
1.8.2 Evaluation of binding via fluorescence polarization

Once we established binding using Nickel-NTA beads we chose to quantitatively assess the binding using fluorescence polarization on a number of fluorophores. We tested both the
rhodamine B electrophiles and the sulforhodamine 101 electrophiles along with their corresponding non-electrophilic fluorophores as controls. It is important to note that we completed multiple time studies to see if the binding between the fluorophore was time dependent. FP measurements were taken at 30 minutes to 24 hours and we did not see a significant effect on the binding results. This suggested to us that the interaction of the small concentration of fluorophore were quick to equilibrate with the high concentration of proteins and that if the result of the binding was covalent bonding, this molecular change did not greatly affect the polarization results. As these reactions took place in buffers containing salts we were unable to utilize MALDI-TOF to monitor the covalent labeling of these reactions.

1.8.2.1 **FP of rhodamine B analogs with EF-Tu proteins containing peptidic inserts**

In the first set of experiments, three proteins were tested: EF-Tu wild type without a peptidic insert, 512 EF-Tu containing peptidic insert 512, and containing peptidic insert 501. These were each tested against the rhodamine B derivatives described in Table 0-2 whose route of synthesis was described in Section 1.3.1, pages 15-22.
Table 0-2 Rhodamine analogs tested for affinity

<table>
<thead>
<tr>
<th>Molecule</th>
<th>Name</th>
<th>Peptide Insert</th>
<th>Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Rhodamine B</td>
<td>501 EF-Tu</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>512 EF-Tu</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>– EF-Tu</td>
<td></td>
</tr>
<tr>
<td>1-11</td>
<td>RhB Piperazine Amide Benzophenone</td>
<td>501 EF-Tu</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>512 EF-Tu</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>– EF-Tu</td>
<td></td>
</tr>
<tr>
<td>1-9</td>
<td>RhB Piperazine Amide Propanol</td>
<td>501 EF-Tu</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>512 EF-Tu</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>– EF-Tu</td>
<td></td>
</tr>
</tbody>
</table>

The propanol derivative showed negligible binding affinity with all EF-Tu proteins. The benzophenone derivative showed binding affinity in the low micromolar range but did not show specificity between EF-Tu wild type and the EF-Tu protein containing the peptidic inserts. Molecule 1-1 (Rhodamine B) as a control without an electrophile was tested and modest affinity was observed, Chart 0-2.

Chart 0-2 Rhodamine B (Molecule 1-1) binding via FP
However, polarization values typically range from 0.33 to 0.55 and the range for anisotropy is -0.25 to 0.4118. When evaluating the data we noted the values of the polarization were within a very narrow range (0.097-0.130) and this led us to question the data as the EC₅₀ data could correlate to a local maxima and at higher concentrations of protein the polarization values would be much higher than background. We changed the y-axis and found that there was no detectible binding, Chart 0-3. The lack of binding observed in rhodamine B (Molecule 1-1) led us to synthesize another rhodamine analog, sulforhodamine 101, the specific xanthene backbone tested in the work by Nolan and co-workers.

Chart 0-3 Rhodamine B (Molecule 1-1) FP values
1.8.2.2 FP of sulforhodamine 101 analogs with proteins containing peptidic inserts

After further testing with rhodamine B analogs we decided to test the binding affinities of sulforhodamine derivatives instead, Chart 0-4.

Chart 0-4 Sulforhodamine 101 analogs affinity binding testing with proteins containing peptidic inserts and peptides without proteins

<table>
<thead>
<tr>
<th>Molecule</th>
<th>Name</th>
<th>Peptide Insert</th>
<th>Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-12</td>
<td>Sulforhodamine 101</td>
<td>512</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>512WPE</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>512</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>501</td>
<td>-</td>
</tr>
<tr>
<td>1-19</td>
<td>Sulforhodamine 101 Tosylate</td>
<td>512</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>512WPE</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>EF-Tu</td>
</tr>
<tr>
<td>1-22</td>
<td>Sulforhodamine 101 Etam</td>
<td>501</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>512</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>EF-Tu</td>
</tr>
</tbody>
</table>

1.8.2.2.1 FP of sulforhodamine 101 with EF-Tu proteins containing peptidic inserts

After further testing with rhodamine B analogs we decided to test the binding affinities of sulforhodamine derivatives instead. First we tested sulforhodamine 101 (Molecule 1-12) and discovered that the binding affinity at best was approximately 50 µM for those proteins containing the 512 or 501 peptide, Chart 0-5. With low binding affinity the concentrations of the proteins needed to evaluate the EC50 needed to be in the high micromolar range. This created multiple problems. Firstly, in this concentration range a lot of protein needed to be
expressed and purified and this can be time intensive. Secondly, proteins begin to precipitate at higher concentrations that increases light scattering during FP measurements obscuring the results and creating falsely high FP values. Regardless there was not a remarkable difference between the correct 512 aptamer and the 512 WPE aptamer, and EF-Tu itself.

Chart 0-5 FP Data sulforhodamine 101 (SR101, Molecule 1-12) with proteins containing peptidic inserts

1.8.2.2.2 FP of sulforhodamine 101 tosylate with EF-Tu containing peptidic inserts

In addition we decided to test if our electrophilic sulforhodamine 101 analogs would show better binding affinity than Molecule 1-12. We did not find enhanced binding, Chart 0-6. Using
Molecule 1-19, the wild type EF-Tu affinity showed a binding affinity of 120.81 µM while our 512 EF-Tu protein had a weaker binding affinity of 191.72 µM.

**Chart 0-6 FP sulforhodamine 101 Tosylate with proteins containing peptidic inserts**

The high error values are attributed to the fact that we were unable to collect data at higher concentrations before precipitation became problematic and resulted in light scattering that introduces error into the polarization measurements. The lack of values at higher concentrations of protein (max, the top of the curve) increased the error associated with applying a four-parameter logistic curve Equation 0-2\(^{119}\).

**Equation 0-2 Four parameter logistic curve used to calculate EC\(_{50}\) values for FP binding affinity**

Data

\[
y = \min + \frac{(\max - \min)}{1 + (x/\text{EC}_{50})^{-\text{Hillslope}}}
\]

Where

- min- bottom of the curve
- max- top of the curve
- EC\(_{50}\)- half maximal effective concentration
- Hillslope- the slope of the curve at its midpoint

Therefore, based on the high error values due to light scattering we were unable to calculate statistically significant EC\(_{50}\) values for those curves requiring high protein concentrations, typically above 0.1 M.
1.8.2.2.3  FP of Molecule 1-22 with proteins containing peptidic inserts

Once we had tuned down the electrophilicity of the leaving group we chose to test another of our analogs, Molecule 1-22. Once again we found that the binding affinity for our electrophilic fluorophore showed low affinity for all Ef-Tu proteins, Chart 0-7. For this fluorophore the binding data was observed to be 129.19 μM for 501 EF-Tu. However, the curves were fit so poorly we were unable to get a statistically significant EC_{50} values for 512 EF-Tu and wild type EF-Tu. Overall, the binding values for this analog are no better than those observed for sulforhodamine 101 and the tosylate analog (Molecule 1-19).

Chart 0-7 FP Binding data for Molecule 1-22 with proteins containing peptidic inserts

1.8.2.2.4  FP of sulforhodamine 101 with peptides and no protein

To determine if the protein was the cause of low binding affinity and not the aptamer itself we synthesized both the 512 and 501 peptides and tested binding affinity using fluorescence polarization.
1.8.2.2.4.1 512 peptide FP studies

Fluorescence polarization studies of 512 peptide with Molecule 1-12, sulforhodamine 101 free acid was found to have an EC$_{50}$ of 42 µM, which was similar to the earlier results with the EF-Tu protein attached to the aptamer, Chart 0-8.

Chart 0-8 FP of 512 peptide with Molecule 1-12, sulforhodamine 101, SR101

In addition, we found that 512 peptide had only a slightly better affinity for our sulforhodamine 101 electrophile, Molecule 1-22 with an EC$_{50}$ of 34.3, Chart 0-9.

Chart 0-9 FP data of 501 peptide Molecule 1-22
At this point we became concerned that the buffer was disrupting the constrained peptide sequence used to enable cyclization of the aptamer we were using to run the experiments in, 1X PBS. Therefore, we tested the Molecule 1-12, SR101, with peptide 512 in three different solvent systems to see if pH and/or buffer components would affect the binding, Chart 0-10. We tested 1X PBS at pH 7.6, 50 mM HEPES at pH 8.1 and 50 mM Tris buffer at pH 8.3 and saw little to no difference in binding affinities, with a binding range of 40.93-55 µM.

**Chart 0-10 FP buffer study of 512 peptide with Molecule 1-12, sulforhodamine 101**

Anisotropy of 512 Peptide with Sulforhodamine 101 (20 nM) in Various Solvents

<table>
<thead>
<tr>
<th>Protein Concentration (µM)</th>
<th>1X PBS pH 7.6</th>
<th>50 mM HEPES pH 8.1</th>
<th>50 mM TRIS pH 8.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>EC50</td>
<td>40.93 ± 0.55</td>
<td>55.42 ± 6.44</td>
<td>43.53 ± 1.46</td>
</tr>
</tbody>
</table>

1.8.2.4.2 501 peptide FP studies

FP studies of the peptide 501 resulted in poor binding affinity such that we were unable to obtain an EC50 curve. Therefore, we conclude that 501 peptide has poorer affinity for Molecule 1-12 than the 512 peptide, Chart 0-11.
1.8.3 Labeling assays

Despite the fact that the fluorophores showed weak binding with the peptides that were reported to have picomolar binding affinity, we wanted to see if the binding might result in selective labeling of cell lysates containing our EF-Tu conjugates. We tested two of our electrophilic fluorophores, Molecule 1-19) and Molecule 1-23 using Molecule 1-12 as a control for both. Details with regards to the experimental design of the labeling experiments can be found in Section 2.5.5. In general, the proteins were mixed with the fluorophores and incubated at 37 °C, mixed with Laemmli Buffer containing BME and boiled for 5 minutes to denature the proteins, resolved on an SDS page gel, then visualized on a fluorescent scanner to ascertain which proteins, if any, were fluorescently labeled. When run on an SDS page gel all small molecule fluorophores were resolved to the bottom of the gel. When non-electrophilic sulforhodamine 101 fluorophores (Molecule 1-12, 1-14) were run on an SDS page gel with our proteins they did not label the proteins. In conjunction with the 5 minute boiling time used to
denature the proteins we were able to use this assay as a method to identify those proteins that had been covalently labeled with our electrophilic proteins.

An example of one of our preliminary control studies can be seen in Figure X.X. In these reactions, 512 EF-Tu, 501 EF-Tu and EF-Tu were each reacted with varying concentrations of either Molecule 1-19 or Molecule 1-12 as a control. The figure contains the gel for 512 EF-Tu as the results for 501 EF-Tu and EF-Tu were the same. Lanes 1-5 contain 512 EF-Tu reacted with 0.63-10 µM of Molecule 1-19, and there was labeling at 1.3-10 µM. Lanes 6-10 contain 512 EF-Tu reacted with 0.63-10 µM of Molecule 1-12, and there was no labeling. Lane 11 was a protein control lane containing only EF-Tu, and Lanes 12-13 were control lanes for the fluorophores with no protein present. The Coomassie stain of the gel showed that the 512 EF-Tu protein was pure, at equivalent concentrations for all reactions tested. The fluorescence scan of the gel showed that the dual color ladder was fluorescent and subsequently we chose not to use a protein ladder in our reaction studies with the fluorophores as they saturated the signal. (When the proteins were purified, the molecular weight was always confirmed with a dual color ladder). The other result of the preliminary study showed that the fluorophores run to the bottom of the gel and also saturate the signal. As such, the fluorophore was run off of the gel such that the small molecules were no longer present on the gel and only proteins that were covalently labeled with the electrophilic fluorophore remained and were therefore scanned with enhanced sensitivity.
Figure 0-2 Preliminary labeling studies used to validate covalent labeling of fluorescent electrophiles with electrophilic fluorophore (Molecule 1-19) and non-electrophilic fluorophore (control, Molecule 1-12)

<table>
<thead>
<tr>
<th>Lanes:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dual Color Ladder</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>S12 EF-Tu (5 uM)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Molecule 1-19 (uM)</td>
<td>10</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Molecule 1-12 Control (uM)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>10</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Fluorescence Scanner at Ex/Em: 580/610 nm 10% SDS Page Gel

Coomassie Stain

Lanes

1  2  3  4  5  6  7  8  9 10 11 12 13 14 15
1.8.3.1 Labeling with sulforhodamine 101 tosylate analog, Molecule 1-19

Initially, we tested the labeling of proteins 512 EF-Tu (lanes 2-7) and EF-Tu (lanes 9-14) at a constant concentration of 10 μM protein with decreasing concentrations of Molecule 1-19 (0-10 μM) for 24 hours. The fluorophore was shown to label both proteins in a concentration dependent manner without selectivity for the 512 EF-Tu over the wild-type EF-Tu, Figure 0-3. As the rate equation of an $S_N2$- type reaction is dependent on the concentration of the electrophile and the nucleophile we expected that halving the concentration of the electrophile would halve the rate of the reaction.

**Figure 0-3 Labeling of proteins using Molecule 1-19 shows concentration dependence**

<table>
<thead>
<tr>
<th>Molecule 1-19 uM</th>
<th>10</th>
<th>5</th>
<th>2.5</th>
<th>1.3</th>
<th>0.6</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluorescence Scanner at 580/610 nm</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

10% SDS Page Gel with a 24 hour incubation time.

The other variable we wanted to evaluate for labeling optimization was the amount of time necessary for incubation. Both 512 EF-Tu and EF-Tu were tested at a constant concentration of 10 μM with two concentrations of Molecule 1-19 and time points were taken at 0.25, 0.5, 1, 2,4, 8 and 24 hours. Labeling of the protein (512 EF-Tu top panel and EF-Tu bottom panel) began at approximately 30 minutes and increased in a time dependent manner, Figure 0-4.
As we did not see a large difference between the labeling of wild type EF-Tu and the 512 EF-Tu we decided to try labeling 512 EF-Tu and EF-Tu in the presence of other proteins to see if there was selectivity for the binding of 512 EF-Tu, Figure 0-5. The 512EF-Tu and EF-Tu concentrations were kept constant at 10 μM. Factor proteins used in mRNA display were varied with concentrations ranging from 1.25-20.25 μM (Table 0-3). These components were tested with a fluorophore concentration at 5 μM.
Table 0-3 Table Factor mix components for competitive labeling study using Molecule 1-19

<table>
<thead>
<tr>
<th>Factor:</th>
<th>Stock (µM)</th>
<th>Final (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EF-G</td>
<td>158.8</td>
<td>3.9</td>
</tr>
<tr>
<td>EF-Ts</td>
<td>294.9</td>
<td>9.9</td>
</tr>
<tr>
<td>IF1</td>
<td>339.8</td>
<td>20.3</td>
</tr>
<tr>
<td>IF2</td>
<td>49.5</td>
<td>3.0</td>
</tr>
<tr>
<td>IF3</td>
<td>131.8</td>
<td>11.3</td>
</tr>
<tr>
<td>MTF</td>
<td>61.6</td>
<td>4.5</td>
</tr>
<tr>
<td>RF1</td>
<td>106.5</td>
<td>2.3</td>
</tr>
<tr>
<td>RF3</td>
<td>285.5</td>
<td>1.3</td>
</tr>
<tr>
<td>RRF</td>
<td>142.9</td>
<td>3.8</td>
</tr>
</tbody>
</table>

We found that Molecule 1-19 labels proteins without discrimination (Figure 0-5) and as such, these results led us to believe that our electrophile may be too reactive and we chose to develop a slightly less electrophilic analog as described in Chapter One.
Figure 0-5 Competition labeling using Factor Mix with Molecule 1-19

<table>
<thead>
<tr>
<th>Factor Mix (1.25-20.25 uM)</th>
<th>+</th>
<th>+</th>
<th>+</th>
<th>+</th>
<th>+</th>
<th>+</th>
<th>+</th>
<th>+</th>
</tr>
</thead>
<tbody>
<tr>
<td>512 EFTU (10 uM)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>EFTU (10 uM)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Fluorophore (5 uM)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Time (hours)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Lanes 1-3 show the factor mix, 512 EF-Tu and EF-Tu without the fluorophore. As such, the proteins were not labeled and the gel shows no band using fluorescent imaging. Lanes 4-6 show the factor mix with the fluorophore at 3 time points and the experiment shows that the factor mix proteins are readily labeled by the electrophilic fluorophore and labeling increases.
with increasing time. Lanes 7-9 of the gel show factor mix with the fluorophore in addition to 512 EF-Tu. We had hoped that in the presence of 512 EF-Tu, the fluorophore would selectively label the 512 EF-Tu protein, as the fluorophore was the limiting reagent in the reaction. However, the components of the factor mix were stained along with the 512 EF-Tu. Lanes 10-12 also shows non-specific labeling when EF-Tu is added to the fluorophore, factor mix and 512 EF-Tu mixture. Finally, lanes 13-15 show a direct competitive labeling experiment between 512 EF-Tu and EF-Tu and the results show that the proteins are both readily fluorescently labeled.

1.8.3.2 Labeling Proteins with Molecule 1-22

In the first experiment we tested 512 EF-Tu (top gel) and EF-Tu (bottom gel) at concentrations of 5 μM with either Molecule 1-22 or the hydrolyzed version of Molecule 1-22 at varying time points at equimolar concentrations, Figure 0-6.
Figure 0-6  *Meta*-ethylamido sulfonate ester analog of sulforhodamine 101 (Molecule 1-22) and the hydrolyzed analog (Molecule 1-24)

*Meta*-ethylamido sulfonate ester analog of sulforhodamine 101 (Molecule 1-22) was shown to label both proteins in a time dependent manner without selectivity for the 512 EF-Tu over the wild type EF-Tu and the hydrolyzed analog of sulforhodamine 101 (Molecule 1-24).
24) did not label the proteins as expected, Figure 0-7. The top panel of Figure 2-6 shows 512 EF-Tu reacted with equimolar concentrations of either electrophilic Molecule 1-22 (Lanes 1-4 at 2,4,7 and 26 hours respectively) or the hydrolyzed isomer, Molecule 1-24 (Lanes 5-9 at 2,4,6 and 26 hours respectively). Lane 9 is a protein control lane containing no fluorophore. The bottom panel is the same as the top panel with EF-Tu wild type as the protein. Please note that in lane 2, the samples were switched between the gel, so the top panel lane 2 contains EF-Tu instead of 512 EF-Tu and the bottom panel contains 512-EF-Tu instead of EF-Tu as can be visualized by the band being lower in the top panel and higher in the lower panel. Electrophilic Molecule 1-22 labeled both 512 EF-Tu and EF-Tu. Molecule 1-24 was later found to be contaminated with a small amount of Molecule 1-22 (determined by MALDI-TOF) and as such, there is evidence of some labeling as visualized by the very faint bands in lanes 5-9. This experiment further showed that our fluorescent electrophile labeled proteins indiscriminately.
Figure 0-7 *Meta*-ethylamido sulfonate ester analog of sulforhodamine 101 (Molecule 1-22) labeling studies with 512 EF-Tu and EF-Tu

Finally we tested *Meta*-ethylamido sulfonate ester analog of sulforhodamine 101 (Molecule 1-22) in cell lysate to see if it would selectively label the 512 EF-Tu protein in the presence of a complex mixture of other biomolecules, Figure 0-8. The experiment was prepared such that
there was a control lane for each component, 512 EF-Tu Protein, EF-Tu protein, the fluorophore and the lysate. In addition, both proteins were mixed with the fluorophore with and without the presence of cell lysate. Lanes 11-14 were control lanes and did not show any labeling without the fluorophore nor did the fluorophore remain on the gel (Lane 15). Lanes 4-5 are the control lanes with the EF-Tu proteins alone. Lanes 7-10 contained cell lysate and Molecule 1-22 was shown to label both cell lysate and EF-Tu proteins when mixed together.
Figure 0-8 *Meta*-ethylamido sulfonate ester analog of sulforhodamine 101 (Molecule 1-22) labeling of cell lysate

<table>
<thead>
<tr>
<th></th>
<th>Lanes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell Lysate 12.5 2 mg/mL</td>
<td>+</td>
</tr>
<tr>
<td>512 EFTU 70 uM</td>
<td>+</td>
</tr>
<tr>
<td>EFTU 70 uM</td>
<td>+</td>
</tr>
<tr>
<td>Fluorophore 10 uM</td>
<td>+</td>
</tr>
</tbody>
</table>

10% SDS Page gel visualized at 580/610 nm of Molecule 1-22 used to label proteins and cell lysate. Lane 1: protein ladder, Lanes 2-3 empty.

1.9 Discussion of labeling results

While Nolan and coworkers reported picomolar affinities on phage and low nanomolar affinities with free peptide in their patent we did not find this level of binding affinity between their aptamers and sulforhodamine 101. Considering the LDT chemistry performed in vivo
containing a sulfonate ester moiety by Hamachi and coworkers and the ability of our
electrophilic fluorophores to label proteins, the next step for this project would be to
incorporate a better affinity partner for the fluorophore. Unfortunately, there is a large gap in
what is available in terms of small peptides with a specific affinity for small molecules.
Biologically speaking, large proteins with binding sites and pockets are much more effective at
producing microenvironments that optimize the conditions necessary for the reactivity in their
active sites. We contacted a large number of vendors attempting to find a monoclonal
antibody that we could use as a binding partner for sulforhodamine 101(Molecule 1-12). While
there are antibodies available for various rhodamine analogs, they specifically do not work for
sulforhodamine 101. We did however find that sulforhodamine 101 (Molecule 1-12) is used in
neurobiology as a label for astroglia and we decided to determine if the binding affinity for our
electrophilic fluorophore might covalently label this type of glial cells, see Chapter 4.

1.10 Experimental details and methods

1.10.1 Synthesis of 501/512 proteins (EF-Tu and GST)

1.10.1.1 Design and synthesis of the insert

146-171 bp ds DNA primers were designed and amplified, called Primer ABC, that encoded for
the overlapping regions of the plasmid of interest and the aptamer of choice, Table 0-4.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Aptamer</th>
<th>Primer ABC</th>
</tr>
</thead>
<tbody>
<tr>
<td>EF-Tu</td>
<td>512c</td>
<td>GACGACCCACCCACCCACCGTGGGAGTACGAAGGTGTAAGGGGATTTTTCGAAGGGCTCCGGCAGCGATGCCGTTCTAAATGGCTAGTAAAAATAGTCATCCTGTTGAGGGGCTGGAGGCCGGCCTGGCAGTAAAGTATATTATTCGAAGGTGGCAGAGGA TGAGATCCGGCTGCTAACAAAGCC</td>
</tr>
<tr>
<td>EF-Tu</td>
<td>501</td>
<td>CACCTGAGCACACACCACCCACCCACCGTGGGAGTTCAAAAGTCATTCCTGTTGAGGGGCTGGAGGCCGGCCTGGCAGTAAAGTATATTATTCGAAGGTGGCAGAGGA TGAGATCCGGCTGCTAACAAAGCCCGAAAGG</td>
</tr>
</tbody>
</table>
The PCR product was run on a 2% Agarose gel to validate purity and size, Figure 0-9.

**Figure 0-9 PCR#1 to make primers ABC for four inserts (2% Agarose Gel, Hyperladder I)**

2% agarose gel representing PCR primers ABC used to make the following protein inserts: Lane 1: DNA Ladder, Hyperladder V; Lane 2: 501 EF-Tu, 171 bp; Lane 3: 501 GST 158 bp; Lane 4: 512c EF-Tu, 159 bp; Lane 5: 512c GST, 146 bp; Lane 6: Negative Control.

**1.10.1.2 Proteins chosen for genetic modification**

Two proteins were chosen for testing. The first was elongation factor thermally unstable protein, EF-Tu for which PET-24(a) was used and EF-Tu was added at restriction sites Nhe I/Not I (166). Overlap extension PCR cloning\textsuperscript{120}, a variant of PCR used to insert specific mutations at specific points in a sequence was used to insert aptamers 501, 512, and 512c before the His Tag at the C terminus. The second protein modified was GST using plasmid pGEX-4T-a. The site of insertion of 501, 512, and 512c was just after the thrombin site at the C terminus. Both proteins are known to express well and contain affinity tags for purification following expression.
1.10.1.3 Protein methods and expression

Overlap extension PCR cloning was used as a method to create the desired recombinant plasmids. Primer ABC previously described underwent a second PCR amplification (Figure 0-10) to amplify the plasmid containing the insert. The new plasmid was then transformed into E. Coli after Dpnl Digest destroyed the parental plasmid.

Figure 0-10 PCR#2 recombinant plasmid (0.8% agarose gel, Hyperladder V)

0.8% agarose gel representing PCR product for the recombinant plasmid needed to make the following proteins with the respective peptidic inserts. Lane 1: 501 EF-Tu, Lane 2: Empty; Lane 3: 501 GST; Lane 4: 512c EF-Tu; Lane 5: 512c GST; Lane 6: MW marker, Hyperladder I.

The recombinant plasmids underwent transformation into MACH1 competent cells, the inserts were verified first by colony PCR and then by sequencing. The proteins were then induced using IPTG, expressed and isolated via affinity tags, dialyzed in enzyme storage buffer, and stored at -80 °C. For the EF-Tu proteins, solubility was an issue for all three inserts. To overcome this problem, induction took place at 18 °C overnight and a suitable concentration of EF-Tu obtained. Solubility did not prove to be a problem for expressing GST. Transformation of plasmid DNA was accomplished using competent E. Coli cells which provided the means to express 512 EF-Tu protein in working concentrations for studies with the Texas Red
electrophile. This same method was applied for the 501 insert and 512WPE inserts in EF-Tu (Table 1).

1.10.1.4 Determination of protein concentrations

Protein concentrations were initially determined using their UV absorbance at 280 nm. Extinction co-efficients were calculated online using Richard’s protein calculator.

1.10.2 Synthesis of 501 and 512 peptides

1.10.2.1 Molecule 2-1

Molecule 0-1 501 Peptide

Amide-GGGSKVILFEGPAG RTIWPKEASNHTGAPGSKVILFEGGPG-Acetyl

MS (MALDI-TOF) (m/z) Calculated for Molecule 2-1 monoisotopic mass [M+H⁺]: 4246.20, observed: 4246.61.

1.10.2.2 Molecule 2-2

Molecule 0-2 512 Peptide

Amide-GGGSKVILFEGPAG RWTWEPISE GAPGSKVILFEGGPG-Acetyl

MS (MALDI-TOF) (m/z) Calculated for Molecule 2-2 monoisotopic mass [M+H⁺]: 3881.0, observed: 3881.88.

1.10.2.3 501/512 Peptide Experimental

Two peptides, 501 and 512, were synthesized using a CEM Liberty Automated Microwave Peptide Synthesizer with 0.6 substitution on Fmoc-PAL-PEG-PS resin for both peptides using N-alpha-FMOC-protected amino acids. Double coupling was performed for valine, isoleucine and
arginine. HBTU was used as the coupling reagent and the peptides were removed from resin by mixing with a cleavage mixture of 9.25 mL TFA, 250 µL water, 250 µL TIS, and 250 µL 2,2’-(ethylenediozy)diethanethiol for three hours followed by filtration and 8X volume ether precipitation in triplicate. The peptide pellet was dissolved in a mixture of 1:1 acetonitrile/water, spin filtered and purified using reverse phase HPLC on a Shimadzu Prominence system with a Vydac (218TP C18 5u) column. The HPLC purification was run from 10-60% gradient of 0.1% TFA water solution and 0.1% TFA acetonitrile solution as the mobile phase monitoring at 215 nm. The gradient was run for 40 minutes to give elution times of approximately 30 minutes. The purified HPLC fractions were verified using MALDI-TOF MS using CHCA as the matrix detecting a mass range from 3000-7500 AMU in linear mode.

A peptide mass calculator v3.2\textsuperscript{122} was used to determine the final mass of the peptides with an acetyl N-terminal group and an amide C-terminal group and the extinction coefficient calculated at 280 nm\textsuperscript{123}, Figure 2-11. The MALDI-TOF MS spectra of the purified peptides are provided in Appendix A.

**Figure 0-11 501 and 512 peptide molecular weight calculations**

501
GGGSKVILFEGPAG RTIWEPEASNHTGAPGSKVILFEGPG
Monoisotopic mass: 4245.1969,
[M+H]+: 4246.204
42 AA
Extinction Coefficient: 5690 cm\(^{-1}\)M\(^{-1}\)

512
GGGSKVILFEGPAG RWTEWPEISE GAPGSKVILFEGPG
Monoisotopic mass: 3879.9946,
[M+H]+: 3881.002
38 AA
Extinction Coefficient: 11380 cm\(^{-1}\)M\(^{-1}\)
1.10.3 Fluorescent electrophiles chosen for covalent labeling studies

A variety of rhodamine B and sulforhodamine 101 analogs were tested against various proteins containing peptidic inserts or against the wild type proteins. The following is the list of analogs whose binding data is provided in the text of this chapter, Table 0-5.

Table 0-5 Fluorescent electrophiles tested for binding with proteins

<table>
<thead>
<tr>
<th>Molecule</th>
<th>Name</th>
<th>Peptide Insert</th>
<th>Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Rhodamine B</td>
<td>501</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>512</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>EF-Tu</td>
</tr>
<tr>
<td>1-11</td>
<td>RhB Piperazine Amide Benzophenone</td>
<td>501</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>512</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>EF-Tu</td>
</tr>
<tr>
<td>1-9</td>
<td>RhB Piperazine Amide Propanol</td>
<td>501</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>512</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>EF-Tu</td>
</tr>
<tr>
<td>1-12</td>
<td>Sulforhodamine 101</td>
<td>512</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>512WPE</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>512</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>501</td>
<td>-</td>
</tr>
<tr>
<td>1-19</td>
<td>Sulforhoamdine 101 Tosylate</td>
<td>512</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>512WPE</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>512</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>501</td>
<td>-</td>
</tr>
<tr>
<td>1-22</td>
<td>Sulforhodamine 101 Etam</td>
<td>501</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>512</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>EF-Tu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>512</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>501</td>
<td>-</td>
</tr>
</tbody>
</table>

1.10.4 Fluorescence polarization experiments

A working range of binding affinities was optimized for the fluorescence polarization studies. In these experiments the concentration of the fluorophore was kept constant against an increasing concentration of protein. At higher concentrations aggregation of the protein can interfere with analysis such that weak interactions between the fluorophore and protein are difficult to analyze due to nonspecific light scattering. In most cases the proteins were tested at
pH 7.4 in 1X PBS and incubated at 37 °C for one hour prior to testing. TRIS buffer and HEPES buffer were also tested to insure that binding was not affected by sodium concentration and/or an interaction with other salts in the media. The samples were analyzed in a 100 µL quart cuvette using a Cary Eclipse Spectrophotometer with a polarizing lens attachment. The G-factor was determined by measuring the fluorophore at the testing concentration in the buffer used to prepare the proteins. The polarization values or anisotropy values were calculated by the Advance reads software and used to determine the EC₅₀ with a four-parameter logistic equations using Sigma Plot.

### 1.10.5 General labeling assays protocol

Proteins were labeled with the fluorescent probes in biologically relevant media, typically 1X PBS at varying concentrations and times and incubated for at 37°C. The proteins were then boiled for five minutes in a 1:1 Laemmli buffer: sample mixture and run on a 10% SDS Page gel at 130 V for the stacking and 200 V for the resolving gel, Figure 0-12. A series of experiments were completed to establish that the quenching of the reaction mixture with BME used in the Laemmli buffer and/or the denaturation by heat was not the source of covalent labeling.

**Figure 0-12 In gel evaluation of reaction between fluorophore and protein**

![Image](image_url)

Left picture: SDS page gel during elution, the blue bands are bromophenol
blue and the pink bands below are the fluorescent electrophile used for labeling. Right picture: SDS page gel visualized at 365 nM UV light.

### 1.10.6 Beads experiment

Seven sets of 100 µL aliquots Ni-NTA Agarose beads (Qiagen) were prepared for this experiment. Sample A served as a control whereby the beads were washed only with the same media used to prepare the fluorophore and the proteins. Samples B and C served as the background control for the Molecule 1-12 and the beads were exposed to the dye solution but no protein. Samples D and E captured EF-Tu protein in 1XPBS from a 5.5 µM solution followed by reaction with Molecule 1-12 at 1.6 µM and 1.6 nM respectively. Samples F and G captured 512 EF-Tu from a 10 µM solution in 1XPBS followed by reaction with Molecule 1-12 in water at 1.6 µM and 1.6 nM respectively. In all cases the beads were washed with double distilled deionized water (ddH₂O) and visualized via fluorescent microscopy to test parameters and background for future covalent binding using a Texas red filter.
Utilizing mRNA display as a tool to find fluorescent binding peptides

This chapter provides a general introduction to utilizing in vitro evolutionary selection towards the identification of a short peptidic sequence with specificity towards a biotinylated small molecule with covalent labeling properties. Achievement of site selective covalent binding of a small labeling molecule (SLM) to a protein would provide a powerful tool that could be used to study protein-protein interactions and elucidate biological pathways. To accomplish this task it was necessary for the small molecule to be selectively reactive with a specific peptidic nucleophile found within the target protein. This reactivity was controlled by the electrophilic nature of a small molecule containing an excellent leaving group that was found to be stable in biological media and shown to react covalently with the biomolecule. Maximizing the reactivity of the small molecule and simultaneously optimizing selectivity of the reaction between the small molecule and a specific peptidic nucleophile, within a protein, would permit site-selective labeling of peptide-tagged proteins within live cells. Elucidation of the sequence of a peptide with this feature was likely to be rare. Utilization of mRNA display to create a peptide library could have provided access to a systematic combination of a large number of peptides and has been found to be a powerful tool used for drug discovery, structural studies, and identification of bioactive peptides and ligand binding activities\textsuperscript{11,124,125}. 

119
1.11 Introduction in vitro selection

For a chemist, utilizing the size of a large set of molecules provided by the in vitro selection of a few key molecules towards a desired target is a powerful tool. It is possible to mimic natural selection and find key molecules without bias that have a relationship with a desired target. Directed evolution of the living world has long been recognized and appreciated by early scientists such as William Paley\textsuperscript{126} and Charles Darwin\textsuperscript{127,128}. With regards to the directed course of evolution, Richard Dawkins said that, ""All appearances to the contrary, the only watchmaker in nature is are the blind forces of physics, albeit deployed in a very special way. A true watchmaker has foresight: he designs his cogs and springs and plans their interconnections, with a future purpose in his mind's eye. Natural selection, the blind, unconscious, automatic process which Darwin discovered, and which we now know is the explanation for the existence and apparently purposeful form of all life, has no purpose in mind. It has no vision, no foresight, no sight at all. It can be said to play the role of watchmaker in nature, it is the blind watchmaker.\textsuperscript{129}""

There have are various methods used to create peptide libraries that include traditional solid phase synthesis, yeast two hybrid system\textsuperscript{130}, phage display\textsuperscript{131,124}, cell surface display using bacteria and yeast\textsuperscript{132,133}, ribosome display\textsuperscript{134} and in vitro compartmentalization\textsuperscript{135}. Ribosome display and mRNA have a distinct advantage over other techniques due to enhanced library size and diversity,
Table 0-1. A larger library size means there is a higher possibility of isolating peptides of interest with higher affinity and specificity\textsuperscript{136}. 

Table 0-1 Comparison of mRNA display versus other peptide libraries.

<table>
<thead>
<tr>
<th>Libraries and Selection Techniques</th>
<th>Library Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yeast Two Hybrid System</td>
<td>$10^6$</td>
</tr>
<tr>
<td>Traditional solid phase synthesis</td>
<td>$10^7$</td>
</tr>
<tr>
<td>Bacteria and Yeast Cell Surface</td>
<td>$10^6-10^{10}$</td>
</tr>
<tr>
<td>Phage Display</td>
<td>$10^8-10^{10}$</td>
</tr>
<tr>
<td>In Vitro Compartmentalization</td>
<td>$10^9 - 10^{10}$</td>
</tr>
<tr>
<td>Ribosome Display</td>
<td>$10^{12} - 10^{13}$</td>
</tr>
<tr>
<td>mRNA Display</td>
<td>$10^{12} - 10^{14}$</td>
</tr>
</tbody>
</table>

1.11.1 Phage display

The most commonly Phage display refers to a pool or library\textsuperscript{137} that utilizes the ability of a phage (a viruses that infects bacterial cells) to insert DNA into a gene responsible for the surface protein of bacteriophage. The DNA is expressed as a polypeptide sequence/protein that appears on the surface coating of the phage and creates a system whereby the peptides are physically attached to its encoding DNA sequence such that isolated clones can be isolated and identified. Various drug targets have been identified using phage display, including D2E7\textsuperscript{138} known as Humira\textsuperscript{™} by Abbott Laboratories. Some limitations of phage display include speed and library diversity (\textasciitilde$<10^9$) compared to mRNA/Ribosome display due to dependence on the efficiency of transformation of the initial library into the host cell\textsuperscript{139}.

1.11.2 Ribosome display

Ribosome display is a cell free type of vitro selection of proteins and peptides whereby the DNA template encoding the randomized peptide is directly added\textsuperscript{140}. Without the limitation of the efficiency of transformation of the phage into the host cell the diversity of the initial library for ribosome display can provide libraries of $>10^{12}$ peptide sequences in size\textsuperscript{134,139,141} and diversity is dependent on translation efficiency and scale of reaction. In ribosome display the ribosome,
mRNA and polypeptide are stabilized into a complex such that the polypeptide protrudes from the complex and can fold and is available for reactivity with a desired partner. To fold away from the complex. The link between the genotype and phenotype is provided by a spacer sequence that lacks a stop codon\textsuperscript{139}.

1.11.3 mRNA display

mRNA display is an in vitro selection technique that permits the identification of new molecules from libraries that can contain up to $10^{13}$ sequences\textsuperscript{12}. These identified molecules discovered by mRNA display have helped to elucidate biological processes\textsuperscript{142}, identify protein-protein interactions\textsuperscript{143}, and find potential antibody\textsuperscript{141,144} and peptide\textsuperscript{11} mimics some of which can bind to potential drug candidates\textsuperscript{145}.

Compared to ribosomal display, mRNA display offers the unique advantage that the linkage between the mRNA and the peptide is a stable, covalent amide bond in lieu of non-covalent binding. The covalent nature of this linkage allows increased flexibility in terms of selection conditions to minimize selection of nonspecific sequences. In addition, the polypeptide in ribosome display is attached to the ribosome versus the puromycin DNA linker in mRNA display; the enormous size difference between the two methods permits less biased results for mRNA display\textsuperscript{139}.

1.12 Strategy using mRNA display

The overall strategy to obtain peptides with a high reactivity/affinity for a specific electrophile will be accomplished by testing a highly diverse peptide library against an electrophilic small labeling molecule (SLM). mRNA display will be used to generate a library of peptides containing
$10^{12}$-$10^{13}$ diverse peptides that are potentially nucleophilic$^{124}$. Each of these peptides is covalently bound to its encoding mRNA, which permits PCR amplification of the corresponding cDNA. The DNA can be sequenced and used to identify those members of the library that are selectively reactive with the electrophile of choice.

1.12.1 Overview for the identification of short peptidic sequence with affinity for rhodamine

In vitro selection utilizing mRNA display was used to create a library of peptide mRNA fusions to test against an electrophilic small labeling molecule (SLM). Scheme 0-1 illustrates the processes during each round of selection to discover reactive peptides of interest. To begin, a diverse library of oligonucleotides is made on a DNA synthesizer and then undergoes transcription to its corresponding mRNA (Step 1). The mRNA is then cross-linked with a puromycin/psoralen oligonucleotide (Step 2). Once the mRNA is cross-linked to puromycin, in vitro translation using a PURE system$^{146}$ translates the mRNA into its corresponding peptide. Puromycin aids in stalling the ribosomal machinery such that competitive amidation takes place and the peptide translated is covalently attached to its encoding mRNA. (Step 3) Reverse transcription of the mRNA strand creates cDNA that serves to stabilize the double stranded DNA (Step 4). At this point, reaction of the library with a target of interest takes place in a step known as selection (Step 5). During selection, the peptide library is incubated with the electrophile, a biotinylated tosyl-sulfonate ester, in aqueous media. Certain peptides become biotinylated. Only those peptides that are reactive with the desired electrophiles yield covalently-bound, biotinylated peptides that are then selectively isolated using magnetic-streptavidin beads (Step 6). The peptide fusion containing the encoding cDNA can be PCR amplified to give multiple copies of the cDNA that encodes for those members of the peptide
library that are reactive (Step 7). The DNA product after each round is then used for the next round and the process is repeated until enrichment of the peptide library is observed. After several rounds of selection an increasingly concentrated pool of reactive peptides is expected to remain and the members of this pool can be sequenced and identified (Step 8). The DNA that encodes for the reactive peptides will be analyzed yielding possible trends associated with each electrophile.

**Scheme 0-1 Generalized schematic of mRNA display (\(^{147}\))**

---

**1.12.2 mRNA display –library design**

The random region of the template encodes for 12 amino acids via a degenerate codon NNB, whereby N=A,T,C or G and B=T, C, or G. The rationalization for this semi-random codon is to decrease the chance of stop codons in addition to increasing the percentage of cysteines that
can act as nucleophiles towards the electrophilic SLM. The components of the library are described in Table 3-2.
<table>
<thead>
<tr>
<th>Section of Construct</th>
<th>Sequence</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>T7 Promoter</td>
<td>TAATACGACTCATATAG</td>
<td>T7 polymerase transcribes DNA downstream of the promoter. Needed for transcription from DNA to mRNA.</td>
</tr>
<tr>
<td>Translation Enhancer Sequence</td>
<td>TTAACCTTAG</td>
<td>A truncated sequence from the Tobacco Mosaic Virus used to enhance translation before initiating methionine.</td>
</tr>
<tr>
<td>Shine Dalgarno</td>
<td>TAAGGAGG</td>
<td>Prokaryotic binding site used to recruit binding of mRNA to the ribosome.</td>
</tr>
<tr>
<td>Spacer</td>
<td>ACAGCTAA</td>
<td></td>
</tr>
<tr>
<td>Initiating Methionine</td>
<td>ATG</td>
<td>A start codon: the initiating methionine (ATG).</td>
</tr>
<tr>
<td>Cys</td>
<td>TGC</td>
<td>Statistically increases the chance of 2 Cysteine residues available for cyclization of peptide.</td>
</tr>
<tr>
<td>Library CX12</td>
<td>(NNB)X12</td>
<td>NNB = semi-random codon Where N = (A, G, C, or T) and B = (G, C, or T) of 12 amino acids encoded by a degenerate codon NNB (B = T, C, or G). This semi-random codon decreases the prevalence of stop codons and increases the percentage of cysteine in the peptide library for later cyclization,</td>
</tr>
<tr>
<td>Flexible Linker</td>
<td>GGCTCCGTTAGCTTAGGC</td>
<td>Peptide attached at C terminus of mRNA, therefore, it is useful to provide structureless amino acids for flexibility. GlySerGlySerLeuGly Flexible Linker with 2 out of frame stops</td>
</tr>
<tr>
<td>His Tag</td>
<td>CACCATCACCATCAC</td>
<td>6 histidines used for Nickel NTA purification following PCR amplification of cDNA.</td>
</tr>
<tr>
<td>Linker Hybridization</td>
<td>CACCGGCTAT</td>
<td>Linker hybridization with first Stop</td>
</tr>
<tr>
<td>3’ UTR</td>
<td>AGGTAGCTAG</td>
<td>3’ Untranslated Region to allow non cross-linked peptides to release at the in frame TAG</td>
</tr>
</tbody>
</table>

During the first round, the template described above was synthesized and purified. The DNA then underwent transcription to mRNA, purified and then cross-linked using UV light, and puromycin functionalized with Psoralen XL Oligo = 5’-PsoC6-(tagccggtg)2’-Ome-15xA-2xspacer9-ACC-Puro-3’. Psoralen is used to intercalate DNA and forms covalent inter-strand cross links with thymine; the O-methylated oligos are advantageous for RNase H resistance; and the polyadenylated tail of the molecule is used for affinity purification using Oligo-dT cellulose.
beads. Puromycin is critical for mRNA display as it causes premature termination during translation and resembles the 3’ end of amino-acylated tRNA, which ultimately results in covalent bonding between the mRNA template and the translated peptide, see Figure 0-1.

**Figure 0-1 mRNA display covalent linkage**

I carried out the translation in the presence of $^{35}$S-radiolabeled methionine (Met) to yield a radioactive peptide fusion. Radioactive labeling of Met-containing peptides provided the means to calculate yields of successive purifications and more importantly, the yield of the selection process used to determine progress of enrichment. Once the fusion was purified via affinity binding of Oligo-dT cellulose beads, the encoding mRNA underwent reverse transcription. RT-PCR covers the mRNA to enhance stability and the peptide fusion then underwent the selection process.

**1.12.3 Library selection against a small molecule**

The peptide library underwent a selection process with an electrophile, Molecule 3-3 whereby nucleophilic members of the peptide that are reactive with the electrophile form a covalent bond and can be captured and isolated on magnetic iron beads (Scheme 0-2). The cDNA tethered to those reactive peptides can be PCR amplified and the process of mRNA display was repeated.
1.12.4 Design and choice of electrophile

Selective reactivity between peptidic nucleophiles within proteins and various small molecule electrophiles has been demonstrated both in vivo and intracellularly. Hamachi and coworkers have shown the effectiveness of using p-methyl-sulfonate ester (tosyl) as a leaving group in biological media in experiments that utilized proximity effect to motivate covalent
binding between a ligand and probe that had known affinity for one another. However, the ligand probe utilized in these studies cannot be generally applied to different proteins. Cravatt and coworkers have recognized patterns of reactivity between certain carbon electrophiles and various enzymes generating a screen to isolate families of enzymes, concluding that there exists disparate reactivity between the five types of electrophiles tested and functional residues in proteomes, possibly dictated by the local protein micro-environment. Our approach was to find a specific reaction between a nucleophilic peptidic sequences that covalently and site-specifically reacted with an electrophilic SLM. This would have allowed us to more generally apply this labeling technique toward a broader range of proteins that had not yet been achieved.

Tosylates were chosen as the electrophile due in part to the success of working with these types of electrophiles in biological environments by Cravatt and co-workers and Hamachi and co-workers. Tosylate esters have been used as model electrophiles because they are stable in biological media, the corresponding leaving group is not cytotoxic, and sulfonate esters can be tuned in terms of reactivity. It has been observed that tosylate esters are moderately reactive as opposed to spiro-epoxides, which were found to be highly reactive and less selective than epoxide containing molecules, which are not reactive enough to be useful.

The electrophile designed for testing with the peptide library contains three components, biotin, a polyethylene oxide linker and a highly reactive sulfonate ester. In the presence of a nucleophilic amino acid, an $\text{SN}_2$-type reaction produces a covalent bond between the reactive amino acid that is then tethered to a flexible linker containing biotin. The flexible
linker is important as it provides spatial distance between the peptide fusion and biotin and this distance serves two purposes: (1) it minimizes steric hindrance between the fusion and the magnetic streptavidin beads, which will serve to magnify capture of the biotinylated peptides, and (2) it minimizes interference of the beads with the fusion when undergoing reverse transcription to PCR amplify the cDNA. Finally, biotin is chosen due to its high affinity for streptavidin. Magnetic beads containing streptavidin are highly selective and the solid phase interface of the beads provides a physical means of easily capturing the biotinylated peptide fusions.

1.13 Results and conclusions

1.13.1 Preparation of Molecule 3-3

Two electrophiles were synthesized (Scheme 0-3) whereby a phenyl-sulfonate ester (Molecule 0-4) was made in addition to the tosyl-sulfonate ester (Molecule 0-3).
To begin, biotin was activated as the NHS ester using standard coupling conditions to give Molecule 3-1 that was subsequently reacted with polyethylene oxide monoamine to give the desired alcohol (Molecule 3-3). The alcohol of the biotinylated linker was then reacted with \textit{para}-toluenesulfonyl chloride or benzene sulfonyl chloride under basic conditions to give the desired electrophile, Molecule 3-3 or Molecule 3-4, respectively.
1.13.2 8 Rounds of mRNA peptide selection with Molecule 3-3

Following reverse transcription, the fusion product was reacted with streptavidin magnetic particles. Reaction of the fusion product with the beads before exposure to the electrophile removed those peptides from the fusion that reacted non-selectively with the magnetic beads. The amount of beads used was determined by the binding capacity of the magnetic beads; 10 µl of streptavidin beads is sufficient to isolate 180 pmol of fusion product per assay. The pre-cleared beads were counted to determine what percentage of fusion was removed during the pre-clearance reaction. The remaining fusion was then reacted with Molecule 3-3 and dialyzed to remove any unreacted electrophile. Scintillation counting of the radioactive methionine ($^{35}$S) incorporated into the peptide sequence was used to determine yields at each step of purification and following the selection process. Additional details with regards to the selection process are described in Appendix B: mRNA display supplemental information.

Eight rounds of selection took place and the percent yield following each round of selection is described in Chart 0-1. Initially, the capture of peptides was low and statistically approximate to background noise. Rounds 4 through 6 gave a low yield of capture while Round 7 indicated enrichment. However, Round 8 did not show enhanced enrichment and led us to believe that the enhanced enrichment for Round 7 may have been an artifact.
In addition Round 8 was repeated nine times without success. 8a and 8h could not be PCR amplified and the results of 8i were impure (Figure 3-2). For round 8i, small scale amplification of cDNA from the captured mRNA peptide fusions was used to increase the yields of the DNA in two ways. First, the number of cycles needed for PCR amplification was optimized. The gel shown in Figure 3-2(a) shows the amplification of DNA at 22, 24, 26, and 28 cycles for both 4 and 5μL. All samples showed a small second band at a lower molecular weight and 26 cycles were chosen. Secondly, a variable volume of beads (1-5 μL) was to each PCR reaction. The gel shown in Figure 3-2(a) represents the reactions of a sample size of 4 and 5 μL of beads, at 22, 24, 26, and 28 cycles respectively. Thirdly, it was discovered that agitation of the beads during the PCR reaction increased the yields dramatically. The sample represented by 4* in Figure 3-2(a) shows that the yield is reduced after 28 cycles if the sample is not agitated. Finally, the optimized reaction conditions were applied to the remainder of the beads and seven random samples were pulled from the multiple PCR reactions to show the final amplified product, Figure 3-2 (b). The results of the amplification show an impure product.
Figure 0-2 Round 8i PCR amplification

Figure 3-2(a)

Figure 3-2(b)

Figure 3-2(a) - Small scale testing of Round 8i cDNA amplification of SA beads post selection. 4 or 5 uL of beads tested at 22, 24, 26, 28 cycles of PCR amplification. 4* represents no agitation of the sample during reaction. Figure 3-2 (b) Lanes 1-7 represent 7 random samples taken from large scale PCR reaction following optimized PCR conditions. In both gels, (+) indicates dsDNA from the peptide fusion library before selection was used as the positive control for and (-) is water used for the PCR reactions as the negative control.

The results of each round for each step are provided in Table 3-3.
**Table 0-3 Summary of Molecule 3-3 selection results using mRNA Display**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Round 1</th>
<th>Round 2</th>
<th>Round 3</th>
<th>Round 4</th>
<th>Round 5</th>
<th>Round 6</th>
<th>Round 7</th>
<th>Round 8</th>
<th>Round 9</th>
<th>Round 10</th>
<th>Round 11</th>
<th>Round 12</th>
<th>Round 13</th>
<th>Round 14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Translation Scale (mL)</td>
<td>0.25</td>
<td>0.5</td>
<td>0.5</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>5x1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oligo Yield (pmol)</td>
<td>140.8</td>
<td>80.0</td>
<td>214.2</td>
<td>33.3</td>
<td>5.0</td>
<td>2.1</td>
<td>3.1</td>
<td>3.1</td>
<td>10.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OligoFiltered Final yield (pmol)</td>
<td>148.2</td>
<td>76.7</td>
<td>198.5</td>
<td>38.8</td>
<td>4.8</td>
<td>2.5</td>
<td>-</td>
<td>2.7</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cpm of Product</td>
<td>1.02E+07</td>
<td>1.69E+06</td>
<td>5.59E+06</td>
<td>1.44E+06</td>
<td>9.84E+04</td>
<td>1.89E+05</td>
<td>1.89E+05</td>
<td>3.10E+05</td>
<td>9.13E+05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cpm/5 uL master mix</td>
<td>4.92E+06</td>
<td>1.69E+07</td>
<td>5.57E+06</td>
<td>3.64E+06</td>
<td>5.80E+05</td>
<td>3.10E+06</td>
<td>1.00E+06</td>
<td>4.47E+06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cpm/pmol</td>
<td>1.27E+05</td>
<td>4.38E+04</td>
<td>1.44E+05</td>
<td>2.90E+05</td>
<td>4.61E+04</td>
<td>6.16E+04</td>
<td>9.94E+04</td>
<td>8.92E+04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># RT Reactions</td>
<td>6.0</td>
<td>6.0</td>
<td>6.0</td>
<td>8.0</td>
<td>1.1</td>
<td>0.5</td>
<td>0.5</td>
<td>0.7</td>
<td>0.7</td>
<td>2.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(mRNA)/ reaction pmol/μL</td>
<td>0.13</td>
<td>0.23</td>
<td>0.05</td>
<td>0.04</td>
<td>0.05</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nickel (pmol)</td>
<td>-</td>
<td>32.5</td>
<td>78.7</td>
<td>4.5</td>
<td>1.0</td>
<td>0.1</td>
<td>0.3</td>
<td>0.1</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Yield (Nickel/Oligo)</td>
<td>-</td>
<td>42.4</td>
<td>39.6</td>
<td>11.6</td>
<td>20.8</td>
<td>5.1</td>
<td>11.0</td>
<td>2.4</td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cpm</td>
<td>-</td>
<td>4.13E+06</td>
<td>3.45E+06</td>
<td>6.49E+04</td>
<td>2.91E+05</td>
<td>5.80E+03</td>
<td>2.08E+04</td>
<td>5.90E+03</td>
<td>1.79E+04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beads (pmol)</td>
<td>-</td>
<td>1.34</td>
<td>0.41</td>
<td>0.06</td>
<td>0.01</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cpm</td>
<td>-</td>
<td>3.42E+03</td>
<td>1.81E+04</td>
<td>8.93E+03</td>
<td>3.96E+03</td>
<td>BG</td>
<td>BG</td>
<td>BG</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fusion (pmol)</td>
<td>-</td>
<td>27.0</td>
<td>23.76</td>
<td>2.04</td>
<td>1.37</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biotin Fe Beads Used (μL)</td>
<td>-</td>
<td>50.0</td>
<td>70.0</td>
<td>50.0</td>
<td>10.0</td>
<td>10.0</td>
<td>10.0</td>
<td>10.0</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[fusion]/beads in pmol/μL</td>
<td>-</td>
<td>0.65</td>
<td>1.12</td>
<td>0.09</td>
<td>0.10</td>
<td>0.01</td>
<td>0.03</td>
<td>0.0</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amount of Electrophile (umol)</td>
<td>-</td>
<td>30.0</td>
<td>30.0</td>
<td>50.0</td>
<td>10.0</td>
<td>10.0</td>
<td>10.0</td>
<td>10.0</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Yield (Beads/Oligo)</td>
<td>-</td>
<td>1.7</td>
<td>0.2</td>
<td>0.2</td>
<td>0.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Yield (Beads/Nickel)</td>
<td>-</td>
<td>4.1</td>
<td>0.5</td>
<td>1.4</td>
<td>1.4</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCR- Beads Used</td>
<td>-</td>
<td>5.0</td>
<td>5.0</td>
<td>10.0</td>
<td>5.0</td>
<td>10.0</td>
<td>10.0</td>
<td>10.0</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># of Cycles</td>
<td>-</td>
<td>28.0</td>
<td>26.0</td>
<td>22.0</td>
<td>22.0</td>
<td>28.0</td>
<td>28.0</td>
<td>24.0</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dsDNA (μM) in 500 μL</td>
<td>-</td>
<td>3.8*</td>
<td>3.1*</td>
<td>2.7</td>
<td>1.5</td>
<td>1.0</td>
<td>0.5</td>
<td>4.2</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mRNA (μM)</td>
<td>-</td>
<td>78.6</td>
<td>12.4</td>
<td>20.7</td>
<td>109.9</td>
<td>24.5</td>
<td>18.8</td>
<td>7.5</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Variables that were addressed for reaction of the peptide fusion with Molecule 3-3 included the following: establishing preclearance parameters, determining ratio of beads relative to biotinylated peptide library, determining the ratio of beads relative to any residual unreacted Molecule 3-3 not removed during dialysis, stability and storage of beads in selection buffer, concentration, time and temperature of reaction, and solubility issues associated with Molecule 3-3 in water.

For successful completion of PCR amplification post selection it was necessary to optimize a number of parameters. For most of the following experiments the purity of the cDNA amplified was determined using a 2% Agarose gel and visualization of the bands using...
ethidium bromide. First, it was important to validate the positive and negative controls used for the PCR amplification for each round. In each round, a small sample of the mRNA peptide fusion before pre-clearance was used as the positive control, Figure 3-3 (a). The distilled, deionized water used for the PCR reaction served as the negative control Figure 3-3(b). In addition, the beads that were used for preclearance were PCR amplified. Those beads that reacted with the fusion during the pre-clearance study theoretically removed non-selective binding to the electrophile.

**Figure 0-3 Controls used for mRNA peptide fusion amplification and selection**

![Image of gel electrophoresis](image)

2% Agarose gel testing cDNA amplification at PCR cycles 22-32 of (a) negative control (b) positive control and (c) SA beads used for preclearance of mRNA peptide fusion before selection with Molecule 3-3.

In the experiment described by Figure 3-4 below, two concentrations of Molecule 3-3 used during selection and the mRNA fusion captured on SA beads were tested. In addition, the volume of beads needed per PCR reaction was tested, 10 versus 20 µL of beads per 100 µL PCR. Finally, the number of times the beads needed to be washed to remove non-biotinylated mRNA fusion was determined. As can be visualized by the results of the cDNA separated via a 2% agarose gel in 1XTBE, the concentration of electrophile during the selection process was tested.
at 0.6 and 60 μM. The results were similar between high and low concentrations of the electrophile with the amplification of the lower concentration being slightly better. In addition, it was clear that 10 μL of beads amplified better than 20 μL. Finally, it was decided that two washes of the beads prior to PCR amplification was better than a single wash.

**Figure 0-4 PCR Optimization of concentration of electrophile exposed to mRNA fusion on magnetic particles**

There existed a lower and upper limit to the number of beads used in each PCR amplification reaction that required optimization at each round. Excess beads were found to impede amplification. Bead optimization for Round 4 is exemplified in Figure 3-5. In these experiments, a volume of 5 μL (Figure 3-5(a)), 10 μL (Figure 3-5(b)), and 20 μL (Figure 3-5(c)), of SA beads were tested from mixture of SA beads exposed to the mRNA peptide fusion after selection with Molecule 3-3 and subsequently washed. The numbers 22-34 represent the number of amplification cycles tested in the thermocycler. From these experiments it is evident that for Round 4, a volume of 20 μL of beads impedes amplification and 10 μl produced
a purer product than 5 μL. The conclusion yielded from these results was that 10 μL of beads were used for the PCR reactions with 22 cycles.

**Figure 0-5 Optimization of volume of beads needed for PCR amplification**

![Image of gel electrophoresis with three lanes labeled (a), (b), and (c). Each lane contains bands labeled with cycle numbers from 22 to 34, and molecular weight markers at 200 bp and 100 bp. The legend explains that 2% agarose gel testing cDNA amplification of mRNA peptide fusion captured during selection using Biotin-(PEO)3-OTs, samples were taken at cycles 22/24/26/28/30/31 and 34 respectively. Volume of SA beads added per 100 uL PCR (a) 5 uL (b) 10 uL (c) 20 uL. The mRNA peptide fusion pre-selection was used as the positive control (+) and water used for the PCR reaction used as the (-) control.]

1.13.3 Stability of electrophile in the selection reaction in the presence of imidazole

A potential problem during the reaction between the mRNA peptide fusion and Molecule 3-3 was the stability of the electrophile in the presence of imidazole. The C-terminus of the peptide fusion contains a His tag that enables Nickel-NTA affinity purification. Imidazole is used to elute the tagged peptides through competitive binding to nickel ions attached to the surface of the beads in the column releasing the His tag from nickel co-ordination and the buffer that contained the mRNA fusion following purification contained imidazole. Imidazole can act as a nucleophile and it was necessary to validate that the electrophile was available for reactivity with the mRNA fusion at the concentrations we had optimized. Therefore, a TLC experiment was run to establish whether or not the presence of imidazole impacted the stability of the electrophile testing comparable concentrations of imidazole and Molecule 3-3 in
the selection buffer used during at various time points, Figure 3-6. It was shown that imidazole in the reaction mixture does not significantly react with the electrophile indicating that the concentration isn’t a changing parameter during reaction with the fusion and these results were confirmed by MALDI-TOF MS. None-the-less, for future experiments, the selection process should take place in more relevant biological media.

**Figure 0-6 Imidazole testing on electrophile stability.**

Reaction of Molecule 3-3 in elution buffer containing imidazole at various time points (20% MeOH/ 90% CHCl3). B= Molecule 3-3 reference standard , I= Imidazole reference standard, EB=Elution Buffer containing imidazole, and the following reaction mixture times points were taken at 5, 20, 40 minutes, 1 hour, and 12 hour. 8 uL of each time point was spotted.
1.14 Discussion

One significant drawback to mRNA display is that the process of obtaining final results is lengthy. Each experimental round is time intensive and takes approximately one week to complete. Despite the lack of enrichment it would have been useful to sequence Rounds 7 and 8 to determine if there were any comparable trends. This would have provided more information about whether the success of Round 7 was due to an increase in enrichment or an artifact. In addition, had there been any repeated peptides in the pool these could have been synthesized and tested against the electrophile to determine the rate of reactivity and selectivity and helped to establish whether or not using mRNA display to find short peptide sequences with high and specific reactivity towards an electrophile is a useful tool and worth repeating with other types of electrophiles.

If this protocol were to be continued I would recommend the following changes. During the selection process, a mixed pot of various biotinylated electrophiles should be mixed together with known moderate or low reactivity. This may result in lower stringency initially, but means we are likely to identify peptides that are more likely to be reactive in general. Once enrichment is established, then multiple copies of cDNA will be present in the library and the library can be divided by the number of electrophiles tested or divided by the number of families of electrophiles tested. For example sulfonate esters, haloacetamides, and unsaturated ketones are known electrophiles used in biological media with known moderate and high reactivity. After additional rounds to further enrich the peptides and their corresponding electrophile or family of electrophiles, the cDNA can be submitted for sequencing and it will be
possible to correlate peptides and their respective electrophiles. Utilizing this mixed approach would likely save months of work, time and resources.

1.14.1 Future direction

1.14.1.1 Selecting for biotin versus the electrophile

A final potential problem is that despite our pre-clearing experiments, we could be selecting for peptides with a natural affinity for biotin, the other end of the electrophilic small molecule. It will be necessary to test other probes when optimizing reaction conditions between the nucleophilic peptide and its corresponding electrophile to see if a change in probe changes the reactivity.

1.14.1.2 Future Directions if sequencing data is obtained

Sequencing data will help us to identify and establish trends within the sequences obtained. The next step will be to synthesize promising peptides via standard solid state coupling conditions, purify them via reverse phase HPLC then react them with Molecule 3-3. The masses of the resulting products can be validated via MALDI-TOF. Identification of the site of reactivity within the peptide can be obtained via MALDI TOF MS/MS.

Once the site of reactivity is determined, then future elucidation of the necessary amino acids needed for reactivity with the electrophile can be accomplished by rationally changing the amino acid sequence closest to the amino acid that covalently bound with the electrophile. Ideally, the size and bulk of the peptide should be minimized, so it will be important to identify which amino acids are necessary to optimize reactivity. In addition, testing to see if cyclization is an important parameter for reactivity should also be performed.
Once a promising peptide sequence has been identified then the parameters of labeling need to be optimized and tested in various reaction conditions that should include establishing rate, yield and selectivity of reaction in water/ PBS/ glutathione (thiol rich environment) / competitive reaction with other electrophiles/ competitive reaction with other peptides/ cellular lysis. Once the goal of having identified and tested a peptide/electrophile pair is accomplished, then it will be possible to move forward to validate the method in various proteins.

The tosylate ester has been used as the model electrophile thus far. However, the phenyl analogue should might prove to be a better leaving group and as such, is a desirable alternative to the tosylate leaving group. Fine-tuning the reactivity of the electrophile can be achieved by further derivatization of the phenyl sulfonate ester.

1.15 Experimental strategies

1.15.1 Synthesis of electrophilic biotin analogues

General Experimental Methods: Anhydrous solvents were dried by refluxing with calcium hydride followed by distillation and stored over 4 angstrom molecular sieves.
1.15.1.1 **Molecule 3-1**

**Molecule 0-1**

![Chemical Structure]

**Chemical Formula:** $C_{14}H_{19}N_3O_5S$

**Exact Mass:** 341.10

**Molecular Weight:** 341.38

To a 10 mL flame dried vial, biotin (400mg, 2.05 mmol) was suspended in 2.7 mL of anhydrous dimethylformamide (3.38 mL, 40.31 mmol) that had been gassed/degassed three times with argon. To this vial was added N-hydroxy succinimide (282.6 mg, 2.46 mmol) and EDC (381.3 mg, 2.46 mmol) gassed/degassed with argon again and stirred at room temperature for one day under an argon balloon. Biotin is initially insoluble in DMF and upon reaction goes into solution, an indication that the reaction is complete. The reaction mixture was monitored via TLC using 20% methanol and 80% chloroform, iodine and vanillin being the preferred stains to visualize the product with an Rf=0.72. The retention factor of biotin in these conditions is baseline. The solvent was removed under vacuum and purification took place via flash column chromatography using 10% v/v methanol in chloroform.

Molecule 3-1: 2,5-dioxopyrrolidin-1-yl 5-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoate. $\equiv 0.5$ (methanol/chloroform= 1/9). White solid. Rf=0.72 using 20% methanol in chloroform. Yield: 0.4907 g, 70.2%. $^1H$ NMR (DMSO-$d_6$, 400 MHz): $\delta = 6.43$ (s, 1H), 6.37 (s, 1H),
4.30 (t, 1H), 4.15 (t, 1H), 3.13-3.06 (m, 1H), 2.85-2.81 (m, 5H), 2.67 (t, J=7.2 Hz, 2H), 2.58 (d, J=11.6 Hz, 1H), 1.682-1.410 (m, 6H). MS (MALDI-TOF) (m/z) Calculated for C_{14}H_{19}N_{3}O_{5}S [M+H]^+: 342.11, observed: 342.39.

1.15.1.2 Molecule 0-2

Chemical Formula: C_{14}H_{19}N_{3}O_{5}S
Exact Mass: 375.18
Molecular Weight: 375.48

Chemoselectivity for the formation of the amide and not the ester was accomplished in high yield and loosely based on a protocol provided by Rotello^{152} and coworkers. To a flame-dried 4 mL glass vial was added a stir bar, Molecule 0-1 (174 mg, 0.51 mmol), anhydrous distilled DMF (4.0 mL), anhydrous distilled dichloromethane (4.0 mL) and diisopropylethylamine (0.2 mL, 1.02 mmol), and the reaction mixture was gassed/degassed three times under Argon. Molecule 3-1 was poorly soluble in the reaction mixture but slightly more soluble upon addition of the base. The contents of the vial were cooled to 4°C and the (PEO)3-monoamine added (76 mg, 0.51 mmol), and allowed to react for two days in the cold room. The reaction was monitored via TLC using 20% methanol in chloroform v/v with an Rf=0.52 of the product. The reaction mixture
was taken up in chloroform and the solvent removed under vacuum and purification took place via flash chromatography. The column was run in a binary gradient with 15% methanol in chloroform (v/v) followed by 20% methanol in chloroform (v/v). The TLC plates were stained with Iodine, KMnO₄, Vanillin and Ninhydrin, KMnO₄ and Vanillin being the preferred stains.

Molecule 3-2: N-(2-(2-(2-hydroxyethoxy)ethoxy)ethyl)-5-((3aR,4R,6aS)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamide. = 0.43 (methanol/chloroform= 1/5). White oily solid. Rf= 0.52 using 20% methanol in chloroform. Yield: 154.2 mg, 80.5%. ¹H NMR (400 MHz, DMSO-d₆): δ = 10.53 (s, 1H), 7.81 (t, J = 5.6 Hz, 1H), 6.41 (s, 1H), 6.35 (s, 1H), 4.58 (t, J = 5.5 Hz, 1H), 4.36 – 4.25 (m, 1H), 4.18 – 4.08 (m, 1H), 3.56 – 3.45 (m, 6H), 3.44 – 3.37 (m, 4H), 3.32 (s, 3H), 3.23 – 3.15 (m, 2H), 3.14 – 3.05 (m, 1H), 2.82 (dd, J = 12.4, 5.1 Hz, 1H), 2.60 (s, 2H), 2.56 (s, 1H), 2.07 (t, J = 7.4 Hz, 2H), 1.68 – 1.20 (m, 7H). MS (MALDI-TOF) (m/z) Calculated for C₁₆H₂₉N₃O₅S [M+H⁺]: 376.19, observed: 376.4.

1.15.1.3 Molecule 3-3

Molecule 0-3

Chemical Formula: C₂₃H₃₅N₃O₇S₂
Exact Mass: 529.19
Molecular Weight: 529.67
To a stirred solution of Molecule 0-2 (25.0 mg, 0.07 mmol) in anhydrous degassed dichloromethane was added *para*-toluenesulfonyl chloride (18.4 mg, 0.1 mmol), diisopropylethylamine (35 µL, 0.20 mmol) and 4-dimethylaminopyridine (DMAP, 1.9 mg, 0.2 mmol), stirred at room temperature overnight. The reaction was monitored via TLC using 20% methanol in chloroform v/v, with a product Rf= 0.87. The solvent was removed under vacuum and purified by preparatory Silica column chromatography using the same gradient. The product (Molecule 3-3) isolated was a sticky clear solid, (8.7 mg, 23.5% yield)

Molecule 3-3: 2-(2-(2-(5-((3aR,4R,6aS)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamido)ethoxy)ethoxy)ethyl 4-methylbenzenesulfonate. Rf=0.63 (methanol/chloroform= 1/5). White oily solid. Rf=0.87 using 20% methanol in chloroform. Yield: 8.7 mg, 23.5%. 

$^1$H NMR (400 MHz, DMSO-$d_6$): $\delta = 7.80$ (d, $J = 8.24$ Hz, 2H), 7.36 (d, $J = 8.24$ Hz, 2H), 6.45 (t, 1H), 6.14 (s, 1H), 5.42 (s, 1H), 4.51-4.48 (m, 1H), 4.33-4.30 (m, 1H), 4.170.16 (m, 2H), 3.70-3.44 (m, 10H), 3.14-3.14 (m, 1H), 2.89-2.88 (m, 1H), 2.74-2.71 (m, 1H), 2.44 (s, 3H), 2.31-2.22 (m, 2H), 1.68-1.44 (m, 6H). MS (MALDI-TOF) (m/z) Calculated for C$_{23}$H$_{35}$N$_3$O$_7$S$_2$ [M+H$^+$]: 530.20, observed: 530.7.
1.15.1.4 Molecule 3-4

Molecule 0-4

To a stirred solution of Molecule 0-2 (25.0 mg, 0.07 mmol) in anhydrous degassed dichloromethane was added benzene-sulfonyl chloride (17.1 mg, 0.1 mmol), N,N-diisopropylethylamine (35 μL, 0.20 mmol) and 4-dimethylaminopyridine (DMAP, 1.9 mg, 0.2 mmol), stirred at room temperature overnight. The reaction was monitored via TLC using 20% methanol in chloroform v/v, with a product Rf= 0.87. The solvent was removed under vacuum and purified by preparatory Silica column chromatography using the same gradient. The product (Molecule 3-4) isolated was a sticky clear solid, (8.4 mg, 24.5% yield)

Molecule 3-4: 2-(2-(2-(5-((3aR,4R,6aS)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamido)ethoxy)ethoxy)ethyl benzenesulfonate. = 0.62 (methanol/chloroform= 1/5).

White oily solid. Rf=0.87 using 20% methanol in chloroform. Yield: 8.4 mg, 24.5%. $^1$H NMR (400 MHz, DMSO-$d_6$): $\delta = 8.19$ (d, $J = 7.4$ Hz, 1H), 7.82 (t, $J = 5.5$ Hz, 1H), 7.65 – 7.53 (m, 1H), 7.37 – 7.24 (m, 2H), 6.89 (d, $J = 7.4$ Hz, 1H), 6.37 (d, $J = 22.3$ Hz, 2H), 4.36 – 4.26 (m, 1H), 3.50 (d, $J = 8.3$ Hz, 7H), 3.44 – 3.36 (m, 5H), 3.24 – 3.17 (m, 3H), 3.10 (dd, $J = 8.8$, 4.3 Hz, 1H), 2.83 (dd, $J = 12.4$,
5.1 Hz, 1H), 2.60 (s, 1H), 2.07 (t, J = 7.4 Hz, 2H), 1.67 – 1.22 (m, 10H). MS [M+H; MALDI-TOF]

Calculated for C_{22}H_{32}N_{3}O_{7}S_{2} [M+H]\]: 516.19, observed: 516.8.

1.15.2 Preparation of mRNA fusions

The in vitro selection protocol has previously been outlined in detail and is further described in Appendix mRNA Display Supplemental Information. The sequence of the mRNA display library is described by Figure 0-7 Oligonucleotide sequence of the template\textsuperscript{153}.

**Figure 0-7 Oligonucleotide sequence of the template\textsuperscript{153}**.

![Oligonucleotide sequence diagram]

The library mRNAs were photo chemically cross-linked via UV light to a Puromycin linker after annealing, and underwent translation to produce peptide –mRNA fusions. A 5 mL scale translation was used to generate the fusion for the first round of selection. The mRNA peptide fusions were purified from the translation mixture using Oligo (dT) resin via the poly-A sequence in the Puromycin linker then ethanol precipitated and yield calculated by scintillation counting of $^{35}$S-Methionine. Reverse transcription of the mRNA fusion was purified with Ni-NTA beads via a His Tag on the peptide and the yield quantified by scintillation counting of $^{35}$S-Methionine. The yield of peptide fusions in the first round after all purification steps was 32.5 pmol and this is equivalent to $1.9 \times 10^{13}$ peptides.
1.15.2.1 Selection

10 µL of streptavidin beads (Roche SA Magnetic Particles, 10mg/mL) was calculated to be sufficient to capture 180 pmol of biotin.

Preparation of SA magnetic beads: 10 µl of SA beads were washed 3 times with 600 µL of selection buffer (50 mM Tris-HCl pH 8, 150 mM Nalco, 4 mM MgCl₂, 0.25 % Triton X-100) using a magnet to isolate the beads from the suspension.

Pre-clear: mRNA peptide fusions were added to SA beads and tumbled at room temperature for 30 minutes.

Selection: The pre-cleared mRNA fusions that had been exposed to the SA beads were reacted with 0.1 mM Molecule 3-3 at 4 °C overnight then tumbled at room temperature for one hour. Molecule 3-3 was removed from the mixture via three exchanges of dialysis using selection buffer in a 2,000 MWCO chamber. The mRNA fusion was then added to the SA beads, incubated and mixed well for 30 minutes and the beads isolated and washed using a magnet to isolate the beads from the suspension. Portions of the pre-cleared beads, fusion, flow through, washes were quantified by scintillation counting of ³⁵S-Met. The isolated beads that had been reacted with the mRNA fusion were PCR amplified and the cDNA was transcribed and used for the next round of selection. The scale of each round of selection and all data quantified by scintillation counting is described in Table 0-3. The protocol followed for each step of in vitro selection using mRNA display is described in detail in Appendix B.

1.15.3 Electrophile stability study- imidazole reactivity with Molecule 3-3

TLC Samples: All samples were thoroughly mixed for 10 seconds then centrifuged at 13k rpm for 30 seconds. Aliquots were taken from organic layer by inverting tube and pulling from
bottom either by spotting a capillary tube 10 times or at the 24-hour time point, by pipetting an
8 µL sample per spot. TLC was run using KMnO4 stain in 80% Chloroform and 20% Methanol
testing Various Timepoints-5/ 20/ 40/ 60 minutes and 12 hours

Control 1. Molecule 3-3 -50 µL aliquot of 1 mM Biotin PEO-OTs added to 100 µL CHCl3
Control 2. Imidazole - 20 µL of 1 M aqueous imidazole added to 100 µL CHCl3
Control 3 Elution Buffer - 20 µL of Ni-NTA Elution Buffer added to 100 µL CHCl3

Reaction mixture samples: 40 µL of a 1 mM Molecule 3-3 was added to 860 µL of Ni-NTA
Elution Buffer containing 250 mM imidazole. For each time point a 50-µL aliquot of the
reaction mixture was added to 100 µL CHCl3.
Labeling astrocytes with fluorescent probes

This chapter describes a unique application for our fluorescent probes: labeling astroglia. While the development of this project is still in its infancy stage preliminary results suggest that our Molecule 1-12 analogs may be superior to current stains used to tag live astrocytes. As such, we are currently collaborating with Kurt Hauser, PhD and Shiping Zou in the Department of Pharmacology/Toxicology and Anatomy/Neurobiology at VCU to investigate the effects of our fluorescent probes on neuronal cells with the hopes that they are specific for astroglia and show improved properties currently in use for characterization and study of astroglia.

1.16 Background: selective labeling of astroglia using sulforhodamine 101

Elucidating the mechanisms and supportive role of glial cells in nervous tissue is a longstanding area of interest in the field of neurobiology\(^\text{154}\). Astrocytes are glial cells that make up approximately 50% of cells in the central nervous system and their role in the nervous system was initially overlooked most likely due to the limitation of techniques available to monitor their dynamic properties\(^\text{155}\). Astrogliosis, an abnormal increase of activated astroglial cells due to nervous system injury, has been linked to neurodegenerative disease states\(^\text{156}\) such as epilepsy\(^\text{157}\), stroke, Alzheimer and Parkinson’s\(^\text{158}\) and the molecular mechanisms of astrogliosis are being studied as potential targets for drug development in these disease...
Transgenic and knockout mice have been instrumental in determining the potential effects of astrocyte dysfunction in these disease states. Techniques used to visualize the morphology of astrocytes include electron microscopy and immunohistochemistry (IHC). Most labeling procedures such as IHC using antibodies against glial fibrillary acidic protein (GFAP) or hematoxylin and eosin stain (H&E) require fixation of the astroglial cells, which are unable to be used for tracking live changes with respect to time. Sulforhodamine 101 (Molecule 1-12) is the only known simple specific molecule that labels astroglia in the neocortex. Nimmerjahn and coworkers discovered this simple procedure based on their observation that sulforhodamine 101 is effectively taken up by protoplasmic astrocytes. Their studies suggest that the dye can be spread to neighboring astrocytes via gap junctions. However, the uptake mechanism has not yet been determined.

Advantages of using Molecule 1-12 as a staining reagent include high sensitivity, high contrast imaging, quick staining time (5 minutes) and long lasting effect (on the order of days). One drawback of using Molecule 1-12 as a stain for astroglia is that it leaks in fixed tissue and an analog of Molecule 1-12, Texas Red hydrazide had to be used for those cells analyzed by immunohistochemistry (IHC). The success of in vitro labeling has also been found to be limited by various experimental parameters including temperatures and developmental stages of studied tissue (animal age). The precise mechanism responsible for Molecule 1-12 uptake has not yet been determined, which leads to lack of understanding of the discrepancy between Molecule 1-12 staining of neocortical rat tissue and other central nervous system regions.
Outside of sulforhodamine 101 hydrazide, no other analogs of sulforhodamine have been reported to selectively stain astrocytes. We rationalize that an electrophilic analog of sulforhodamine 101 could be used to covalently label astrocytes, leading to a more permanent fluorescent tag, this could aid in identification of the protein target responsible for its selective uptake. Permanent labeling of astrocytes amongst a mixed sample of neuronal cell types without IHC staining would also be useful for intravital imaging. In addition, the signals regulating astrocyte development are currently not well understood. Working with a permanent stain at different points of maturation could provide a method to culture astrocytes at specific states of development. This would be particularly helpful for working with astrocytes in their non-reactive form as astrocytes are cultured in their reactive states, but the non-reactive state is considered to be reflective of normal in vivo properties. The current methodology employed to define mature astrocytes include examination of the morphology of individual cells, protein markers (GFAP, EZRIN), and dynamic properties. GFAP in particular is a putative biomarker for mature astrocytes.

1.17 Goals and objectives for astroglial staining

We planned to expose neuronal cells including astroglia to our fluorescent probe, Molecule 1-22 and Molecule 1-12 (sulforhodamine 101 used as a control). We wanted to compare staining properties and specificity for astroglia relative to other types of neuronal cells.
1.18 Astrocytes results and conclusions

Both young and mature murine astrocytes cultures were incubated with Molecule 1-22 and Molecule 1-12. In addition, an oligodendroglia cell sample containing a small percentage (5-10%) of astrocytes were also incubated with Molecule 1-22 and Molecule 1-12.

1.18.1 Imaging of young astrocytes

Live cell imaging after a two-hour incubation time revealed that neither Molecule 1-22 nor Molecule 1-12 effectively labeled Day 3 astrocyte cultures. These results were not unexpected, as published characterizations of Molecule 1-12 staining of astrocytes have shown developmental dependence173.

1.18.2 Imaging of mature astrocytes

1.18.2.1 Live cell imaging

Mature astrocytes were exposed to either Molecule 1-22 or Molecule 1-12 for two hours.
Figure 0-1. In this case, the astrocytes were cultured for 8 days from P0-P1 pups, re-suspended and plated again for an additional 3 days. Three individually cultured samples were used for Molecule 1-22 or Molecule 1-12 (n=3) and for each sample, five fields were randomly picked for imaging.

Both Molecule 1-22 (0.45 µM) and Molecule 1-12 (0.50 µM) were found to be cell permeable and in the cytoplasm of the cells. Qualitatively there is a remarkable difference in the staining between the two fluorophores. Molecule 1-22 is not as bright as Molecule 1-12. However, the effective coverage of the cells is enhanced for Molecule 1-22 relative to Molecule 1-12. Within the Molecule 1-12 samples there were uneven pockets of strong fluorescence on some cells but not others. For the Molecule 1-22 stained astrocytes, the coverage showed improved uniformity and less overall background noise for the areas of the sample that did contain cells. In addition, the outline of the nucleus was possible to be visualized, as the stains were very cytosolic. This effect was enhanced and clearly distinct for the Molecule 1-22 fluorophore relative to the Molecule 1-12 stain.
Figure 0-1 Mature astroglia staining with 2 hour incubation time


1.18.2.1.1 Assessment of staining time needed for live cell imaging

The astrocytes that had been stained were checked for residual fluorescence three days later. For both fluorophores, the stains had diffused such that they needed to be re-stained
prior to fixation. Therefore, for both fluorophores at 1 µM concentration with a two hour incubation time the window of effective staining is less than three days.

1.18.2.1.2 Assessment of re-staining properties for live cell imaging

Two astrocyte samples, stained with Molecule 1-12 or Molecule 1-22, were re-stained using the same method 72 hours later. Imaging of the re-stained astrocytes showed that the cells were healthy and viable and it was possible to effectively re-stain the astrocytes with similar staining properties as compared to the initial staining,

Figure 0-2.

Figure 0-2 Re-staining astroglia with fluorescent probes
Examples showing fluorescent re-staining of astrocytes after 72 hours of initial stain and a 2 hour re-stain period. **Left Panels:** Molecule 1-12 (SR101). **Right panels:** Molecule 1-22. **Top Panel:** bright field contrast imaging. **Middle panel:** Texas Red fluorescent filter. **Bottom panel:** overlay of bright field contrast and Texas Red filter.
1.18.2.2 Astrocyte fluorescence post fixation

The astrocytes that had been re-stained with either Molecule 1-12 or Molecule 1-22 were fixed and immunostained using antibodies specific against GFPA. Hoechst were also used to counterstain the cell nuclei.

1.18.2.2.1 Wide field fluorescence imaging of astrocytes post fixation

Fluorescent imaging showed that the astrocyte samples stained with Molecule 1-12 or Molecule 1-22 failed to remain fluorescent after fixation. While this is not an unexpected result for Molecule 1-12, it does raise questions as to the covalency of Molecule 1-22 to an intracellular protein. We would expect that covalent binding would render the cell fluorescent after fixation. This experiment was repeated with an overnight incubation time. However, the astrocytes did not remain fluorescent after being fixed despite the extended labeling time.

1.18.2.2.2 Confocal Imaging of astrocytes post fixation

The mature astrocytes that were stained for two hours as described were imaged via confocal microscopy. Confocal microscopy uses point illumination to provide enhanced optical resolution and contrast utilizing a single plane of vision within a sample. Using the DAPI/GFAP to positively identify the number of cells in a specific field of vision would have made it possible to provide quantitative data to compare staining between Molecule 1-12 and Molecule 1-22. However, the fixed cells were not stained well by either fluorophore, Figure 4-3. As can be seen for both Molecule 1-12 and Molecule 1-22 in Figure 4-5, Panels A and D, there is very little magenta coloration indicating a low level of fluorophore in the astrocyte sample. Panels B and E of Figure 4-5 shows little to none red fluorescence, suggesting no binding of either Molecule 1-12 or Molecule 1-22 remains in fixed cells. Panels C and F, shows GFAP staining and DAPI of
the same field, suggesting that lack of red fluorescence is not due to lack of astrocytes in the field, but rather the loss of the fluorophore. Finally, the bottom panel shows overlap of all filters, rhodamine was pseudo-colored as magenta because red pixelation was used by the confocal software to indicate of GFAP saturation.
Figure 0-3 Confocal imaging of fixed astrocytes with Molecule 1-12 and Molecule 1-22

Examples showing confocal imaging of IHC staining of astrocytes following a 2 hour incubation with Molecule 1-12 (40X magnification) and Molecule 1-22 (10X magnification). Left Panels: Molecule 1-12 Right panels: Molecule 1-22. Panels A/D: Overlay of bright field (gray), Texas Red fluorescent filter (red) and DAPI (blue) microscopy imaging. Panels B/E: Texas Red fluorescent filter. Panels C/F: overlay of GFAP and DAPI imaging. Panels D/G: Overlap of bright field and Texas Red fluorescent filter.
1.18.2.3 Live Cell Imaging overlap with fixed immunohistochemical staining of mature astrocytes.

1.18.2.3.1 Molecule 1-22

Due to the absence of fluorescent staining following fixation, two additional samples of mature astrocytes were prepared as in previous experiments. However, this time the samples underwent live cell imaging, then fixed and stained with GFAP and the same filed that was used in live imaging was found and re-imaged. This provided a direct comparison of astrocytes between Molecule 1-22 and GFAP. To begin, mature astrocytes were incubated with Molecule 1-12 or Molecule 1-22, respectively, for 2 hours at the same concentration (1 µM) as previously tested. 20 random fields of vision taken of Molecule 1-22 treated sample and 8 random fields of Molecule 1-12 treated samples were imaged (see section 4.3.2.3.2). The cells were then fixed and stained with DAPI and GFAP and the same fields were then imaged using GFAP and DAPI staining with the same microscopic settings. Figure 4-4 and Figure 4-5 show sample images (19 of 20 and 6 of 20 respectively).

Both sets of images were examined and the number of astrocytes positively identified by Molecule 1-22 was compared to the number of astrocytes validated by GFAP and DAPI staining. For example, in Figure 4-4, IHC staining showed thirteen astrocytes positively identified by DAPI and GFAP staining. In comparison, the live cell staining identified twelve astrocytes positively using Molecule 1-22. A positive identification for an astrocyte requires both a blue nuclear stain by DAPI and surrounding cytosolic GFAP staining. Live imaging with Molecule 1-22 shows cytosolic staining with a shadow around the nuclei of the astrocytes. In combination with bright field imaging it is possible to positively identify astrocytes without IHC staining.
**Figure 0-4** Molecule 1-22 live cell and fixed cell imaging of stained astrocytes first example

Field of vision 6/20 showing Molecule 1-22 labeled astrocytes using live cell imaging (left panels) and IHC staining (right panels) at 20X magnification. Live cell imaging; Panel A: bright field (gray), Panel B: overlay bright field (gray) and Texas Red filter, Panel C: Texas Red filter. IHC staining; Panel D: bright field (gray), Panel D: Overlay of bright field (gray), GFAP (green) and DAPI (blue), Panel E: GFAP (green), Panel F: DAPI(blue).

In addition, it is possible to identify dead cells and unhealthy cells because they are highly dense in red coloration disproportionately to the healthy astrocytes. This observation is
made more evident by the lack of nuclear shadow. In this example the astrocytes were dense in number and it was still possible to distinguish each cell from one another based on the nuclear shadow patterns seen for those red fluorescent cells.

Field of vision 6 of 20 provided in Figure 4-5 is a second example of astrocytes stained by Molecule 1-22. Eight astrocytes were GFAP positive in this field. Similarly, the live imaging showed eight astrocytes positively labeled by Molecule 1-22. For example, in Figure 4-4, the mid center left most cell is brightly fluorescent without a nuclear shadow, panel C. When compared to imaging after the cells were fixed this cell shows a DAPI stain, but shows a minimal GFAP staining and is not counted as a healthy astrocyte. In contrast, healthy cells stained by Molecule 1-22 show bright cytosolic red fluorescence similar in shape to the GFAP staining but in addition exhibit nuclei that appear shadowed and in the overlay with bright field contrast imaging appear spotted and/or punctuated in coloration.

In comparison, Molecule 1-12 stained astrocytes display much brighter fluorescence. However, the homogeneity of staining renders it difficult to differentiate astrocytes from one another in dense samples. In addition, the nuclear shadowing for Molecule 1-12 staining is far diminished in comparison to Molecule 1-22 staining.
Figure 0-5 Molecule 1-22 Live cell and fixed cell imaging of stained astrocytes second example

Live Cell Imaging

A. Brightfield
B. Overlay Brightfield SR101 Filter
C. SR101 Filter

Post fixation

D. Brightfield
E. Overlay Brightfield GFAP (Green) DAPI (Blue)
F. GFAP (Green) DAPI (Blue)
G. DAPI (Blue)

Field of vision 19/20 showing Molecule 1-22 labeled astrocytes using live cell imaging (left panels) and IHC staining (right panels) at 20X magnification). Live cell imaging; Panel A: bright field (gray), Panel B: overlay bright field (gray) and Texas Red filter, Panel C: Texas Red filter. IHC staining; Panel D: bright field (gray), Panel D: Overlay of bright field (gray), GFAP (green) and DAPI (blue), Panel E: GFAP (green), Panel F: DAPI (blue).
1.18.2.3.2 Molecule 1-12

Molecule 1-12 Contrast of live Cell imaging versus post IHC staining took place in the same manner as the astrocytes labeled with Molecule 1-22 in 4.3.2.3.1. For astrocytes labeled with Molecule 1-12 (Figure 4-6) it was very difficult to differentiate between samples using fluorescence alone. As seen before, the staining is heterogeneous amongst astrocytes clearly labeled by DAPI and GFAP staining. In addition, the nuclear shadowing seen for Molecule 1-22 is greatly diminished.

Figure 0-6 Molecule 1-12 Contrast of live cell imaging versus post immunohistochemical staining first example

Field of vision 6/8 showing Molecule 1-12 labeled astrocytes using live cell
imaging (left panels) and IHC staining (right panels) at 20 X magnification. Live cell imaging; Panel A: bright field (gray), Panel B: overlay bright field (gray) and Texas Red filter, Panel C: Texas Red filter. IHC staining; Panel D: DAPI (blue), Panel E: Overlay of GFAP (green) and DAPI (blue), Panel F: GFAP (green).

Another example from the eight fields of vision sampled for astrocytes labeled with Molecule 1-12 is provided in Figure 4-7.

**Figure 4-7 Molecule 1-12 Contrast of live cell imaging versus post immunohistochemical staining second example**

Field of vision 6/8 showing Molecule 1-12 labeled astrocytes using live cell imaging (left panels) and IHC staining (right panels) at 20X magnification. Live cell imaging; Panel A: bright field (gray), Panel B: overlay bright field (gray) and Texas Red filter, Panel C: Texas Red filter. IHC staining; Panel D: DAPI (blue), Panel E: Overlay of GFAP (green) and DAPI (blue), Panel F: GFAP (green).
1.18.3 Imaging of mixed sample of oligodendrocytes and young astrocytes

In order to determine if Molecule 1-22 can be used as a specific stain for astrocytes in the presence of other neuronal cell types a mixed sample of oligodendrocytes and young astrocytes were incubated for 2 hours with Molecule 1-22 and Molecule 1-12, respectively. Neither fluorophore stained the oligodendrocytes samples. This would be expected as Molecule 1-12 has been shown to be selective for astroglia in the presence of other types of neuronal cells and any astrocytes present in the sample were immature and not expected to stain. However, the results for the Molecule 1-22 were skewed because a large number of cells were no longer viable upon imaging. This experiment needs to be repeated to determine if Molecule 1-22 is cytotoxic for other types of neuronal cells as this is not a desired property.

1.19 Discussion of astrocyte binding with Molecule 1-22

In pursuing labeling of astroglia with our fluorescent probes we have a lot of questions that need to be addressed. First, is there a higher affinity towards labeling astrocytes for our electrophile relative to Molecule 1-12? Is there covalent labeling? If so, why did the cells not remain stained after IHC staining? If we can establish that there is covalent labeling then it might be possible to isolate the protein/biomolecule to which it is attached. This would be highly valuable because the current mechanism for the binding of Molecule 1-12 to astroglia is currently unknown and there are differing theories and conflicting research\textsuperscript{161,174,175}. We can postulate covalent labeling if fixated cells show better staining that diffuse Molecule 1-12 but the preliminary studies did not show this. As such, we need to repeat the labeling, and possibly optimize certain reaction conditions including time of incubation, concentration of fluorophore, and concentration of DMSO. If there is enhanced labeling we can use quantitative analysis by
cell counting to evaluate a difference between Molecule 1-12/sulforhodamine 101-hydrazide and our fluorophore. In addition, we attempted to view the cells without washing them but there was too much background noise, if we can bring the concentration down is affinity high enough to remove free Molecule 1-22 from solution to provide wash free conditions? What about selectivity? We need to address toxicity/death of cells at 1 μM and retest in a mixed oligodendrocytes sample to see if newly differentiated astrocytes are stained in presence of other cell types. Finally, if we modulate the electrophilicity of the leaving group on our Molecule 1-12 analog would this aid in controlling the labeling properties?

1.20 Astroglia experimental procedures

1.20.1 Preparation of fluorophores

Sulforhodamine 101 (SR101, Molecule 1-12) and Molecule 1-22 were prepared as 1 mM stock solutions in biological grade endotoxin free DMSO and concentrations determined by UV.

1.20.2 Preparation of astrocytes

Primary cultures were prepared from mice (CD-1, Charles River Laboratory, Wilmington, MA) at postnatal day 0-to-1 as previously published In brief, whole brains were dissected and minced before being incubated (37 °C, 5% CO₂) with 2.5 mg/ml trypsin (Sigma, St. Louis, MO) and 0.015mg/ml DBase (Sigma) in Dulbecco’s Modified Eagle’s Medium (DMEM) (Life Tech, Carlsbad, CA) for 30 min. Tissue was triturated and re-suspended in DMEM supplemented with 10% fetal bovine serum (Thermo Scientific, Waltham, MA), 6 g/L glucose (Sigma), sodium bicarbonate, and penicillin/streptomycin (Life Tech) and filtered through 135 μm and 45 μm nylon mesh, and then plated in 35mm culture dish (Corning Inc., Corning, NY) pre-coated with
poly-L-lysine (Sigma) with a density of $1 \times 10^6$ cells/dish. Two samples types of astrocytes were tested for culture. Sample set one, young astrocytes, were cultured 3 days from P0-P1 mice pups. The second sample set, mature astrocytes, are cultured 8 days from P0-P1 Pups first, then re-suspended and plated again for an additional 3 days.

**1.20.3 Incubation of astrocytes with fluorophore**

For each sample of cells containing either astroglia or a mixture of oligodendrocytes, 6 µL of a 1 mM of fluorophore solution was taken up into 12 mL of media (DMEM supplemented with 10% fetal bovine serum, Thermo-Scientific) and 2 mL dispersed for each sample to give a final 0.5 µM concentration of Molecule 1-12 with 0.05% DMSO in media and a 0.45 µM concentration of Molecule 1-22 with 0.05% DMSO in media. The samples were incubated for two hours and washed twice with an additional 2 mL of media before imaging took place.

**1.20.4 Qualitative assessment of astrocytes using live cell imaging and fixated cells**

After a 2 hour treatment, the dishes containing the samples were transferred to a heating insert (Pecan, Erbach, Germany) and put on the scanning stage of a Zeiss Z1 Observer microscope (Carl Zeiss Microscopy, LLC, Thornwood, NY) equipped with an environmental incubator (Pecon) that keeps a constant temperature (37 °C) and CO$_2$ level (5%) during the experimental period. Both the heating insert and the environmental incubator were pre-warmed to 37 °C at least 1 h before the start of the experiment. For each treatment, more than 8 fields were systematically yet arbitrarily chosen and imaged using an automated stage controlled by the Zeiss Axiovision 4.8 software (Carl Zeiss). Fluorescence emission of Molecule 1-12 and Molecule 1-22 were detected between 610 and 630 nm.
1.20.5 Assessment of fixated astrocytes using confocal imaging

Confocal immunofluorescence images were taken using a Zeiss LSM700 laser scanning confocal head configures to a Zeiss Axio Observer Z.1 and processed using Zen 2010 software.

1.20.6 Immunohistochemical staining

Both DAPI and GFAP were used for overlay experiments on fixed cells.
Overall Conclusions/ Discussion

Our aim to find a short nucleophilic peptide sequence that would covalently react with a small labeling molecule was a complex challenge. To date, there have been no short peptide sequences known to covalently label a small molecule without the aid of an enzymatic process or metal chelation (outside of peptides 501/512 peptides which were discovered to have micromolar binding affinities). Our initial attempt to elucidate the structure of such a peptide utilized mRNA display to create a peptide library that provided access to a systematic combination of a large number of peptides. Following 8 rounds of selection we did not find enrichment and sequencing results were not obtained. One significant drawback to mRNA display is that the process of obtaining results is lengthy. Despite the lack of enrichment it would have been useful to sequence Rounds 7 and 8 to determine if there were any comparable trends. In addition, had there been any repeated peptides in the pool these could have been synthesized and tested against the electrophile to determine the rate of reactivity and selectivity and helped to establish whether or not using mRNA display to find short peptide sequences with high and specific reactivity towards an electrophile is a useful tool and worth repeating with other types of electrophiles.

Having realized the challenge of finding high specificity of a peptidic nucleophile towards a specific small molecule we chose to utilize the binding affinity between peptides 501
and 512 with rhodamine B and sulforhodamine 101. The reported picomolar binding affinities and nanomolar solution phases binding affinities were hypothesized to provide us the selectivity needed to utilize a proximity effect and this in turn would create a flexible tool to label proteins of interest with a short peptidic aptamer using a small fluorescent molecule. Towards this goal, we developed several electrophilic fluorescent small molecules containing three key components. The synthetic routes we developed to access our electrophilic rhodamine B and sulforhodamine 101 fluorophores were optimized. Key intermediates were produced that served as flexible points of modification to make various analogs of our desired electrophilic fluorophores and allowed us to attach other probes such as biotin.

Development of a rhodamine based electrophile proved difficult using rhodamine B (Molecule 1-1) as the starting material. For rhodamine B the competitive intramolecular cyclization as a competing reaction for derivatization led to synthesis of the piperazine intermediate (Molecule 1-6) with a tertiary amide that could not cyclize. However, utilizing a PEG linker as the source of the hydroxyl functional group for tosylation led to an unstable rhodamine electrophile. Based on these results we concluded that the lack of stability of the product formed based on degradation with time observed on isolated products using NMR and MALDI-TOF analysis was believed to be the source of the challenges in isolating the desired sulfonate esters.

The difficulties associated with tosylation of the PEG linker on the rhodamine B analogs were also observed for the sulforhodamine 101 analogs despite the multiple methods of optimization. Thereafter, we chose to bypass formation of the tosylate via the PEG alcohol and instead attached the sulfonate ester directly via a [3+2] cycloaddition reaction. The sulfonate
ester linked to sulforhodamine 101 via a triazole was found to be stable. Another complication in the route of synthesis towards our sulforhodamine 101 analogs was formation of a positively charged bis-azole sulforhodamine 101 salt that co-eluted with the desired mono-azole using HPLC and flash column chromatography. The best method to minimize these purification difficulties was to first attempt to minimize the formation of the bis-sulfonyl chloride and secondly to utilize a sterically hindered alcohol (isopropanol) with a lower polarity than methanol with dichloromethane to optimize separation of the bis analog as well as providing a means of separation for the ortho and para- mono-substituted isomers.

Optimization of the route of synthesis towards electrophilic sulforhodamine 101 analogs was accomplished. In addition, the purification methods provide a means to separate the ortho/para isomers from one another. Commercially, the sulforhodamine 101 analogs that are currently available for derivatization are economically restrictive and are not isomerically resolved. In addition, the azide analog prepared creates a highly flexibly intermediate that provides the means to derivatization of sulforhodamine 101 with nucleophiles, electrophiles and other useful analogs. Once we had made the electrophilic fluorophores we tested their ability to function as a protein tag by testing them with proteins that contained peptides 512 and 501.

The affinity between the proteins containing peptides 501 and 512 and the fluorescent electrophiles were determined by fluorescence polarization and were found to be in the low to mid micromolar range and the resulting affinity between the electrophilic fluorophore was not high enough to produce selective labeling\(^\text{14,16}\). While Nolan and coworkers reported picomolar affinities utilizing phage display and low nanomolar affinities with the peptides in solution
phase we did not find this level of binding affinity between the 501/512 aptamers and sulforhodamine 101. Our experiments suggest that the binding affinities in solution phase between the peptides and the fluorescent electrophiles were no higher than the 40-54 µM range.

The next step for this project would be to incorporate a better affinity partner for the fluorophore. Unfortunately, there is a large gap in what is available in terms of small peptides with a specific affinity for small molecules. Biologically speaking, large proteins with binding sites and pockets are much more effective at producing microenvironments that optimize the conditions necessary for the reactivity in their active sites. We had considered using a monoclonal antibody that could be used as a binding partner for sulforhodamine 101 (Molecule 1-12). While there are antibodies available for various rhodamine analogs, they specifically do not work for sulforhodamine 101.

While we were attempting to test and optimize the binding affinities between the our electrophilic fluorophore and peptides 501 and 512 we wanted to determine if covalent labeling was in fact taking place between the two entities. We found that two of our sulforhodamine 101 sulfonate ester analogs (Molecule 1-22 and 1-19) labeled proteins non-selectively in a time and concentration dependent manner in biologically applicable media leading us to the conclusion that sulfonate esters can be utilized as labeling reagents for in vitro protein studies.

In addition, we found that sulforhodamine 101 (Molecule 1-12) is used in neurobiology as a label for astroglia and we decided to determine if the binding affinity for our electrophilic fluorophore might covalently label this type of glial cell. Our electrophilic fluorophore was
found superior to sulforhodamine 101 in live cell imaging of astrocytes. Sulforhodamine 101 is a stain currently used to selectively label astrocytes\textsuperscript{161,162,177} in the presence of other neuronal cell types for live cell imaging applications. These recent findings lead us to postulate additional questions about the mechanism by which sulforhodamine 101 stains astrocytes and we would like to determine if covalent labeling of the astrocytes could provide information that would help elucidate the mechanism or provide additional information about astroglia selectively in mixed samples utilizing live cell imaging without the need for post immunohistochemical staining.
List of references

(1) Shulman, J. Survey of Ph.D. Programs in Chemistry

(2) Valerie J. Kuck. Women Physicists and Chemists are Making Slow Progress in Academe


(39) The Nobel Prize in Chemistry 2008


(75) Chemical Suppliers - CHEMCATS http://www.cas.org/content/chemical-suppliers (accessed Jul 26, 2014).
(77) SciFinder - The choice for chemistry research


(78) Molecular Probes Inc Legal Notices and Trademark Attributions


(86) Ledgard, J. A Laboratory History of Chemical Warfare Agents; Lulu.com, 2007; p. 259.

(87) The Nobel Prize in Chemistry 2001


(92) Chen, Z.; Popp, B. V.; Bovet, C. L. *ACS Chem. Biol.* **2011**.


(96) Lin, M. Z.; Wang, L. *Physiology (Bethesda).* **2008**, *23*, 131–141.


(105) BioTechniques - Direct pH measurements by using subcellular targeting of 5(and 6-) carboxyseminaphthorhodafluor in mammalian cells


(118) ISS | Technical Notes | Fluorescence Polarization

(119) SigmaPlot: Standard Curves Analysis

(120) Matsumura, I. Biotechniques 2010.


(123) Peptide Properties Calculator


(126) Paley, W. Natural theology or, evidence of the existence and attributes of the deity, collected from the appearances of nature, 2006.


(153) White, E. R. Selection of a Non-Phosphorylated Peptide Inhibitor of BRCA1’s (BRCT)2 Domain - ProQuest Dissertations & Theses Full Text - ProQuest, Virginia Commonwealth University, 2013.


(168) *NEUROINFLAMMATION IN NEURONAL DEATH AND REPAIR*; Elsevier, 2007; p. 504.


(178) Oligonucleotide Properties Calculator


Appendix A: NMR and MALDI Spectra

Molecule 1-1: Rhodamine B (RhB)

Molecule 1-1: 1H NMR 400 MHz DMSO-d6
Molecule 1-1: C-13 NMR, DMSO-d$_6$
Molecule 1-1: MALDI-TOF CHCA Matrix Linear Positive Mode
No TFA Used in matrix preparation

SDS-4-21-RhBNon Sample spot=G3
SDS-4-21-RhBNon 10 (1.115) Sb (15,10,00), Sm (Mn, 1x5,00), Cm (2,14)

TOF LD+
7.59e3
Molecule 1-2: Rhodamine B Isothiocyanate Starting Material

Molecule 1-2: 1H NMR 400 MHz DMSO-d6
Molecule 1-3

Molecule 1-3: 1H NMR 400 MHz DMSO-d6
Molecule 1-3: MALDI-TOF CHCA Matrix Linear Positive Mode

Chemical Formula: \( C_{35}H_{45}N_4O_6S^{+} \)

Exact Mass: 649.31
Molecular Weight: 649.82
Molecule 1-4

Molecule 1-4: 1H NMR 400 MHz DMSO-d$_6$
Molecule 1-4: MALDI-TOF CHCA Matrix Linear Positive Mode
(Crude Reaction mixture)
sds-141-RMdlute Sample spot = F5
sds-14 l-RMdlute  S (2.874) Sb (15.10.00) Sm (Mn, 1x5.00), Cm (2.9)
Molecule 1-5: Rhodamine B Base

Molecule 1-5: 1H NMR 400 MHz DMSO-d6

\( ^1H \) NMR (400 MHz, DMSO-d6) δ 7.98 (dd, \( J = 7.5, 1.1 \) Hz, 1H), 7.75 (dd, \( J = 28.7, 7.5, 1.1 \) Hz, 1H), 7.25 (d, \( J = 7.6 \) Hz, 1H), 6.52–6.33 (m, 4H), 3.35 (q, \( J = 7.9 \) Hz, 2H), 1.09 (t, \( J = 7.0 \) Hz, 4H).

Molecule 1-5: Rhodamine B Base C13 Carbon NMR in DMSO-d6

\( ^13C \) NMR (100 MHz, DMSO-d6) δ 143.38, 138.36, 129.70, 140.77, 126.94, 129.22, 128.95, 128.82, 128.80, 119.99, 118.19, 117.57, 118.08, 95.49, 65.36, 56.03, 52.78, 13.86.
Molecules 1-1 and 1-5 Overlap Proton NMR with Rhodamine Starting Material in DMSO-d6
Molecules 1-1 and 1-5, respectively: MALDI-TOF CHCA Matrix Linear Positive Mode

Chemical Formula: C_{28}H_{31}N_{2}O_{3}^+
Exact Mass: 443.23
Molecular Weight: 443.56
Molecule 1-6: 1H NMR 400 MHz DMSO-d6

1H NMR (400 MHz, DMSO-d6) δ 10.17 (s, 2H), 7.91 – 7.68 (m, 2H), 7.52 (d, J = 5.8, 5.1 Hz, 1H), 7.19 – 7.04 (m, 4H), 6.94 (d, J = 2.2 Hz, 2H), 3.68 (dq, J = 17.9, 10.0, 7.1 Hz, 2H), 3.00 (s, 2H), 2.93 – 2.79 (m, 3H), 1.21 (s, J = 7.0 Hz, 12H).
Molecules 1-6 (bottom) and 1-5 (top): DMSO 400 MHz Proton NMR
Molecule 1-6: $^{13}$C NMR 400 MHz DMSO-d$_6$
Molecule 1-6: MALDI-TOF CHCA Matrix Linear Positive Mode

SDS-5-22-SM-NOM Sample spot=D.3
SDS-5-22-SM-NOM 11 (0.612) 5b (15.10.00); Sm (Mn, h<5 90); Cm (2-14)

Chemical Formula: C_{32}H_{39}N_{4}O_{2}^{+}
Exact Mass: 511.31
Molecular Weight: 511.68
Molecule 1-7

Molecule 1-7 (bottom) and triethylene glycol (top): 400 MHz Proton NMR DMSO-d6
Molecule 1-7: C13 HNMR 400 MHz in DMSO-d6
Molecule 1-8

400 MHz 1H Proton NMR in DMSO-d6

Molecule 1-7 (top)
Molecule 1-6 (middle)
Molecule 1-8 (bottom)
400 MHz C13 Carbon NMR in DMSO-d6
Molecule 1-7 (top)
Molecule 1-6 (middle)
Molecule 1-8 (bottom)
Molecule 1-8: MALDI-TOF CHCA Matrix Linear Positive Mode
SDS-102512-134-LCP-Vydac Sample spot=G,5
SDS-102512-134-LCP-Vydac 5 (0.414) Sb (15,10.00); Sm (Mn, 1x5.00); Cr (1:15)
TOF LD+
1.1864
Molecule 1-9

Molecule 1-9: 400 MHz Proton NMR DMSO-d$_6$
Molecule 1-9: MALDI-TOF CHCA Matrix Linear Positive Mode
SDS:5-44-RhBPipOHRM1 Sample spot=B,3
SDS:5-44-RhBPipOHRM1 7 (1.647) S6 (15.1000); Sm (Mn, 1x5.00); Cm (1.7)

TOF LD+
2.18e3
Molecule 1-10

Molecule 1-10: 400 MHz 1H Proton NMR in DMSO-d6
Molecule 1-11

Molecule 1-11: 400 MHz Proton NMR CDCl₃
Molecule 1-11: MALDI-TOF CHCA Matrix Linear Positive Mode

SDS-120312-6-42-FP-CHCA Sample spot=G,3
SDS-120312-6-42-FP-CHCA 9 (0.713) Sb (15.10.00 ); Sm (Mn, 1x5.00); Cm (2:15)
Molecule 1-12: Sulforhodamine 101 (SR101)

Molecule 1-12: 400 MHz Proton NMR CDCl₃, commercially available (Biotium)
Molecule 1-12: MALDI-TOF CHCA Matrix Linear Positive Mode

(Magnification of above)

607.5, 51297]
Molecule 1-13: Sulforhodamine 101 Sulfonyl Chloride (Texas Red)

Molecule 1-13: 400 MHz Proton NMR 1:1 CDCl₃:DMSO
Molecule 1-14: 400 MHz Proton NMR CDCl₃
Molecule 1-14: MALDI-TOF CHCA Matrix Linear Positive Mode (Magnification Below)
Molecule 1-15

Molecule 1–15: 400 MHz Proton NMR CDCl₃
Molecule 1-15: MALDI-TOF CHCA Matrix Linear Positive Mode
Molecule 1-17

Molecule 1-17: 400 MHz Proton NMR CDCl₃
Molecule 1-17: MALDI-TOF CHCA Matrix Linear Positive Mode (Magnification below)
Molecule 1-18

Molecule 1-18: 1H-NMR: CDCl$_3$, 400 MHz
Molecule 1-18: MALDI-TOF CHCA Matrix Linear Positive Mode
Molecule 1-19

Molecule 1-19: 400 MHz Proton NMR CDCl₃
Molecule 1-19: MALDI-TOF CHCA Matrix Linear Positive Mode (Magnification Below)
Molecule 1-20

Molecule 1-20: 400 MHz Proton NMR CDCl₃

$^1$H NMR (400 MHz, Chloroform-d): 8.83 (s, 1H), 8.17 (dd, J = 11.2, 7.9 Hz, 2H), 7.71 (d, J = 7.9 Hz, 1H), 6.52 (s, 1H), 3.54 (p, J = 7.2 Hz, 2H), 1.31 (s, 3H).
Molecule 1-21: 400 MHz Proton NMR CDCl₃
Molecule 1-21: MALDI-TOF CHCA Matrix Linear Positive Mode (Magnification Below)
Molecule 1-22: 400 MHz Proton NMR DMSO-d6 (Top) with Molecule 1-24 (Bottom)
Molecule 1-22: MALDI-TOF MS CHCA Positive Linear Mode
Molecule 1-23

Molecule 1-23: 400 MHz Proton NMR CDCl₃
Molecule 1-23: MALDI-TOF MS CHCA Positive Linear Mode (Magnification Below)
Molecule 1-24

Molecule 1-24: 400 MHz Proton NMR DMSO-d6 (Bottom) with Molecule 1-22 (Top)
Molecule 1-24: MALDI-TOF MS CHCA Positive Linear Mode (Magnification Below)
Molecule 2-1 501 Peptide

Molecule 2-1: MALDI-TOF MS CHCA Positive Linear Mode (Magnification Below)
Molecule 2-2 512 Peptide

Molecule 2-2: MALDI-TOF MS CHCA Positive Linear Mode (Magnification Below)
Molecule 3-1

1H NMR 400 MHz DMSO-d6
Biotin Starting Material for Synthesis of Molecule 3-1:
Molecule 3-1: MALDI-TOF MS DHB Matrix Positive Linear Mode
Molecule 3-2

Molecule 3-2: 1H NMR 400 MHz DMSO-d6
Molecule 3-2: MALDI-TOF MS DHB Matrix Positive Linear Mode
Molecule 3-3

Molecule 3-2: 1H NMR 400 MHz CDCl₃

Molecule 3-3: MALDI-TOF MS DHB Matrix Positive Linear Mode
Molecule 3-4

Molecule 3-2: 1H NMR 400 MHz CDCl₃
Appendix B: mRNA display supplemental information

DNA template

The cDNA construct was made on a DNA synthesizer. The components of the construct are described in further detail in Figure 3-8.

Transcription (DNA to mRNA)

The DNA template (Figure B-0-1) undergoes transcription from DNA to mRNA, single nucleotide strands that carry the “code” for protein synthesis from the DNA to the ribosome.

Figure B-0-1 Random DNA sequence

CX12: 123-mer:
CTA GCT ACC TAT AGC CGG TGG TGA TGG TGA TGG TGG CCT AAG CTA CCG GAG CCG GAC CGC TGA GCC ACC TTA ATG GCG CAG AAC CAC ATG CAC ATT TAG CTG TCC TCC TTA CTA AAG TTA ACC

The concentration of the dsDNA PCR product was measured with the spectrophotometer, and then calculated using the Oligonucleotide Properties Calculator (http://www.unc.edu/~cail/biotool/oligo/index.html). The following components were combined based on the measured concentration and incubated at 37 °C overnight (Table B-0-1)

Table B-0-1 Transcription reaction conditions

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Stock</th>
<th>Final</th>
<th>1 ml Rxn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td></td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>Tris/Triton</td>
<td>400 mM/0.1% v/v</td>
<td>40 mM/0.01% v/v</td>
<td>100</td>
</tr>
</tbody>
</table>
The contents of the transcription reaction then underwent DNase treatment by removing the transcription reaction from the incubator to which was added 1 µL turbo DNase/ 20 µL transcription (50 µl for 1 ml) and incubated for 15 min at 37 °C. (Turbo DNase: Ambion #2238)

**mRNA purification via urea gel**

An 8% Sequagel in 1X TBE was used to purify the transcription reaction. The reaction was prepared by the addition of 8M Urea (1mg/2.3µl) and a 10 µl /500 µl of Urea Loading Buffer Dye whereby a 10ul CX12 mRNA was used as a reference in a second sample well. The band was visualized with a 254 nM UV lamp and cut out.

**mRNA isolation from urea gel via electroelution**

The mRNA in the sequagel was purified via electroelution for 2 hours at 300 V, the electrodes switched for the last minute and the purified solution further purified by ethanol precipitation.

**Crosslink psoralen to mRNA**

mRNA is cross-linked to XL Oligo:5’-PsoC6-(tagccggtg)2’-Ome-15xA-2xspacer9-ACC-Puro-3’, (Figure B-2).
The mRNA concentration was determined using a Oligonucleotide Properties Calculator inputting the sequence as ssRNA with the measured OD from the spectrophotometer as described by Figure B-3.

**Figure B-0-3 Random CX12 mRNA sequence: (123 mar)**

| CUA GCU ACC UAU AGC CGG UGG UGA UGG UGG CCU CCA CCG GAG CCG GAC CGC UGA GCC ACC UUA AUG GCG CAG AAG CAC AUG CAC AUU UAG CUG UCC UCA UUA CUA AAG UUA ACC |

Distilled, deionized water was added to give a final concentration of 3 µM to which the following compounds were used for the crosslinking, Table B-0-2:

**Table B-0-2 Psoralen crosslinking reagents**

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Stock Concentration</th>
<th>Final Concentration</th>
<th>µL added (100 µL rxn)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mRNA</td>
<td>Variable</td>
<td>3 µM</td>
<td>Variable</td>
</tr>
<tr>
<td>XL Oligo</td>
<td>133 µM</td>
<td>7.5 µM</td>
<td>5.64</td>
</tr>
<tr>
<td>HEPES</td>
<td>1 M</td>
<td>20 mM</td>
<td>2</td>
</tr>
<tr>
<td>KCl</td>
<td>1 M</td>
<td>100 mM</td>
<td>10</td>
</tr>
<tr>
<td>Spermidine</td>
<td>25 mM</td>
<td>1 mM</td>
<td>4</td>
</tr>
<tr>
<td>EDTA</td>
<td>500 mM</td>
<td>1 mM</td>
<td>0.2</td>
</tr>
<tr>
<td>ddH$_2$O</td>
<td>Variable</td>
<td>Variable</td>
<td>Variable</td>
</tr>
</tbody>
</table>
Each 100 µL reaction was sufficient for a 250 µL translation reaction. This material was added to PCR tubes and heated to 70 °C for 5 min, cooled to 25 °C over 5 min (total run time was 20 min). The reaction mixture was then transferred to a clear crosslink plate and irradiated with a filtered UV light at 366 nm for 20 minutes at 4°C. Multiple samples were combined and precipitated in ethanol using 10% 3M KOAc, 3X volume 100% EtOH, frozen for 30 min, spun at 4 °C for 20 min, supernatant discarded and washed twice with 500 µl of 70% EtOH in water, followed by air drying for at least ten minutes.

**Translation (mRNA to peptide fusion library)**

A translational library uses mRNA to make a peptide library. The advantages of this process compared to other in vitro libraries such as phage display and ribosome display libraries include enhanced diversity, reduced cost, and time efficiency in addition to the flexibility in terms of scale up.

**Preparation of working area**

Initially, pipette tips, gloves and surface of working area were washed with RNAse free spray. Small molecules and buffers were prepared first and then enzymes were added last to prevent loss of activity. To begin, polymix buffer 3X, a buffer that promotes translation was freshly prepared.

**Preparation of polymix 3X**

Polymix 3X is an inorganic mixture of small molecules and buffers that promotes translation and prepared in advance of enzyme addition to minimize loss of enzyme activity, Table B-0-3.
Table B-0-3 Polymix components

<table>
<thead>
<tr>
<th>3X Polymix</th>
<th>3X</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 µL 2.4 M putrescine</td>
<td>24 mM</td>
<td>8 mM</td>
</tr>
<tr>
<td>100 µL 300 mM spermidine</td>
<td>3 mM</td>
<td>1 mM</td>
</tr>
<tr>
<td>100 µL 1.5 M KH₂PO₄</td>
<td>15 mM</td>
<td>5 mM</td>
</tr>
<tr>
<td>1 mL 2.85 M KCl</td>
<td>285 mM</td>
<td>95 mM</td>
</tr>
<tr>
<td>100 µL 1.45 M NH₄Cl</td>
<td>15 mM</td>
<td>5 mM</td>
</tr>
<tr>
<td>pH Adjust 600 mM HCl</td>
<td>~135 µL</td>
<td></td>
</tr>
<tr>
<td>100 µL 1.5 M MgOAc</td>
<td>15 mM</td>
<td>5 mM</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>Variable</td>
<td>Variable</td>
</tr>
<tr>
<td>100 µL 150 mM CaCl₂</td>
<td>1.5 mM</td>
<td>0.5 mM</td>
</tr>
</tbody>
</table>

Polymix Procedure: To a 15 ml conical tube with 4 mL of water were added putrescine, spermidine, KCl and KH₂PO₄ and NH₄Cl and mixed well to give a pH of approximately 9.1 and the pH adjusted to 7.70-7.80 using 600 mM HCl. If excess HCl was added, potassium hydroxide was then used be added to increase the pH as a potassium base is preferred as sodium salts can decrease translation yields.

Once the solution was pH adjusted, MgOac was added to the polymix tube. To a separate 15 mL conical vial were added 3.0 mL water and the calcium chloride. The two parts were then mixed. The calcium chloride was added after dilution to prevent transient precipitation upon mixing. The solution was brought to 9.75 mL in the Falcon tube and filtered with a syringe tip filter (Acrodisk 25 mm) into a new 15 mL conical tube. The 3X polymix buffer was prepared fresh and stored at room temperature as all reagents are not temperature or vortex sensitive but components may precipitate out of solution if cooled.

Preparation of master mix

The translation Master Mix was prepared as described in Table B-0-4. Note that the Master Mix is a mixture of all components minus those components that act as the variable and are added...
in varying concentrations. Thus the master mix can be added to multiple translations as a homogenous mixture. Preparing a master mix saves time as opposed to adding multiple small quantities of components, and maximizes enzyme quality and accuracy with regards to volume measurement using a pipette.
### Table B-0-4 Components of a translation master mix

<table>
<thead>
<tr>
<th>Master Mix components</th>
<th>Location/ Direction for preparation</th>
</tr>
</thead>
<tbody>
<tr>
<td>water</td>
<td>Calculated last as headspace to give the final desired concentration of all components. Use distilled water.</td>
</tr>
<tr>
<td>1. Polymix buffer</td>
<td>fresh (3X)- Preparation and location of components described previously.</td>
</tr>
<tr>
<td>2a.DTT (fresh)</td>
<td>Prepared fresh from pre-weighed solids, add 10X prescribed amount of ddH₂O water to get 100mM solution. DO NOT VORTEX. Gently finger tap to avoid oxidation.</td>
</tr>
<tr>
<td>2b.Methionine (fresh)</td>
<td>Prepared fresh from pre-weighed solids add 10X prescribed amount of ddH₂O water to get 100mM solution, then dilute 1:100 to get a final 1 mM concentration. DO NOT VORTEX. Gently finger tap to avoid oxidation.</td>
</tr>
<tr>
<td>Labeled Met</td>
<td>Prepared fresh from pre-weighed solids add prescribed amount of ddH₂O water to get 100mM solution, DO NOT VORTEX. Gently finger tap to avoid oxidation.</td>
</tr>
<tr>
<td>Labeled Cyst</td>
<td>Prepared fresh from pre-weighed solids add prescribed amount of ddH₂O water to get 100mM solution, DO NOT VORTEX. Gently finger tap to avoid oxidation.</td>
</tr>
<tr>
<td>3. creatine Kinase</td>
<td>Defrost 5X simple buffer from -80°C freezer and dilute to 1X. Add a couple of flakes with a clean spatula and transfer into a 600 µL centrifuge tube. Add 10 µL of 1X buffer to the creatine kinase and then obtain the optical density using the spectrophotometer.</td>
</tr>
<tr>
<td>4. nucleotide PP kinase</td>
<td>Sigma Aldrich, as is.</td>
</tr>
<tr>
<td>4. myokinase (rabbit)</td>
<td>Sigma Aldrich, as is.</td>
</tr>
<tr>
<td>5. Inorganic Pyrophosphatase</td>
<td>Located in -20°C freezer in a cryobox.</td>
</tr>
<tr>
<td>6. 18 AA Mix (Melissa)</td>
<td>-20°C Freezer, can be reused.</td>
</tr>
<tr>
<td>6. methyl tetrahydrofolate</td>
<td>-80°C freezer, can be reused, stored in individual boxes</td>
</tr>
<tr>
<td>6. creatine phosphate (CrP)</td>
<td>-80°C freezer, can be reused, stored in individual boxes</td>
</tr>
<tr>
<td>6. ATP +GTP</td>
<td>-80°C freezer, can be reused, stored in individual boxes</td>
</tr>
<tr>
<td>6. normal/deacylated E. coli tRNA (labeled 2 hours)</td>
<td>Normally used for natural AA translations, labeled tRNA, Deacylated can be used for either, labeled 2 hour tRNA, both stored at -80°C, can be reused, stored in individual boxes</td>
</tr>
<tr>
<td>9 Factor mix</td>
<td>-80°C freezer, cannot be reused.</td>
</tr>
<tr>
<td>EF-Tu (labeled TU)</td>
<td>-80°C freezer, cannot be reused.</td>
</tr>
<tr>
<td>Natural AARS(labeled DOT),</td>
<td>-80°C freezer, cannot be reused. For natural amino acid translation, add MRS AND Phe</td>
</tr>
<tr>
<td>Un-nat AARS (labeled U)</td>
<td>-80°C freezer, cannot be reused.</td>
</tr>
<tr>
<td>Ribosome (labeled R)</td>
<td>-80°C freezer, cannot be reused.</td>
</tr>
<tr>
<td>35S Met</td>
<td>-80°C freezer, cannot be reused.</td>
</tr>
<tr>
<td>Master Mix</td>
<td>All of the components above EXCEPT for the variables (typically hot vs. cold 35-S Methionine</td>
</tr>
<tr>
<td>mRNA</td>
<td>XL cross-linked mRNA, Tim, -20, OD 1.6</td>
</tr>
<tr>
<td>MVHM-His6</td>
<td>Positive Control</td>
</tr>
</tbody>
</table>
Translation of mRNA fusion reaction mixture

The order of addition was such that the small molecules were prepared first, followed by the more sensitive enzymes. Once the polymix buffer was prepared then all of the materials were mixed together with the exception of the ribosome. To insure homogeneity, the enzymes were mixed each time by gently finger flicking the tube(s). Mixing was accomplished by tube inversion followed by a gentle spin down. For all translations once everything was premixed premixing was complete, mRNA thawed on ice was added last and the tubes incubated for 60 minutes at 37 °C. The reaction was quenched by adding a 30% volume to volume MgCl2/KCl mixture.

Note that water was used as headspace such that concentrations of components could be changed. Any components added with a volume of less than a microliter were diluted to increase the volume of the component and increase accuracy. Below is the translation spreadsheet used with typical stock and final concentrations of each component, Table B-5.
# Table B-0-5 Translation components

<table>
<thead>
<tr>
<th>Master Mix components</th>
<th>Stock Source and/or Concentration</th>
<th>Final Concentration</th>
<th>Sample 1 (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>water</td>
<td></td>
<td></td>
<td>31.5</td>
</tr>
<tr>
<td>XL cross-linked mRNA (Trans9)</td>
<td>OD(1.1)</td>
<td></td>
<td>6.66</td>
</tr>
<tr>
<td>Polymix buffer</td>
<td>fresh (3X)</td>
<td>1X</td>
<td>333.4</td>
</tr>
<tr>
<td>DTT (fresh)</td>
<td>100 mM</td>
<td>1 mM</td>
<td>10</td>
</tr>
<tr>
<td>Methionine (fresh)</td>
<td>1 mM</td>
<td>10 µM</td>
<td>10</td>
</tr>
<tr>
<td>Cysteine (freshly prepared)</td>
<td>100 mM</td>
<td>200 µM</td>
<td>2</td>
</tr>
<tr>
<td>Creatine Kinase C</td>
<td>fresh OD 10</td>
<td></td>
<td>0.8</td>
</tr>
<tr>
<td>nucleotide PP kinase</td>
<td>Sigma</td>
<td></td>
<td>1.6</td>
</tr>
<tr>
<td>myokinase</td>
<td>Sigma</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Inorganic Pyrophosphatase (Ppase)</td>
<td>Stock 05-09</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>18 AA Mix</td>
<td>10 mM</td>
<td>200 µM</td>
<td>20</td>
</tr>
<tr>
<td>methyl tetrahydrofolate</td>
<td>6 mM</td>
<td>30 µM</td>
<td>5</td>
</tr>
<tr>
<td>creatine phosphate CrP small molecule</td>
<td>6/14/07</td>
<td>20 mM</td>
<td>40</td>
</tr>
<tr>
<td>ATP +GTP</td>
<td>9/23/07</td>
<td>1 mM each</td>
<td>40</td>
</tr>
<tr>
<td>normal E. coli tRNA 2 hr.</td>
<td>9/22/09</td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>Factor mix (10/09) SLASH</td>
<td>40x</td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>EF-Tu (10/09) TU</td>
<td>Tim 12.18.10</td>
<td></td>
<td>60</td>
</tr>
<tr>
<td>Nat DOT AARS mix</td>
<td>31-Oct</td>
<td></td>
<td>32</td>
</tr>
<tr>
<td>MRS</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Ribosome R (TR/SDS/Tested1/21/11)</td>
<td>6uL/50uL rxn</td>
<td></td>
<td>210</td>
</tr>
<tr>
<td>35S Met</td>
<td>10 µM</td>
<td>0.2 µM</td>
<td>60.0</td>
</tr>
<tr>
<td>Master Mix</td>
<td></td>
<td></td>
<td>901.8</td>
</tr>
<tr>
<td>Total Volume (µL)</td>
<td></td>
<td></td>
<td>1000.0</td>
</tr>
</tbody>
</table>

---

**Purification of translation mixture using Oligo dT**

The following volumes/Oligo dT were used for a 500 µl translation reaction. First, the binding and elution buffers were prepared containing TCEP.

**Column preparation**

A 20 mL Bio-Rad column was rinsed with ddH₂O to which was added 50 mg of Oligo dT (stored at -20 °C desiccator) into the column and ddH₂O was added and mixed well to the column to
swell the cellulose, about 5 mL. The cap was removed and the water removed using gravity filtration in all steps.

**Binding of Oligo dT to mRNA fusion**

To the Bio-Rad column was added at least 4 mL of Oligo dT binding buffer (Table B-0-6) containing TCEP (to cover all of the cellulose) then filtered and this was repeated. If precipitation occurred, the translation reaction was thoroughly mixed (to suspend any protein that precipitated) and 5 µl was removed for scintillation counting. The rest of the translation reaction was transferred to the column and the translation tube was rinsed 3 times with 0.8 mL binding buffer and added to the column. Sufficient binding buffer was added to give a final volume of 4 mL and this mixture was gently mixed in a 4 °C cold room for 30 minutes. The binding buffer was then drained from the column. The mRNA fusion bound to the Oligo dT was then rinsed twice with 4 ml washing buffer (Table B-0-6).

**Table B-0-6 Buffer recipes for oligo dT purification**

<table>
<thead>
<tr>
<th>Oligo dT Binding Buffer:</th>
<th>Stock Concentration</th>
<th>Volume for 500ml reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 mM EDTA pH 8</td>
<td>500 mM</td>
<td>10 ml</td>
</tr>
<tr>
<td>20 mM Tris pH 7.8</td>
<td>1 M</td>
<td>10 ml</td>
</tr>
<tr>
<td>1 M NaCl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2 % triton x-100</td>
<td>10% v/v</td>
<td>10 ml</td>
</tr>
<tr>
<td>0.5 mM TCEP (Added Fresh)</td>
<td>0.5 M</td>
<td>1:1000 dilution</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Oligo dT Wash Buffer:</th>
<th>Stock Conc.</th>
<th>Volume for 500ml</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 mM Tris pH 7.8</td>
<td>1 M</td>
<td>10 ml</td>
</tr>
<tr>
<td>0.3 M NaCl</td>
<td></td>
<td>8.766 g</td>
</tr>
<tr>
<td>0.1 % triton x-100</td>
<td>10% v/v</td>
<td>5 ml</td>
</tr>
<tr>
<td>0.5 mM TCEP (Added Fresh)</td>
<td>0.5 M</td>
<td>1:1000 dilution</td>
</tr>
</tbody>
</table>

**Elution of mRNA fusion from Oligo dT:**

The mRNA fusion bound to the oligo dT on the Bio-Red column was eluted over 8 fractions using 200 µl water. From each fraction, 1 µl was set aside for scintillation counting in addition
to a sample of ddH₂O water for a background count. At this point, the 5 µl from the original mRNA fusion prior to purification, the 8 elutions and the background sample are measured for scintillation counts. All elution fractions were then combined and filtered through a centrifugal filter for one minute at 9000 rpm.

**Quantification of mRNA fusion yields**

An example of typical concentrations of mRNA fusion from the wash and elution phases of Round 6 are shown in Table B-0-7.
Table B-0-7 Typical elution results of mRNA fusion from Oligo dT

<table>
<thead>
<tr>
<th>Sample</th>
<th>cpm</th>
<th>cpm-Control</th>
<th>Tested V</th>
<th>Total V</th>
<th>cpm total</th>
<th>pmols total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background</td>
<td>22</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Flow Through</td>
<td>376576</td>
<td>376554</td>
<td>5</td>
<td>4000</td>
<td>301242960</td>
<td>6871.48</td>
</tr>
<tr>
<td>Wash 1</td>
<td>50726</td>
<td>50704</td>
<td>5</td>
<td>4000</td>
<td>40563200</td>
<td>925.26</td>
</tr>
<tr>
<td>Wash 2</td>
<td>2674</td>
<td>2652</td>
<td>5</td>
<td>4000</td>
<td>2121600</td>
<td>48.39</td>
</tr>
<tr>
<td>Wash 3</td>
<td>871</td>
<td>849</td>
<td>5</td>
<td>4000</td>
<td>679200</td>
<td>15.49</td>
</tr>
<tr>
<td>Elution 1</td>
<td>11964</td>
<td>11942</td>
<td>1</td>
<td>200</td>
<td>2388380</td>
<td>54.48</td>
</tr>
<tr>
<td>Elution 2</td>
<td>14186</td>
<td>14164</td>
<td>1</td>
<td>200</td>
<td>2832800</td>
<td>64.62</td>
</tr>
<tr>
<td>Elution 3</td>
<td>11272</td>
<td>11250</td>
<td>1</td>
<td>200</td>
<td>2250000</td>
<td>51.32</td>
</tr>
<tr>
<td>Elution 4</td>
<td>4520</td>
<td>4498</td>
<td>1</td>
<td>200</td>
<td>899600</td>
<td>20.52</td>
</tr>
<tr>
<td>Elution 5</td>
<td>3873</td>
<td>3851</td>
<td>1</td>
<td>200</td>
<td>770280</td>
<td>17.57</td>
</tr>
<tr>
<td>Elution 6</td>
<td>814</td>
<td>792</td>
<td>1</td>
<td>200</td>
<td>158400</td>
<td>3.61</td>
</tr>
<tr>
<td>Elution 7</td>
<td>341</td>
<td>319</td>
<td>1</td>
<td>200</td>
<td>63800</td>
<td>1.46</td>
</tr>
<tr>
<td>Elution 8</td>
<td>153</td>
<td>131</td>
<td>1</td>
<td>200</td>
<td>26200</td>
<td>0.60</td>
</tr>
<tr>
<td>Combined Filt</td>
<td>5461</td>
<td>5439</td>
<td>1</td>
<td>1600</td>
<td>8702096</td>
<td>198.50</td>
</tr>
</tbody>
</table>

**Ethanol precipitation of purified mRNA fusion with glycogen**

The filtered elutions were then ethanol precipitated with glycogen. To the combined filtered elutions was added 10% volume 3M KOAc pH 5.5 and 3 volumes 100% EtOH with glycogen (1:100 dilution of glycogen in proportion to sample volume). The mixture was mixed and cooled in the -20 °C freezer for 30 minutes followed by centrifugation at 12,000 rpm for 20 minutes at 4 °C. The pellet was visualized and the supernatant discarded in the appropriate radioactive waste container. The pellet was resuspended in 500 µl 70% EtOH and transferred
to 2 ml tube then centrifuged for 2 minutes at max speed. The supernatant was discarded and the pellet was again washed with 500ul 70% EtOH, centrifuged, and the supernatant discarded. The pellet was spun once again and any residual ethanol was removed and the pellet was dried for 10 minutes in a fume hood. The pellet was then re-dissolved in water such that the final concentration should be about was approximately 100 nM (0.1 pmol/µL).

**Reverse transcription (mRNA to cDNA)**

After EtOH precipitation of mRNA fusion, the pellet was dissolved in a volume of water suitable for RT (see Table B-0-8 below). The RT Oligo, dNTP and the mRNA fusion was added to a PCR tube and run with an RT program, Table B-0-8. After step 1 (5 min at 65 °C for annealing), the PCR tube was placed in ice for at least 1 minute, and the remainders of the reagents were added in the order listed on in the table. (DTT always before RNAse). The PCR tube was placed back into the thermocycler and the program run to completion for 15 min at 55 °C for elongation, then 15 min at 70 °C to inactivate the RT enzyme. In the RT reaction, the final concentrations of fusions were 0.04 pmol/µl, which is the optimal concentration.

**Table B-0-8 Reverse transcription protocol**

<table>
<thead>
<tr>
<th>RT Protocol</th>
<th>STOCK</th>
<th>FINAL</th>
<th>Sample (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mRNA Fusion</td>
<td>Variable</td>
<td>0.04 pmol/µL</td>
<td>Variable</td>
</tr>
<tr>
<td>RT Oligo</td>
<td>100 µM</td>
<td>0.5 µM</td>
<td>0.5</td>
</tr>
<tr>
<td>dNTP</td>
<td>10 mM</td>
<td>0.5 mM</td>
<td>5</td>
</tr>
<tr>
<td>10XRTbuf</td>
<td>10x</td>
<td>1x</td>
<td>10</td>
</tr>
<tr>
<td>MgCl2</td>
<td>25 mM</td>
<td>5 mM</td>
<td>20</td>
</tr>
<tr>
<td>DTT</td>
<td>0.1 M</td>
<td>0.01 M</td>
<td>10</td>
</tr>
<tr>
<td>RNAse OUT</td>
<td>40 u/µL</td>
<td>2 u/µL</td>
<td>5</td>
</tr>
<tr>
<td>Superscript III</td>
<td>200 u/µL</td>
<td>5 u/µL</td>
<td>2.5</td>
</tr>
<tr>
<td>Water</td>
<td></td>
<td></td>
<td>Variable</td>
</tr>
</tbody>
</table>
**Ni-NTA purification of RT reaction**

The C-terminus of the resulting peptide fusion each contain a His tag that enables Nickel-NTA affinity purification. Following the elution of the mRNA-peptide fusion the yields can then be calculated prior to selection. The general protocol for Ni-NTA purification is as follows. Fresh binding buffer was prepared by adding 1.05 µl of BME per 3 mL of stock solution (see Table B-0-9 for the binding buffer recipe). To a centrifuge spin filter was added 100 µL Nickel-NTA solution (mixed well) and spin filtered to remove ethanol storage solution from beads. Each 100 µL RT reaction was diluted 5 fold with denaturing binding buffer and added to the spin filtered Ni-NTA beads and bound for one hour whilst mixing at 4 °C. The mixture was then spin filtered via centrifugation at 5900 rpm for 1 minute, 10 µL for was removed for a scintillation count of the flow through. The resin was washed three times with 400 µL wash buffer (spin filtered via centrifugation each time) and 10 µL of each wash was set aside for scintillation counting. The purified mRNA fusion was eluted in six fractions with 50 µL elution buffer. At each elution, 5 minutes were allowed to pass between addition of the elution buffer and centrifugation. For each fraction 0.5 µL was removed for scintillation counting and the eluted fractions with high counts were combined.

**Ni-NTA buffers for purification of RT**

**Table B-0-9 Ni-NTA buffers**

<table>
<thead>
<tr>
<th>Denaturing Binding- Stock</th>
<th>Vol. for 50 mL Solution</th>
<th>Wash Buffer: Stock</th>
<th>Vol. for 50 mL Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH 8.0</td>
<td></td>
<td>pH 8.0</td>
<td></td>
</tr>
<tr>
<td>Stock</td>
<td></td>
<td>Stock</td>
<td></td>
</tr>
</tbody>
</table>

262
<table>
<thead>
<tr>
<th>100 mM NaH₂PO₄</th>
<th>1 M</th>
<th>5 mL</th>
<th>100 mM NaH₂PO₄</th>
<th>1 M</th>
<th>5 mL</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 mM Tris HCl</td>
<td>1 M</td>
<td>0.5 mL</td>
<td>300 mM NaCl</td>
<td>5 M</td>
<td>3 mL</td>
</tr>
<tr>
<td>6 M guanidinium hydrochloride solid</td>
<td>28.65 g</td>
<td></td>
<td>0.2% Triton X-100</td>
<td>10% v/v</td>
<td>1 mL</td>
</tr>
<tr>
<td>0.2% Triton X-100</td>
<td>10% v/v</td>
<td>1 mL</td>
<td>5 mM BME reagent</td>
<td>1 mL</td>
<td></td>
</tr>
<tr>
<td>5 mM BME reagent</td>
<td>1 mL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Native Elution with imidazole**

<table>
<thead>
<tr>
<th>pH 8.0</th>
<th>Stock</th>
<th>Volume for 50 mL Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 mM NaH₂PO₄</td>
<td>1 M</td>
<td>2.5 mL</td>
</tr>
<tr>
<td>300 mM NaCl</td>
<td>5 M</td>
<td>3 mL</td>
</tr>
<tr>
<td>250 mM imidazole</td>
<td>1 M</td>
<td>12.5 mL</td>
</tr>
<tr>
<td>0.2% Triton X-100</td>
<td>1 mL</td>
<td>1 mL</td>
</tr>
<tr>
<td>5 mM BME</td>
<td>10%</td>
<td>1 mL</td>
</tr>
</tbody>
</table>

**Selection Buffer**

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Final Conc.</th>
<th>Stock Conc.</th>
<th>Volume for 1L Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tris-HCl, pH 8</td>
<td>50 mM</td>
<td>1 M</td>
<td>50 mL</td>
</tr>
<tr>
<td>NaCl</td>
<td>150 mM</td>
<td>5 M</td>
<td>8.78 g</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>4 mM</td>
<td>1 M</td>
<td>4 mL</td>
</tr>
<tr>
<td>Triton X-100</td>
<td>0.25%</td>
<td>10%</td>
<td>25 mL</td>
</tr>
<tr>
<td>ddH₂O</td>
<td></td>
<td></td>
<td>to 1 L</td>
</tr>
</tbody>
</table>

**Selection**

The idea of the selection process was to expose the mRNA peptide fusion to Molecule 3-3, which acted as an electrophile against nucleophiles in the library. Subsequently, any reactive peptides that were reactive and subsequently biotinylated within the library were then isolated using streptavidin fused to magnetic beads. Once these peptides were isolated on the beads the cDNA of the mRNA peptide fusion were amplified by PCR.

**Preparation of Streptavidin magnetic particles/beads**

The streptavidin (SA) beads used were Roche Streptavidin Magnetic Particles, Catalog Number 11 641 778 001. SA beads were supplied in a suspension that contains 10 mg of particles per mL.
at pH 7.4 and stored at 2-8 °C but can be stored at cooler temperatures permitted they are not subjected to constant freezing/thawing. Binding capacity for small molecules is listed as greater than 1800 pmol per 1 mg of SA magnetic particles. Therefore 10 µl of streptavidin beads was sufficient to isolate 180 pmol of material per assay.

The beads were prepared for each study by using a sample size of 1X 10 µl of streptavidin beads in solution with 200 µl distilled deionized water that were exposed to the magnet for 1 minute and the solution pipetted away from the beads carefully. The washing protocol for the beads was as follows: First, 200 µl of additional water was added, the beads were removed from the magnet, mixed, then the magnet reapplied for one minute and the water carefully removed, this process was repeated three times in total with 600 µL of selection buffer (Figure B-0-4)

Figure B-0-4 Streptavidin magnetic particles

Magnetic particles for pre-clearance studies.
The purpose of this experiment is to remove any peptides in the mRNA fusion that react with the beads and not selectively with the electrophile. To the washed beads described above (without buffer) was added 500 µl mRNA fusion plus a 20 µl selection buffer rinse of the
centrifuge tube that was used to carry the fusion. The beads were incubated at room temperature in a tumbler for 30 minutes. After incubation the particles were washed three times with buffer. The beads were stored in buffer and set aside in the -20 °C freezer. The mRNA fusion was divided into 4 separated aliquots in 1.5 mL centrifuge tubes.

**Reaction with Molecule 3-3**

Three experiments were prepared such that the mRNA fusion was exposed to a high concentration and a low concentration of Molecule 3-3 and a third sample was prepared as a control without any exposure to the electrophile. In sample 1, one third of mRNA fusion was placed in a 1.7 mL centrifuge tube and 10µL of 1.0 mM Molecule 3-3 dissolved in water was added. In sample 2, one third of the RNA fusion was placed in a 1.7 mL centrifuge tube and 10 µL of 0.1 mM Molecule 3-3 dissolved in water was added. In sample 3, one third of the mRNA fusion was placed in a 1.7 mL centrifuge tube. The three experiments were placed in a 4 °C refrigerator overnight and then mixed in a tumbler for one hour.

**Removal of residual small molecules via dialysis**

Once the peptide mRNA fusion was reacted with Molecule 3-3, the unreacted electrophile needed to be removed as it would saturate the streptavidin binding sites on the magnetic particles and prevent complete capture of those mRNA fusion peptides that had been biotinylated, see Table B-0-10. Therefore, once the reaction between the peptide fusion and Molecule 3-3 was complete, the contents of each experiment were transferred into a 2,000 MWCO dialysis chamber (0.1-0.5 mL) and dialyzed into 1x selection buffer (1:100 sample to dialysis volume) in the cold room overnight. The buffer was then exchanged and the samples
were dialyzed for a second time. The samples were then removed from each dialysis chamber into separate 1.7 mL centrifuge tubes.

**Table B-0-10 Magnetic beads capture limitations required dialysis of unreacted Molecule 3-3**

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Concentration of Electrophile</th>
<th>Molecule 3-3</th>
<th>Reaction Volume</th>
<th>After dialysis</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>High</td>
<td>1 mmol/L</td>
<td>X 10 μL</td>
<td>0.01 mmol</td>
<td>1% of Molecule 3-3 Left to give 0.1 mol \n but 10 μL SA beads react with 180 pmol biotin, double dialysis required.</td>
</tr>
<tr>
<td>2</td>
<td>Low</td>
<td>0.01 mmol/L</td>
<td>X 10 μL</td>
<td>0.0001 mmol</td>
<td>1% of Molecule 3-3 Left to give 100 pmol, within limitations of beads for capture</td>
</tr>
<tr>
<td>3</td>
<td>Only have non-specific binding of mRNA fusion, .069 pmol within limitation of 10 μL beads</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**mRNA Fusion isolation using Streptavidin beads.**

10 μL of streptavidin beads were washed as described above to which was added followed by addition of the dialyzed mRNA fusion and incubation at room temperature in a tumbler for thirty minutes. After incubation the solution was separated from the particles and set aside for scintillation counting. The particles were washed three times with selection buffer and the captured cDNA of the mRNA fusion were amplified via PCR.

An example of the measurement of the extent of capture of biotinylated peptides from the mRNA fusion is provided in Chart B-0-1. The scintillation counts were measured using 200 μL samples from the fusion that had been separated from the beads, from each subsequent wash of the beads, and finally from the beads themselves.
Finally, the cDNA of the mRNA fusion peptides that were biotinylated and captured on the beads were amplified using PCR.

**PCR amplification of cDNA from selected mRNA fusion peptides**

**Small scale optimization**

As each sample contained a varying amount of captured peptides from mRNA fusion it was necessary to optimize PCR conditions. Optimization of parameters included volume of beads, optimization of primers, number of bead washes, and determination of the impact of the concentration of the Molecule 3-3 and number of PCR cycles. A preliminary 100 µl scale test was run. The thermocycler/PCR machine was preheated to 94°C before sample placement; starting in Cycle 14 an aliquot of 4 µL was taken from each sample and added to 4 µl of Orange G Buffer, these samples were run on a 2% agarose gel in 1XTBE with ethidium bromide. An
additional lane was run with 5 µl of the reference DNA ladder (Hyperladder V). The gel was run at 120V for 30-40 minutes.

**Large scale amplification**

Based on the optimized concentrations of beads and optimized number of cycles, the beads were amplified by scaling up the number of PCR reactions. For validation purposes 5 µL from the final products were set aside and run on a 2% Agarose gel to insure the success of the scale-up reactions. The samples were then combined and extracted following a phenol/chloroform extraction procedure.

**Purification of PCR product**

The samples were then combined and purified following a phenol/chloroform extraction. To the combined samples in a 15 mL centrifuge tube was added 100% reaction volume of phenol/chloroform/ isoamyl alcohol 25:24:1 solution. This mixture was thoroughly mixed and spun down at 4 °C at 4000rpm for 2 minutes. The upper layer (aqueous) was transferred to a new 15 ml centrifuge tube and a 100% reaction volume of ice cold chloroform was added. The mixture was again thoroughly mixed and spun down at 4°C at 4000rpm for 2 minutes and the aqueous layer transferred to a new centrifuge tube and ethanol precipitated. The resulting precipitate was dissolved into a final volume of 100 µl, and run through a NAP-5 column with water as the eluting buffer to de-salt the DNA. At this point the dsDNA underwent transcription and another round of in vitro selection took place.
Vita

Susan Daniela Selaya was born on May 31, 1979 in Sucre, Bolivia and is an American citizen. She graduated from North Hollywood High School from the Zoological/Biological Sciences magnet program in 1997. She received her Bachelor of Science in Chemical Education from University of California San Diego in La Jolla, California. She subsequently worked for 2 years at Arena Pharmaceuticals, Inc. in San Diego, California then completed a Master’s of Science in coursework in chemistry at University of California Los Angeles in Westwood, California 2006. She then worked for 3 years at AI Biotech formerly known as Commonwealth Biotechnologies, Inc. before completing her graduate studies at Virginia Commonwealth University.

Academic Background

1997-2001 University of California San Diego La Jolla, CA
  • B.S. Chemistry Education

2004-2006 University of California Los Angeles Westwood, CA
  • Research Advisor: Professor Miguel Garcia-Garibay
  • M.S.c. Organic Chemistry

2008-present Virginia Commonwealth University Richmond, VA
  • Research Advisor: Matthew Hartman, PhD
• Ph.D. Candidate Organic chemistry

Employment & Research Experience

Aug 2014-present    AI BioTech, Division of Bostwick Laboratories, Inc.  Richmond, VA

Senior Scientist

• Preparation/synthesis, isolation and characterization of small bioorganic molecules and polymers using related characterization techniques. (NMR/ GC-MS/ Q-TOF/ MALDI-TOF/ HPLC/ Biotage Isolera Chromatography, etc.)

Jan 2010-present    Virginia Commonwealth University  Richmond, VA

Graduate Student Researcher

• Peptide library produced by mRNA display to find a small nucleophilic pepticid sequence that covalently and specifically reacts with an electrophilic small molecule.
• Synthesis and development of electrophilic fluorescent small molecules designed as molecular probes to covalently label proteins.
• Functionalize of Y$_2$O$_3$ scintillation nanoparticles with an azide that can potentially be coupled to a photocaged drug.

May-Aug 2011    AI BioTech, Division of Bostwick Laboratories, Inc.  Richmond, VA

Scientist II

• Preparation/synthesis, isolation and characterization of small/ bioorganic molecules using related characterization techniques. (NMR/ GC-MS/ HPLC)
Aug 2006- July 2008   Commonwealth Biotechnologies, Inc. Richmond, VA

Laboratory Specialist, Senior

- Preparation/synthesis, isolation and characterization of small/ bioorganic molecules using related characterization techniques. (NMR/ GC-MS/ HPLC)
  - Peptide Synthesis including solution and solid phase and peptide synthesizer.
  - Preparation of biological agent analogues.
  - Large-scale preparation of protein analogues requiring high purity content.
- Prepared estimates and written proposals for incoming project opportunities.
- Prepared monthly presentations and formal reports to summarize progress of each assigned project.
- Methods development of analytical parameters for HPLC/ TLC/ and GC-MS qualifications.
  - Various modes of compliance were followed such as GLP/ non-GLP.
- Quantification, classification and characterization of various biological and biodefense agents.
  - Required background check at state and local level.

2004-2006   University of California, Los Angeles Westwood, CA

Graduate Student Researcher

- Developed and synthesized novel solid-state organic materials that mimic macro scale compasses and gyroscopes.
Completed coursework included spectroscopic methods of organic chemistry, organic synthesis, structure and mechanism in organic chemistry, and laboratory safety training.

2002-2004 Arena Pharmaceuticals, Inc. San Diego, CA

Research Associate

• Synthesis of new analogs for the 5HT2A receptor program in the medicinal chemistry department.
  ▪ Accessed multi gram quantities of key intermediates and high purity material for ex vivo evaluation
  ▪ Completed multi step organic syntheses discretely and/or in a parallel manner, investigating alternative synthetic routes to target molecules.
• Completion of large-scale synthesis of a lead compound for in vivo evaluation within the process and development department.
  ▪ Prepared monthly presentations and formal reports to summarize progress of work.

2000-2001 University of California, San Diego La Jolla, CA

Undergraduate Student Researcher

• Research Advisor: Professor Murray Goodman
• Synthesis of a penta-peptide opioid derivative using standard coupling reagents and a specific protecting group found to cleave amide bonds.
**Teaching Experience**

2008-2012 Virginia Commonwealth University Richmond, VA

**Teaching Assistant**

- Taught Organic Chemistry Laboratories: CHEZ301/ CHEZ302/ Honors 301 and 302

2012 (May-Aug) Richmond, VA

**Private Tutor**

- Tutored organic chemistry 20 hours/week for organic chemistry 301/302 and organic laboratories CHEZ301/302

2008-2010 Club Z Tutoring Richmond, VA

**Math/ Science Tutor**

- Tutored mathematics for various age groups including elementary and high school.
  - Pre-algebra/ trigonometry/ honors calculus/ physical science/ geology.

2004-2006 University of California, Los Angeles Los Angeles, CA

**Teaching Assistant**


**Mentoring Experience**
2009-2012 Virginia Commonwealth University Richmond, VA

- Sylvia Clay, VCU Undergraduate 2009
- Taja Stancil, Bennett College (NSF-REU) 2009 Summer
- Niyant Jain, VCU Undergraduate 2010
- Scott Dickens, VCU Undergraduate 2011-2012

**Technical and Specialized Skills**

- **Molecular Biology**: Isolation and purification of DNA/RNA, in vitro transcription, radiochemical usage and handling, protein expression and purification, PCR, bacterial transformation, plasmid insertion using PCR overlap extension.


- **Health & Safety** and Q.A. training.

**United States Security Clearance**

- **Industrial Secrets Act (ISA) Clearance**
  - ISA/Secret Level (Debiefed 2008, reactivated 2011, debriefed 2011)

**Professional/ Academic Honors and Awards**

2005 **Excellence in Teaching** Westwood, CA

- Department of Chemistry and Biochemistry at University of California, Los Angeles

2012 **Outstanding Teaching Assistant Award** Richmond, VA

- Department of Chemistry at Virginia Commonwealth University
2012-2013  **VCU Dept. of Chemistry Altria Fellow**  Richmond, VA

- Department of Chemistry at Virginia Commonwealth University

2013  **Billy Stump/Ray Ottenbrite Scholarship**  Richmond, VA

- Department of Chemistry at Virginia Commonwealth University

**Publications**

Teegarden, Bradley R.; Li, Hongmei; Jayakumar Honnappa; Strah-Pleynet, Sonja; Dosa, Peter, I.; Selaya, Susan D.; et al. 1-[3-(4-Bromo-2-methyl-2H-pyrazol-3-yl)-4-methoxyphenyl]-3-(2,4-difluorophenyl)urea (Nelotanserin) and Related 5-HT2A Inverse Agonists for the Treatment of Insomnia. *J. Med. Chem.*, 2010, 53 (5), pp. 1923–1936.

**Presentations**


Affiliations/Memberships

American Chemical Society (2002-present)

Chemistry Graduate Student Organization, VCU (2011-2013)

• Fundraising Chair and Event Organizer.