

Copyright

by

Katharine Louise Diehl

2015

The Dissertation Committee for Katharine Louise Diehl Certifies that this is the approved version of the following dissertation:

Development of Cross-Reactive Receptors Based on Serum Albumin

Committee:

Eric V. Anslyn, Supervisor

Jennifer S. Brodbelt

Brent L. Iverson

Sean M. Kerwin

Jonathan L. Sessler

Development of Cross-Reactive Receptors Based on Serum Albumin

by

Katharine Louise Diehl, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2015

Acknowledgements

First, I would like to thank Eric Anslyn for being an excellent adviser and mentor to me during my doctoral studies. I have truly appreciated his kindness, patience, and optimism over the years. I would also like to thank my undergraduate adviser and mentor, Marcey Waters, who served as an important female role model to me. She showed me what it is possible for a woman chemist to accomplish.

I would like to thank Anslyn group members past and present for being wonderful colleagues and friends. I could not have made it through without you! In particular, I want to thank past member Michelle Ivy for teaching me everything she knows about serum albumin and for being a model of patience and compassion to me. I am grateful to past members Leo Joyce and Reid Long for their friendship and for their mentorship to me during my early years in the group. I would also like to especially thank past member Alex Gade for being my best friend in the lab. I would like to thank the following current Anslyn group members for making every day in lab fun: Brette Chapin, Maggie Meadows, Hannah Jo, John Lin, Igor Kolesnichenko, Ram Edupuganti, Rogelio Escamilla, and James Bachman. I would like to give a special thanks to Brette for reading every word of this dissertation out of the kindness of her heart. Finally, I would like to thank all of my friends in the department for making these past five years so much fun!

I would like to thank my family for their love and support during my graduate studies. A big thanks to Adam- for putting up with me when I was stressed out (i.e. all of the time) and helping me with papers, presentations, and even my research, despite not being a chemist; my mom- for calling to check on me and sending me candy on every holiday; and my dad- for encouraging me to be a scientist in the first place.

Development of Cross-Reactive Receptors Based on Serum Albumin

Katharine Louise Diehl, Ph.D.

The University of Texas at Austin, 2015

Supervisor: Eric V. Anslyn

In recent years, differential sensing has become an increasingly popular approach to molecular recognition. Mimicking the mammalian senses of taste and smell, arrays of semi-selective sensors generate a fingerprint for each analyte. Pattern recognition algorithms allow these arrays to be used for discriminating analytes and for predicting the identity of unknowns. Arrays of cross-reactive receptors have found use in a variety of sensing applications, including the differentiation of biologically-relevant analytes (Chapter 1).

Serum albumin has previously been used in an array format for the discrimination of hydrophobic analytes such as terpenes, plasticizers, and fatty acids. This protein is a versatile cross-reactive receptor because of its ability to bind hydrophobic analytes with different affinities and in different modes. In Chapter 2, the use of serum albumin to pattern hydrophobic analytes was further expanded to include glycerides. Glycerides are challenging analytes because they are structurally similar to one another. Due to difficulties in identifying the regio- and stereochemistry of the unsaturated glycerides, a sample pretreatment consisting of olefin cross metathesis was used prior to array analysis. Using the array, twenty glycerides were discriminated, including stereo- and regioisomeric pairs. Further, glycerides in mixtures were quantitated.

Due to the success with using serum albumins as receptors for hydrophobic analytes, it was hypothesized that serum albumin could be used in a different way to develop receptors for other types of analytes. The serum albumin functions as a scaffold onto which species dynamically assemble through hydrophobic interactions. In Chapter 3, the use of fatty acid-appended recognition units was explored as a way of building such receptors. While these conjugates did not bind to the protein as well as expected, an alternative strategy was explored in Chapter 4 with squaraines and thiols. As a starting point toward this goal, a series of squaraines were synthesized, and their reactivity to thiols and affinity for serum albumin were investigated. Finally, in Chapter 5, the reversible covalent attachment of recognition units to serum albumin with thia-Michael chemistry was explored. A bifunctional conjugate acceptor was used to reversibly label proteins with thiols as well as to generate resin-bound dynamic combinatorial libraries.

Table of Contents

List of Tables	xvi
List of Figures	xvii
List of Schemes	xxvii
1.1 Introduction.....	1
1.1.1 Chemometrics	2
1.1.1.1 Principal Component Analysis.....	3
1.1.1.2 Linear Discriminant Analysis.....	4
1.1.1.3 Hierarchical Cluster Analysis.....	6
1.1.1.4 Artificial Neural Networks.....	7
1.1.2 Optical Detection in Chemical Sensing	8
1.1.2.1 Colorimetric Detection	9
1.1.2.2 Fluorescence Detection	10
1.2 Arrays for Small Biomolecules.....	11
1.2.1 Sugars.....	12
1.2.2 Amino Acids	17
1.2.3 Nucleotides	20
1.3 Arrays for Proteins.....	22
1.4 Arrays for Whole Cells	29
1.4.1 Prokaryotic cells.....	29
1.4.2 Eukaryotic cells.....	33
1.5 Summary and Outlook	37
1.6 Additional Information	39
1.7 References.....	39
Chapter 2: Differential Sensing For the Regio- and Stereoselective Identification of Glycerides: A Qualitative Fingerprinting and Quantitative Analysis Study	43
2.1 Glycerides	43
2.1.1 Importance of Glyceride Determination	44

2.1.2	Glyceride Determination Methods.....	44
2.2	An Array-Based Approach to Glyceride Detection.....	47
2.3	Serum albumin.....	50
2.3.1	General Properties.....	51
2.3.2	Ligand Binding.....	53
2.4	Previous Work Using Serum Albumin as a Cross-Reactive Receptor	53
2.5	Development of an Array for Glycerides.....	56
2.5.1	Indicator Uptake Experiments with Serum Albumin and Triglyceride	56
2.5.2	First Iteration of the Well Plate Array.....	58
2.6	Olefin Metathesis of Glycerides.....	60
2.6.1	Synthesis of Allyl Fluorescein.....	60
2.6.2	Optimizing the Metathesis Reaction with Monoerucin.....	61
2.6.3	Evaluating the Olefin Metathesis with All of the Unsaturated Glycerides in the Panel.....	64
2.6.4	Binding Studies of AF and 2.1 to Serum Albumin.....	66
2.6.5	Exploring FRET with Dansyl Amide and Allyl Fluorescein.....	67
2.6.6	Second Iteration of the Well-Plate Array.....	68
2.7	Array Reproducibility.....	71
2.8	Analysis of the Full Glyceride Panel.....	73
2.9	Prediction of Glyceride Structural Features.....	76
2.10	Adipocytes.....	81
2.10.1	Extracts from Human and Rat Adipocytes.....	83
2.10.2	Application of the Array to the Human Fat.....	84
2.10.3	Application of the Array to the Rat Fat.....	91
2.10.4	Mass Spectral and ¹ H-NMR Data from Fat Samples.....	96
2.11	Quantitation of Trilinolein (T5) in Triglyceride Mixtures Using Standard Addition.....	100
2.11.1	Introduction to standard addition and net analyte signal (NAS)	100
2.11.2	Applying SANAS to mixtures of glycerides.....	102

2.12	Conclusions.....	106
2.13	Future Work.....	107
2.14	Contributions.....	107
2.15	Experimental Methods.....	108
2.15.1	General Information.....	108
2.15.2	Synthesis.....	108
2.15.2.1	Synthesis of NBD-FA.....	108
2.15.2.2	Synthesis of allyl fluorescein.....	109
2.15.2.3	Synthesis of 1.....	110
2.15.3	Array Procedures.....	110
2.15.3.1	Glyceride Array (Part 1).....	110
2.15.3.2	Well Plate Metathesis Reaction.....	111
2.15.3.3	Glyceride Array with Metathesized Glycerides (Part 2)	111
2.15.4	Application of the Array to Adipocyte Fat Extracts.....	112
2.15.4.1	Preparation of the fat samples.....	112
2.15.4.2	Use of the fat in the array.....	113
2.15.4.3	¹ H-NMR Spectra of the Fat Samples.....	113
2.15.5	Quantitation of Trilinolein in Triglyceride Mixtures.....	115
2.15.5.1	Preparation of the Glyceride Mixtures.....	115
2.15.5.2	Standard Addition Procedure.....	116
2.16	Additional Figures.....	116
2.17	References.....	117
 Chapter 3: Development of Dynamic Supramolecular Receptors from Serum Albumin and Fatty Acid-Appended Peptides.....		
3.1	Introduction.....	123
3.2	Synthesis of Fatty Acid-Peptides.....	125
3.2.1	Attaching the Fatty Acid to the Resin.....	125
3.2.2	Synthesis of the Fatty Acid-Peptide Conjugates.....	126
3.3	Binding Studies of Fatty Acid-Peptides to Serum Albumin.....	127

3.3.1	Indicator Displacement and Uptake Titrations with NBD-FA	127
3.3.2	Indicator Displacement and Uptake Titrations with Other Fluorophores	133
3.3.2.1	Dansyl Amide.....	133
3.3.2.2	Dansyl Proline.....	134
3.3.2.3	1,8-Anilinonaphthalene Sulfonate.....	135
3.3.2.4	2-Anthracenecarboxylate	136
3.3.2.5	BSPOTPE.....	137
3.3.3	Intrinsic Fluorescence	137
3.4	Conclusions and Future Work	139
3.5	Experimental Methods	140
3.5.1	General.....	140
3.5.2	Synthesis	141
3.5.2.1	Fmoc-Protected 12-Aminododecanoic Acid.....	141
3.5.2.2	Loading Fmoc-Fatty Acid to Wang Resin	141
3.5.2.3	Peptide Synthesis.....	142
3.5.2.4	Acetylation of ACKGGRK-FA.....	143
3.5.2.5	Peptide Purification.....	143
3.5.3	Titrations.....	143
3.5.4	Additional Figures	144
3.6	References.....	145
Chapter 4: Investigation of Serum Albumin Binding and Thiol Addition to Water-Soluble Squaraines.....		148
4.1	Introduction.....	148
4.1.1	Synthesis, Structure, and Optical Properties of Squaraines	148
4.1.2	Aggregation Behavior of Squaraines in Solution	150
4.1.3	Reaction of Squaraines with Nucleophiles	150
4.1.4	Binding of Squaraines to Serum Albumin.....	153
4.1.5	Project Goals and Design.....	155
4.2	Synthesis of Squaraines	156

4.2.1	Indolenine-Based (SQ2)	156
4.2.2	Phloroglucinol-Based (SQ3 and SQ3Br)	157
4.2.3	Aniline-Based (SQ1, SQ4-9)	157
4.3	Optical Properties of the Aniline-Based Squaraines	161
4.4	Studies of Thiol Addition to Squaraines	163
4.4.1	SQ2	163
4.4.2	SQ3Br	165
4.4.3	SQ6	167
4.4.4	Thiol Addition to the Aniline-Derived Squaraine Series (SQ4, 5, 7, 8, and 9)	168
4.4.5	¹ H NMR of Thiol Addition to SQ5	172
4.4.6	Testing Reactivity of Squaraines to Amines	173
4.5	Studies of Squaraine Binding to Serum Albumin	174
4.5.1	SQ1	174
4.5.2	SQ3Br	176
4.5.3	SQ4	176
4.5.4	SQ5	178
4.5.5	SQ6	186
4.5.6	SQ7	187
4.5.7	SQ8	189
4.5.8	SQ9	190
4.5.9	Summary	194
4.6	Thiol Addition to Squaraines in the Presence of Serum Albumin	194
4.7	Conclusions	201
4.8	Future Work	202
4.9	Contributions	202
4.10	Experimental Methods	203
4.10.1	General	203
4.10.2	Synthesis	203
4.10.2.1	SQ1	204
4.10.2.2	SQ2	205

4.10.2.3	SQ3Br.....	206
4.10.2.4	SQ4.....	207
4.10.2.5	SQ5.....	208
4.10.2.6	SQ6.....	209
4.10.2.7	SQ7.....	209
4.10.2.8	SQ8.....	210
4.10.2.9	SQ9.....	211
4.10.2.10	c-BSA.....	213
4.10.3	Titration.....	213
4.11	Additional Figures and Experiments.....	214
4.11.1	¹ H NMR Monitoring of Thiol Addition.....	214
4.11.2	Titration of Thiol into BSA/SQ.....	217
4.11.3	Titration of G5 PAMAM Dendrimer into SQ5.....	219
4.12	References.....	221
Chapter 5: Development of Dynamic Receptors from Serum Albumin and Thia-		
Michael Addition Chemistry.....		
5.1	Introduction.....	225
5.1.1	Dynamic Combinatorial Chemistry.....	225
5.1.1.1	Reversible Covalent Bonding.....	226
5.1.2	Protein Modification and Conjugation.....	227
5.1.3	Characterization of Modified Proteins and Protein Conjugates.....	228
5.2	Partial Reduction of Serum Albumin.....	231
5.3	Thiolation of Serum Albumin.....	232
5.3.1	<i>N</i> -acetylhomocysteine thiolactone (AHTL).....	232
5.3.1.1	Capping with Iodoacetamide.....	232
5.3.1.1	Capping with Methyl Vinyl Ketone.....	235
5.3.2	<i>N</i> -succinimidyl <i>S</i> -acetylthioacetate (SATA).....	237
5.3.3	<i>N</i> -succinimidyl <i>S</i> -tritylthioacetate (STTA).....	242
5.3.4	Synthesis of a Conjugate Acceptor.....	243
5.3.5	Summary.....	244

5.4	Labelling of Serum Albumin by a Conjugate Acceptor (EvaCA).....	244
5.4.1	Reactions of EvaCA with Lysine and Cysteine.....	245
5.4.1.1	Lysine and Cysteine	245
5.4.1.2	Lysine	246
5.4.1.2	Cysteine.....	247
5.4.2	Synthesis of an EvaCA Derivative	248
5.4.2.1	Synthesis of PyrCA via 2-(Bromomethyl)pyridine....	249
5.4.2.2	Synthesis of BenzCA via Benzylbromide.....	249
5.4.2.3	Synthesis of QuinCA via 2-(Bromomethyl)quinolone.....	250
5.4.2.4	Synthesis of PyrCA via Pyridin-2-ylmethanethiol.....	250
5.4.3	Synthesis of a Model BSA Peptide and Cysteine-Containing Peptides.....	251
5.4.4	Reaction of the Model Peptide with a Cysteine-Containing Peptide	252
5.4.5	Reaction of EvaCA with BSA	253
5.4.6	Alternative Thiols	255
5.4.6.1	2-Mercaptoethanol	255
5.4.6.2	2-Azidoethanethiol.....	256
5.4.6.3	2-{2-[2-(2-Mercaptoethoxy)ethoxy]ethoxy}ethanol..	257
5.4.7	Reaction of EvaCA with Myoglobin	258
5.4.8	Removal of Amine and Thiol from EvaCA with Dithiothreitol.....	259
5.4.8.1	Structure of DTT-EvaCA Adduct and Mechanism of Formation.....	260
5.4.8.2	Removal of Label from MyoEvaMEEEE	260
5.5	Triamine Compound as a Scaffold	262
5.5.1	Reaction of the Triamine Compound with EvaCA and Cysteine ...	262
5.5.2	Reversibility of the Thiol Addition to PWEva	263
5.5.3	Resin-Bound Triamine Compound.....	264
5.5.4	Dynamic Combinatorial Libraries	265

5.6	Conclusions.....	268
5.7	Future Work.....	268
5.8	Contributions.....	269
5.9	Experimental Methods.....	269
5.9.1	General.....	269
5.9.2	Synthesis	270
5.9.2.1	N-succinimidyl S-tritylthioacetate (STTA).....	270
5.9.2.2	1-(4-azidophenyl)prop-2-en-1-one (attempted)	271
5.9.2.3	PyrCA via 2-(bromomethyl)pyridine (attempted).....	271
5.9.2.4	EvaCA	272
5.9.2.5	BenzCA (attempted).....	272
5.9.2.6	QuinCA (attempted).....	272
5.9.2.7	PyrCA via 2-pyridinemethanethiol (attempted).....	273
5.9.2.8	2-Azidoethanethiol	274
5.9.2.9	Triamine compound	275
5.9.3	Peptide Synthesis	276
5.9.3.1	Model Peptide	276
5.9.3.2	Purification.....	277
5.9.4	Small Molecule Test Reactions	278
5.9.4.1	LysEvaCys	278
5.9.4.2	LysEvaCys + DTT	279
5.9.4.3	CysEvaCys	279
5.9.4.4	LysEvaBME	281
5.9.4.5	MPEvaVVKLKC	281
5.9.4.6	MPEvaVVKLKC + DTT	282
5.9.4.7	LysEvaLys and LysEvaLys + DTT.....	283
5.9.4.8	LysEvaAz and CuAAC reaction	285
5.9.4.9	LysEvaMEEEE and LysEvaMEEEE + DTT	286
5.9.4.10	MPEvaMEEEE.....	287
5.9.4.11	DTT-Eva Adduct.....	288

5.9.4.12	EDT-Eva Adduct.....	288
5.9.4.13	PE ₃ C ₃	289
5.9.4.14	PE ₃ C ₃ + DTT	290
5.9.4.15	PE ₃ C ₃ + 3-Mercaptopropionic Acid.....	291
5.9.4.16	Coupling of Triamine to Tentagel S Br (Resin-P) ...	292
5.9.4.17	Kaiser Test.....	292
5.9.4.18	Resin-PE ₂	293
5.9.4.19	Resin-PE ₂ C ₂	293
5.9.4.20	DCL.....	294
5.9.5	Protein Modification Reactions	295
5.9.5.1	BSA Reduction by DTT	295
5.9.5.2	BSA Thiolation with AHTL.....	295
5.9.5.3	BSA Thiolation with SATA and Deprotection	296
5.9.5.4	Reaction of t-BSA with Methyl Vinyl Ketone.....	297
5.9.5.5	BSAEvaVVKLKC	297
5.9.5.6	BSAEvaVVKLKC + DTT	298
5.9.5.7	BSAEvaCys.....	298
5.9.5.8	MyoEvaMEEEE.....	298
5.9.6	Protein Characterization.....	300
5.9.6.1	BSA Concentration Measurement.....	300
5.9.6.2	Ellman test.....	300
5.9.6.3	Quantitative Ninhydrin Test.....	301
5.9.6.4	Mass Spectrometric Analysis	301
5.9.7	Titrations.....	302
5.10	References.....	302
	Bibliography	306
	Vita	320

List of Tables

Table 1.1	<i>Structures of the 30 dual-ligand nanoparticles.</i>	36
Table 2.1	<i>Reaction conditions for the metathesis reaction and AF and M1.</i>	63
Table 2.2	<i>Summary of the olefin metathesis reaction for the unsaturated glycerides and AF.</i>	66
Table 2.3	<i>Summary of prediction accuracy for the six combinations of the three repetitions in Figure 2.11.</i>	73
Table 2.4	<i>Human fat samples.</i>	84
Table 2.5	<i>Rat fat samples including time courses.</i>	84
Table 2.6	<i>Results for T5 concentration in triglyceride mixtures from SANAS.</i>	104
Table 2.7	<i>Plate read parameters for part 1 of the glyceride array</i>	111
Table 2.8	<i>Plate read parameters for part 2 of the glyceride array</i>	112
Table 2.9	<i>Composition of glyceride mixtures by weight</i>	115
Table 4.1	<i>Summary of optical properties of aniline-based squaraines.</i>	162
Table 4.2	<i>Fitting of binding curves of NAC added to SQ to 1:1 regime.</i>	172
Table 4.3	<i>Relevant σ values for the squaraine series.</i>	172
Table 5.1	<i>Summary of the lysine modifications of the p-t-BSA replicate samples.</i>	240
Table 5.2	<i>Summary of peptide solutions and DCL setup.</i>	295

List of Figures

Figure 1.1	<i>General outline of array sensing process.</i>	2
Figure 1.2	<i>Principal component analysis.</i>	4
Figure 1.3	<i>Linear discriminant analysis.</i>	6
Figure 1.4	<i>Agglomerative clustering.</i>	7
Figure 1.5	<i>The layers of ANN.</i>	8
Figure 1.6	<i>Discrimination of saccharides by boronic acids/pH indicators by Chang et al.</i>	13
Figure 1.7	<i>Color difference maps of the array by Suslick and coworkers.</i>	14
Figure 1.8	<i>Proposed indicator displacement mechanism for the BBV receptor.</i>	15
Figure 1.9	<i>LDA plot for the discrimination of twelve sugars by the BBV- based system of Singaram et al.</i>	16
Figure 1.10	<i>Discrimination of sugars with a boronic acid-peptide receptor array.</i>	17
Figure 1.11	<i>Discrimination of natural amino acids by IDA-based array by Severin et al.</i>	19
Figure 1.12	<i>Discrimination of nucleotides by an array of MIPs by Mirsky et al.</i>	21
Figure 1.13	<i>Differentiation of nucleotides by the gold nanoparticle/fluorescent dye array of Prins et al.</i>	22
Figure 1.14	<i>Protein fingerprints from porphyrin receptor array by Hamilton et al.</i>	24
Figure 1.15	<i>Differentiation of 17 proteins by a fluorescent polymer array by Rotello et al.</i>	25

Figure 1.16 <i>Discrimination of MAPK isoforms by SOX-peptide receptors by Anslyn et al.</i>	28
Figure 1.17 <i>PCA plot from response of array by Suslick et al. to ten bacterial strains.</i>	30
Figure 1.18 <i>Color response of liposome solution in the presence of LPS.</i>	32
Figure 1.19 <i>Schematic of fluorescent polymer/gold nanoparticle array for bacterial cell surfaces.</i>	33
Figure 1.20 <i>Discrimination of cancer cell lines with the array from Rotello et al.</i>	34
Figure 1.21 <i>Discrimination of cell lines with array by Yan et al.</i>	37
Figure 2.1 <i>MS⁵ experiment with silver cationization to discriminate regioisomers.</i>	46
Figure 2.2 <i>Crystal structure of bovine serum albumin.</i>	52
Figure 2.3 <i>Crystal structure of human serum albumin.</i>	52
Figure 2.4 <i>Schematic of the sensing ensemble of serum albumin, PRODAN, and a hydrophobic additive in the presence of terpene.</i>	54
Figure 2.5 <i>Discrimination of contaminated explosive mixtures by the serum albumin array.</i>	55
Figure 2.6 <i>NBD-FA uptake.</i>	58
Figure 2.7 <i>Discrimination of eight glycerides by the six-receptor array.</i>	59
Figure 2.8 <i>Evidence of FRET between DNSA and 2.1.</i>	67
Figure 2.9 <i>Discrimination of eight glycerides by the fourteen-receptor array.</i> ...	70
Figure 2.10 <i>Loading plot corresponding to the LDA in Figure 2.9.</i>	70

Figure 2.11 <i>LDA plots of data collected from 96-well plates in three independent repetitions.</i>	72
Figure 2.12 <i>Discrimination of twenty glycerides by the array.</i>	74
Figure 2.13 <i>Discrimination of twenty glycerides using PCA.</i>	76
Figure 2.14 <i>Independent repetition of array discrimination of twenty glycerides.</i>	77
Figure 2.15 <i>LDA plots of training sets with glyceride classes designated as structural features.</i>	78
Figure 2.16 <i>LDA plots of the training sets with glyceride classes designated as saturated, unsaturated, and polyunsaturated.</i>	79
Figure 2.17 <i>LDA plots of the training sets of unsaturated glycerides with classes designated as shown in the key.</i>	80
Figure 2.18 <i>LDA plots of the training sets of saturated glycerides with analyte classes designated as shown in the key.</i>	81
Figure 2.19 <i>Lipolysis in an adipocyte under basal and stimulated conditions.</i> ...	82
Figure 2.20 <i>LDA plot of array data from analysis of HA1-HA8.</i>	85
Figure 2.21 <i>PCA plot of array data from analysis of HA1-HA8.</i>	86
Figure 2.22 <i>LDA plot of array data from analysis of HA3-HA8.</i>	87
Figure 2.23 <i>PCA plot of array data from analysis of HA3-HA8.</i>	88
Figure 2.24 <i>LDA plot of array data from analysis of HA9-HA14.</i>	89
Figure 2.25 <i>PCA plot of array data from analysis of HA9-HA14.</i>	90
Figure 2.26 <i>LDA plot of array data from analysis of rat samples 1, 4, and 16.</i>	91
Figure 2.27 <i>PCA plot of array data from analysis of rat samples 1, 4, and 16.</i>	92

Figure 2.28	<i>LDA plot of array data from analysis of all of the rat samples.</i>93
Figure 2.29	<i>PCA plot of array data from analysis of all of the rat samples.</i>93
Figure 2.30	<i>LDA plot of array data from the repetition of the analysis of the rat samples.</i>95
Figure 2.31	<i>PCA plot of array data from the repetition of the analysis of the rat samples.</i>95
Figure 2.32	<i>Chromatograms of metathesis reactions of human fat samples.</i>97
Figure 2.33	<i>Chromatograms of metathesis reactions of rat fat samples.</i>98
Figure 2.34	<i>SANAS plot generated from the standard addition of mixture A.</i>	...104
Figure 2.35	<i>SANAS plot generated from the standard addition of mixture B.</i>	...105
Figure 2.36	<i>SANAS plot generated from the standard addition of mixture C.</i>	...105
Figure 2.37	<i>SANAS plot generated from the standard addition of mixture D.</i>	...106
Figure 2.38	<i>¹H-NMR spectra of control (top), insulin-treated (middle), and isoproterenol (bottom) rat fat samples in CDCl₃.</i>115
Figure 2.39	<i>Excitation spectra for AF and 2.1.</i>117
Figure 2.40	<i>Titration of DNSA into AF/BSA.</i>117
Figure 3.1	<i>Cartoon representation of the fatty acid-peptides assembled on the serum albumin scaffold.</i>124
Figure 3.2	<i>Displacement of NBD-FA from BSA by KAGLK-FA.</i>	128
Figure 3.3	<i>Indicator uptake titrations with NBD-FA, BSA, and KAGLK-FA.</i>129
Figure 3.4	<i>Indicator uptake titrations with NBD-FA, BSA, and KGGRK-FA.</i>130
Figure 3.5	<i>Indicator uptake titrations with NBD-FA, BSA, palmitic acid (PA), and monopalmitin (MP).</i>131

Figure 3.6	<i>Indicator uptake titrations with NBD-FA, BSA, and 12-aminododecanoic acid.</i>	132
Figure 3.7	<i>Indicator uptake titrations with DNSA, BSA, ACKGGRK-FA, and KAGLK-FA.</i>	133
Figure 3.8	<i>Displacement of DNSA from BSA by KAGLK-FA.</i>	134
Figure 3.9	<i>Displacement of DP from BSA by ACKGGRK-FA.</i>	135
Figure 3.10	<i>Indicator uptake titrations with ANS, BSA, and ACKGGRK-FA.</i>	136
Figure 3.11	<i>Displacement experiment of AC from BSA by KAGLK-FA.</i>	136
Figure 3.12	<i>Indicator uptake titrations with BSPOTPE, BSA, and ACKGGRK-FA.</i>	137
Figure 3.13	<i>Titration of palmitic acid into BSA monitoring intrinsic fluorescence of BSA.</i>	138
Figure 3.14	<i>Titration of ACKGGRK-FA into BSA monitoring intrinsic fluorescence of BSA.</i>	138
Figure 3.15	<i>Titration of BSA into NBD-FA.</i>	144
Figure 3.16	<i>Titration of BSA into AC.</i>	144
Figure 3.17	<i>Titration of BSA into ANS.</i>	145
Figure 3.18	<i>Titration of BSA into BSPOTPE.</i>	145
Figure 4.1	<i>Simple schematic of H- and J-type aggregates.</i>	150
Figure 4.2	<i>Solutions of squaraine and ten equivalents of amino acids.</i>	151
Figure 4.3	<i>Addition of BSA to phloroglucinol-derived squaraine of Ramaiah et al.</i>	154
Figure 4.4	<i>Addition of BSA to benzoindolenine-derived squaraine of Belfield et al.</i>	155

Figure 4.5	<i>Cartoon representation of the squaraines and peptides assembled on the serum albumin scaffold.</i>	156
Figure 4.6	<i>Absorbance spectra of SQ4, 5, 7, 8, and 9.</i>	162
Figure 4.7	<i>Addition of thiol to SQ2.</i>	164
Figure 4.8	<i>Addition of selenol to SQ2.</i>	165
Figure 4.9	<i>Absorbance monitoring of thiol addition to SQ3Br.</i>	166
Figure 4.10	<i>Addition of thiol to SQ3Br with fluorescence monitoring.</i>	166
Figure 4.11	<i>Fluorescence monitoring of SQ3Br alone over time.</i>	167
Figure 4.12	<i>Absorbance monitoring of thiol addition to SQ6.</i>	167
Figure 4.13	<i>Absorbance titration of thiol into SQ4.</i>	168
Figure 4.14	<i>Absorbance titration of thiol into SQ5.</i>	169
Figure 4.15	<i>Absorbance titration of thiol into SQ7.</i>	169
Figure 4.16	<i>Absorbance titration of thiol into SQ8.</i>	170
Figure 4.17	<i>Absorbance titration of thiol into SQ9.</i>	170
Figure 4.18	<i>Isotherms for thiol addition to SQ4, 5, 7, 8, and 9 overlaid.</i>	171
Figure 4.19	<i>Addition of an amine to SQ7.</i>	174
Figure 4.20	<i>Fluorescence titration of BSA into SQ1.</i>	175
Figure 4.21	<i>Fluorescence titration of BSA into SQ3Br.</i>	176
Figure 4.22	<i>Absorbance titration of BSA into SQ4.</i>	177
Figure 4.23	<i>Fluorescence titration of BSA into SQ4.</i>	177
Figure 4.24	<i>Absorbance titration of HSA into SQ4.</i>	178
Figure 4.25	<i>Absorbance titration of BSA into SQ5.</i>	179
Figure 4.26	<i>Fluorescence titration of BSA into SQ5.</i>	180
Figure 4.27	<i>Fluorescence titration of BSA and c-BSA into SQ5.</i>	181
Figure 4.28	<i>Job plot of BSA and SQ5.</i>	182

Figure 4.29	<i>Fluorescence titrations of SQ5 into BSA/DNSA and BSA/DP.</i>	183
Figure 4.30	<i>Absorbance titration of DNSA into BSA/SQ5.</i>	184
Figure 4.31	<i>Fluorescence titration of DNSA into BSA/SQ5.</i>	185
Figure 4.32	<i>Absorbance titration of DP into BSA/SQ5.</i>	186
Figure 4.33	<i>Fluorescence titration of DP into BSA/SQ5.</i>	186
Figure 4.34	<i>Fluorescence titration of BSA into SQ6.</i>	187
Figure 4.35	<i>Absorbance titration of BSA into SQ7.</i>	188
Figure 4.36	<i>Fluorescence titration of BSA into SQ7.</i>	188
Figure 4.37	<i>Absorbance titration of BSA into SQ8.</i>	189
Figure 4.38	<i>Fluorescence titration of BSA into SQ8.</i>	189
Figure 4.39	<i>Absorbance titration of BSA into SQ9.</i>	190
Figure 4.40	<i>Fluorescence titration of BSA into SQ9.</i>	191
Figure 4.41	<i>Absorbance titration of DNSA into BSA/SQ9.</i>	192
Figure 4.42	<i>Fluorescence titration of DNSA into BSA/SQ9.</i>	192
Figure 4.43	<i>Absorbance titration of DP into BSA/SQ9.</i>	193
Figure 4.44	<i>Fluorescence titration of DP into BSA/SQ9.</i>	193
Figure 4.45	<i>Titration of SQ5, SQ5/NAC, SQ5/VVCLKC, and NAC into BSA/DP.</i>	195
Figure 4.46	<i>Titration of SQ5, SQ5/NAC, SQ5/VVCLKC, and NAC into BSA/DNSA.</i>	196
Figure 4.47	<i>Isotherms for thiol addition to all of the SQ/BSA combinations overlaid (monomer).</i>	197
Figure 4.48	<i>Isotherms for thiol addition to all of the SQ/BSA combinations overlaid (dimer).</i>	197
Figure 4.49	<i>Isotherms for thiol addition to SQ4 and SQ4/BSA.</i>	198

Figure 4.50	<i>Isotherms for thiol addition to SQ5 and SQ5/BSA.</i>	199
Figure 4.51	<i>Isotherms for thiol addition to SQ7 and SQ7/BSA.</i>	199
Figure 4.52	<i>Isotherms for thiol addition to SQ8 and SQ8/BSA.</i>	200
Figure 4.53	<i>Isotherms for thiol addition to SQ9 and SQ9/BSA.</i>	200
Figure 4.54	<i>¹H NMR of SQ5 in D₂O. SQ5 800 μM</i>	214
Figure 4.55	<i>¹H NMR of SQ5 and thiol in D₂O taken immediately after preparation.</i>	215
Figure 4.56	<i>¹HNMR of SQ5 and thiol in D₂O taken 24 hr after the sample was prepared.</i>	216
Figure 4.57	<i>Titration of thiol into SQ4/BSA.</i>	217
Figure 4.58	<i>Titration of thiol into SQ5/BSA.</i>	217
Figure 4.59	<i>Titration of thiol into SQ7/BSA.</i>	218
Figure 4.60	<i>Titration of thiol into SQ8/BSA.</i>	218
Figure 4.61	<i>Titration of thiol into SQ9/BSA.</i>	219
Figure 4.62	<i>Fluorescence titration of G5 dendrimer into SQ5.</i>	220
Figure 4.63	<i>Displacement of carboxyfluorescein from G5 dendrimer by SQ5, SQ5/NAC, and SQ/VVCLKC.</i>	221
Figure 5.1	<i>Illustration of a small dynamic combinatorial library (DCL) and the effect of a template on the free energy of library species.</i>	226
Figure 5.2	<i>Bottom-up mass spectrometric analysis of proteins.</i>	230
Figure 5.3	<i>Collision-induced dissociation by helium gas.</i>	230
Figure 5.4	<i>Collision-induced dissociation yields b- and y-ions.</i>	231
Figure 5.5	<i>Titration of BSA and c-t-BSA into NBD-FA.</i>	234
Figure 5.6	<i>Titration of BSA and c-t-BSA into CF.</i>	234
Figure 5.7	<i>Titration of BSA and c-t-BSA into AC.</i>	235

Figure 5.8	<i>Titration of BSA and k-t-BSA into ANS.</i>	236
Figure 5.9	<i>Titration of BSA and p-t-BSA into NBD-FA.</i>	238
Figure 5.10	<i>Titration of BSA and k-t-BSA into NBD-FA.</i>	240
Figure 5.11	<i>Titration of BSA and k-t-BSA into ANS.</i>	241
Figure 5.12	<i>Titration of BSA and k-t-BSA into AC.</i>	241
Figure 5.13	<i>Crystal structure of native BSA with the modified lysines in BSAEva and BSAEvaVVKLKC highlighted.</i>	255
Figure 5.14	<i>Crystal structure of native myoglobin with the modified lysines highlighted.</i>	259
Figure 5.15	<i>ESI spectrum of MyoEvaMEEEE after DTT treatment.</i>	261
Figure 5.16	<i>ESI spectrum of MyoEvaMEEEE.</i>	261
Figure 5.17	<i>Separation of the peptide mixture.</i>	266
Figure 5.18	<i>Separation of the untemplated library.</i>	267
Figure 5.19	<i>Separation of the library templated against BSA.</i>	267
Figure 5.20	<i>LC-MS analysis of LysEva reaction after 3 hr.</i>	278
Figure 5.21	<i>LC-MS analysis of LysEva reaction after 72 hr.</i>	279
Figure 5.22	<i>LC-MS analysis of LysEvaCys reaction after 8 days.</i>	279
Figure 5.23	<i>LC-MS analysis of CysEvaCys after 18 hrs.</i>	280
Figure 5.24	<i>LC-MS analysis of LysEvaCys after 3 days.</i>	280
Figure 5.25	<i>LC-MS analysis of LysEvaBME after 5 hr.</i>	281
Figure 5.26	<i>LC-MS analysis of MPEvaVVKLKC after 2 weeks.</i>	282
Figure 5.27	<i>MS analysis of MPEvaVVKLKC + DTT after 3 days.</i>	283
Figure 5.28	<i>LC-MS analysis of LysEvaLys after 3 days.</i>	284
Figure 5.29	<i>LC-MS analysis of LysEvaLys + DTT after 24 hr.</i>	284
Figure 5.30	<i>LC-MS analysis of LysEvaAz after 18 hr.</i>	285

Figure 5.31 <i>LC-MS analysis of Click reaction with LysEvaAz and 3-butyn-1-ol after 24 hr.</i>	286
Figure 5.32 <i>LC-MS analysis of LysEvaMEEEE reaction after 18 hr.</i>	286
Figure 5.33 <i>LC-MS analysis of LysEvaMEEEE + DTT after 18 hr.</i>	287
Figure 5.34 <i>MS analysis of MPEvaMEEEE after 20 hr.</i>	288
Figure 5.35 <i>LC-MS analysis of PE₃ reaction with NAC after 48 hr.</i>	289
Figure 5.36 <i>LC-MS analysis of PE₃ reaction with NAC after 4 days.</i>	290
Figure 5.37 <i>MS analysis of PE₃C₃ reaction with DTT after 3 days.</i>	291
Figure 5.38 <i>MS analysis of reaction of PE₃C₃ with 3-mercaptopropionic acid.</i>	292
Figure 5.39 <i>LC-MS analysis of the flow-through from the DTT reaction with resin-PE₂C₂.</i>	293
Figure 5.40 <i>LC-MS analysis of the flow-through from the DTT reaction with resin-PE₂C₂.</i>	294
Figure 5.41 <i>LC-MS analysis of flow-through from reaction of MyoEvaMEEEE and DTT.</i>	299

List of Schemes

Scheme 1.1	<i>Boronic acid-diol binding.</i>	12
Scheme 1.2	<i>General structure of pentapeptides from combinatorial library.</i>	17
Scheme 1.3	<i>Rhodium species and dyes.</i>	18
Scheme 1.4	<i>Structure of porphyrin scaffold.</i>	23
Scheme 1.5	<i>Structures of the PPEs used by Rotello et al.</i>	25
Scheme 1.6	<i>Boronic acid-functionalized peptides.</i>	26
Scheme 1.7	<i>SOX-peptides with the kinase docking sites indicated.</i>	28
Scheme 2.1	<i>Structure of glycerol, a fatty acid, and a triglyceride.</i>	43
Scheme 2.2	<i>Triglyceride structures</i>	48
Scheme 2.3	<i>Diglyceride structures</i>	49
Scheme 2.4	<i>Monoglyceride structures</i>	50
Scheme 2.5	<i>Structures of the fluorophores screened for use in the array.</i>	57
Scheme 2.6	<i>Synthesis of NBD-FA.</i>	57
Scheme 2.7	<i>Synthesis of allyl fluorescein (AF).</i>	61
Scheme 2.8	<i>Olefin metathesis of monoerucin.</i>	62
Scheme 2.9	<i>Metathesis products detected by mass spectrometry from reaction of AF with each of the unsaturated glycerides.</i>	65
Scheme 2.10	<i>Isoproterenol.</i>	83
Scheme 3.1	<i>FMOC protection of 12-aminododecanoic acid.</i>	125
Scheme 3.2	<i>Loading of FMOC-protected 12-aminododecanoic acid to Wang resin</i>	125
Scheme 3.3	<i>Structures of the fatty acid-peptides.</i>	126

Scheme 3.4	<i>Structures of the fluorescent probes used in this work.</i>	127
Scheme 4.1	<i>Examples of squaraines derived from different electron-rich aromatic and heterocyclic molecules.</i>	149
Scheme 4.2	<i>Resonance of the squaraine zwitterion</i>	149
Scheme 4.3	<i>Thiol addition to a squaraine</i>	151
Scheme 4.4	<i>Structures of the squaraines studied in their binding to serum albumin.</i>	153
Scheme 4.5	<i>Synthesis of SQ2</i>	157
Scheme 4.6	<i>Synthesis of SQ3 and SQ3Br</i>	157
Scheme 4.7	<i>Synthesis of SQ1</i>	158
Scheme 4.8	<i>Synthesis of SQ6</i>	159
Scheme 4.9	<i>Synthesis of SQ4, 5, 7, 8, and 9</i>	160
Scheme 5.1	<i>Thia-Michael reaction.</i>	227
Scheme 5.2	<i>Ellman's test for thiol using DTNB reagent.</i>	228
Scheme 5.3	<i>Ninhydrin test for primary amines.</i>	228
Scheme 5.4	<i>Reduction of disulfides by dithiothreitol (DTT).</i>	231
Scheme 5.5	<i>Thiolation of BSA by AHTL and silver nitrate.</i>	232
Scheme 5.6	<i>Alkylation of thiol by iodoacetamide.</i>	233
Scheme 5.7	<i>Alkylation of thiol by methyl vinyl ketone (MVK).</i>	235
Scheme 5.8	<i>Cleavage of endoproteinase Glu-C.</i>	237
Scheme 5.9	<i>Thiolation of BSA by SATA and hydroxylamine.</i>	237
Scheme 5.10	<i>Synthesis of STTA.</i>	242
Scheme 5.11	<i>Thiolation of BSA by STTA and subsequent deprotection and reaction with MVK.</i>	243
Scheme 5.12	<i>Attempted synthesis of the conjugate acceptor.</i>	244

Scheme 5.13	<i>Synthesis of EvaCA.</i>	245
Scheme 5.14	<i>Reaction of EvaCA with α-Boc-lysine to form LysEva.</i>	245
Scheme 5.15	<i>Reaction of LysEva with N-acetylcysteine (NAC) to form LysEvaCys.</i>	246
Scheme 5.16	<i>Reaction of EvaCA with excess α-Boc-lysine to form LysEvaLys. .</i>	247
Scheme 5.17	<i>Reaction of EvaCa with NAC to form CysEvaCys and then with α- Boc-lysine to form LysEvaCys.</i>	248
Scheme 5.18	<i>Structure of PyrCA (left) and pyridin-2-ylmethanethiol coordinated to copper(II) (right).</i>	248
Scheme 5.19	<i>Synthesis of PyrCA via 2-(bromomethyl)pyridine.</i>	249
Scheme 5.20	<i>Synthesis of BenzCA via benzylbromide.</i>	250
Scheme 5.21	<i>Synthesis of QuinCA via 2-(bromomethyl)quinolone.</i>	250
Scheme 5.22	<i>Synthesis of 2-pyridinemethanethiol.</i>	251
Scheme 5.23	<i>Synthesis of PyrCA via 2-pyridinemethanethiol.</i>	251
Scheme 5.24	<i>Synthesis of MPEvaVVKLKC.</i>	253
Scheme 5.25	<i>Synthesis of BSAEva.</i>	254
Scheme 5.26	<i>Synthesis of BSAEvaVVKLKC.</i>	254
Scheme 5.27	<i>Synthesis of LysEvaBME.</i>	256
Scheme 5.28	<i>Synthesis of 2-azidoethanethiol.</i>	256
Scheme 5.29	<i>Synthesis of LysEvaAz.</i>	257
Scheme 5.30	<i>Click reaction of 3-butyn-1-ol with LysEvaAz.</i>	257
Scheme 5.31	<i>Synthesis of LysEvaMEEEE.</i>	258
Scheme 5.32	<i>Synthesis of MPEvaMEEEE.</i>	258
Scheme 5.33	<i>Removal of amine and thiol from EvaCA by DTT.</i>	260

Scheme 5.34	<i>Synthesis of 1,3,5-benzenetrimethanamine.</i>	262
Scheme 5.35	<i>Synthesis of PE₃.</i>	263
Scheme 5.36	<i>Synthesis of PE₃C₃.</i>	263
Scheme 5.37	<i>Thiol exchange on PE₃C₃ with 2-mercaptopropionic acid.</i>	264
Scheme 5.38	<i>Immobilization of the triamine on a Tentagel S Br resin and subsequent reaction with EvaCA.</i>	265

Chapter 1: Array Sensing Using Optical Methods for Chemical Analysis¹

1.1 INTRODUCTION

Nearly all biological systems employ a lock-and-key approach to molecular recognition in the form of antibodies and enzymes.¹ Nature's high specificity approach has been used successfully for synthetic receptor design in some cases, but it has several drawbacks that limit its applicability to many sensing challenges. Firstly, the rational design of receptors is impractical for analytes that have not been fully characterized, such as many large biomolecules. Furthermore, even if an analyte's structure is known and a highly specific receptor can be designed, the synthetic work required to actually make the receptor can be an overwhelming task in itself. Finally, analyzing complex mixtures of analytes using a lock-and-key approach requires the design and synthesis of receptors for each component in the mixture, which is often an incredibly time-consuming undertaking.

For these reasons, an increasingly popular approach that also mimics biology is differential sensing. Mammalian olfaction and gustation employ cross-reactive receptors that interact differentially with odorants and tastants.^{2,3,4} Instead of identifying an odorant or tastant molecule by its strong affinity for one particular receptor, recognition is achieved by the composite response of the odorant or tastant to the entire array of semi-selective receptors in the nose or on the tongue.⁵ The result is a characteristic pattern – or fingerprint– for that odor or taste that the brain can perceive and store in an organism's memory.⁶ An important advantage of this approach to molecule identification is that the system is limited by the number of different patterns possible for an array of receptors rather than by the number of different receptors. Hence, the different receptors need not be designed to be highly specific for any one analyte. Additionally, this array approach

¹ Parts of this chapter were adapted from Diehl, K. L.; Anslyn, E. V. *Chem. Soc. Rev.* **2013**, *42*, 8596–8611. Katharine Diehl wrote the review, and Eric Anslyn edited the review.

allows the discrimination of analytes or analyte mixtures that have not been exhaustively characterized.

In the laboratory, the differential sensing approach described above is accomplished using sets of cross-reactive receptors and pattern recognition algorithms to process the large amount of data generated from this method (**Figure 1.1**). Both vapors and solutions can be exposed to the sensor array by either an active flow system that samples an analyte at a particular point in time (e.g. headspace sampling) or a static system that allows for equilibration between the receptors and analyte (e.g. microtiter plates).⁷ Whatever method is used to introduce the analyte to the sensor array, a measurable signal is generated upon the interaction of an analyte with the receptors. Many different kinds of signals can be employed, but most can be classified as electrical, gravimetric, or optical.¹

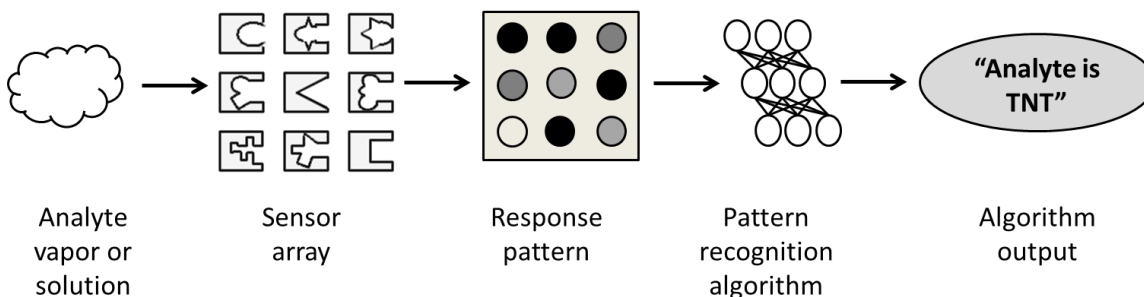


Figure 1.1 *General outline of array sensing process.* (Reproduced from Ref. 42: © Royal Society of Chemistry, 2013)

1.1.1 Chemometrics

Regardless of the type of signal output, array approaches produce a large amount of data across multiple axes. This type of data is usually not interpretable by visual inspection of the data set or by using basic calibrations like a simple linear regression. Therefore, more advanced chemometric methods are routinely used to reduce the dimensionality of the data and present it in graphical form for visual interpretation. There

are many statistical analysis tools available to process the data, but four common methods are principal component analysis (PCA), linear discriminant analysis (LDA), hierarchical cluster analysis (HCA), and artificial neural networks (ANN).

1.1.1.1 Principal Component Analysis

PCA is a cluster analysis method that reduces the dimensionality of a data set by decomposing the data into eigenvectors and eigenvalues.^{8, 9} The magnitudes of the eigenvalues represent the variance in the data. The variance can then be displayed graphically with principal component (PC) axes with the first principal component axis expressing the most variance in the data (**Figure 1.2**). The subsequent PCs are, by definition, orthogonal to one another. This orthogonality is particularly important when large numbers of input variables are used that may be correlated with one another.¹⁰ Requiring the PCs to be orthogonal to one another prevents redundant information in the data set from biasing the visualization of the data in the new space.

PCs are often referred to as latent variables because they represent relevant factors that may not be apparent in the original data space. For example, multiple chemical receptors in the array may respond to a similar property of the analytes, such as degree of hydrophobicity or charge state. Sometimes the PCs can be assigned to physical properties. However, depending on the array, it is not always possible to clearly connect each latent variable to one particular chemical property, nor is it necessary to do so in order to effect an accurate classification of an analyte.

When a visual inspection of a PCA plot detects not only close clustering between data points that represent repetitions of the same analyte class but also good separation between data points that represent different analyte classes, one can conclude that discrimination of the analytes has been successful. Loading values for each receptor

represent the contribution of that receptor to each PC axis. These can be used to evaluate the importance of the different receptors in an assay. Since PCA considers only the variance in the data matrix and not any *a priori* information about groupings or classes, this method is often referred to as unsupervised. Thus, in addition to dimensionality reduction and data visualization, PCA can be employed as an exploratory analysis method to try to identify latent variables as described above or to discover groupings that are not apparent in a data set.

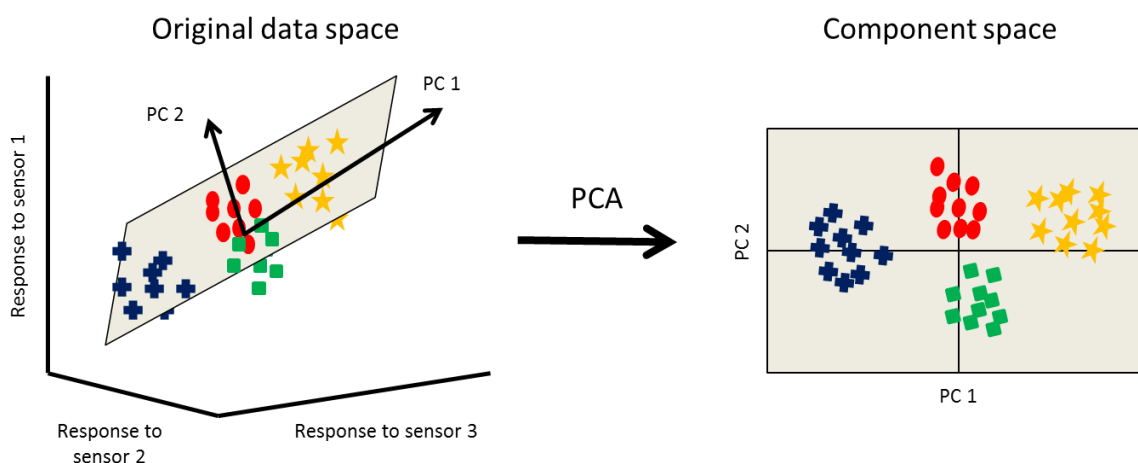


Figure 1.2 *Principal component analysis.* PCA displays the variance in the data set with PC 1 representing the largest variance. Subsequent PCs are orthogonal to one another. (Reproduced from Ref. 42: © Royal Society of Chemistry, 2013)

1.1.1.2 Linear Discriminant Analysis

LDA is a classification method typically used to assign an unknown analyte to its appropriate class.⁸ Both the data from the array (i.e. the optical responses from all of the receptors) and the analyte classes are used as inputs, making LDA a supervised method. From these inputs, discriminant functions, which are linear combinations of these input variables, are calculated with the goals of maximizing separation between classes and

minimizing separation within classes (**Figure 1.3**). The resultant variable space is described by new factor axes that are orthogonal to one another. These factor axes are similar in principle to the PCs in PCA and can also be thought of as latent variables.

As with PCA, a successful LDA plot shows good separation between different classes. The quality of the model built by LDA can be tested in several ways. Within a single data set, a cross validation can be employed that iteratively removes random segments of the data set, recalculates a model from the remaining data, and then uses the new model to assign the left-out data points to classes. When each data point is individually excluded, it is called a leave-one-out cross validation. The output from the cross validation is the accuracy of the assignments obtained through this process, which can be used as a metric for the stability and the quality of fit of the model. For prediction between two data sets, a training set is provided to the algorithm that includes the responses of all of the analytes of interest to the array in order to teach the algorithm. Since the classes (i.e. analyte identities) in the training set are known, an analyte of unknown identity can then be assigned to one of the classes based on the similarity of its response to the responses of the analytes in the training set. In this way, the predictive power of an array can be determined by the frequency of correct versus incorrect classifications.

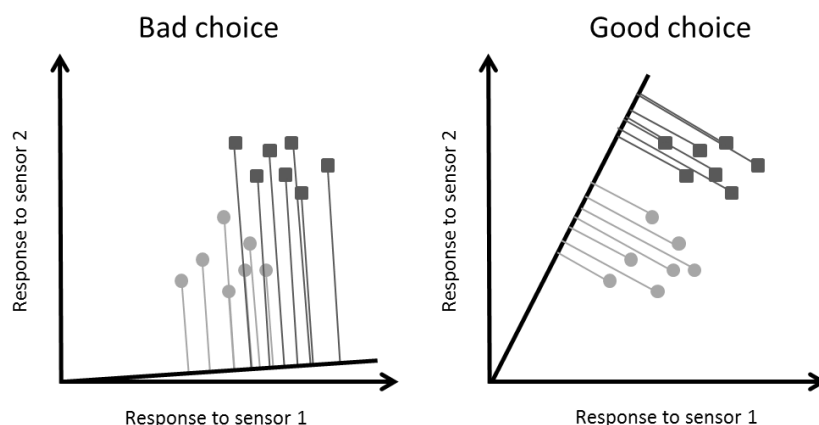


Figure 1.3 *Linear discriminant analysis.* In LDA, the algorithm finds the axis that best separates the given classes. The axis chosen in the example on the right separates the classes better than the axis chosen on the left. (Reproduced from Ref. 42: © Royal Society of Chemistry, 2013)

1.1.1.3 Hierarchical Cluster Analysis

HCA is a method that clusters data points based on the relative distances between all pairs of data points in the data set.^{8,11} This end can be achieved either by an agglomerative (“bottom up”) approach in which the algorithm begins with all of the data in separate clusters and works to merge clusters based on their similarity (**Figure 1.4**) or by a divisive (“top down”) approach in which the algorithm begins with all of the data in a single cluster and works to separate the clusters until each data point is in its own cluster. The Euclidean distance is typically used as the distance metric between data points. The Euclidean distance metric calculates the distance between two data points with N dimensions (N usually being the number of different sensor responses). The graphical result is a dendrogram. When using this method with data from an array, it is up to the researcher to make determinations about the meaningfulness of the clustering, since the algorithm will always eventually place all of the data into either one cluster in the case of

an agglomerative method or into as many clusters as there are data points in the case of a divisive method.

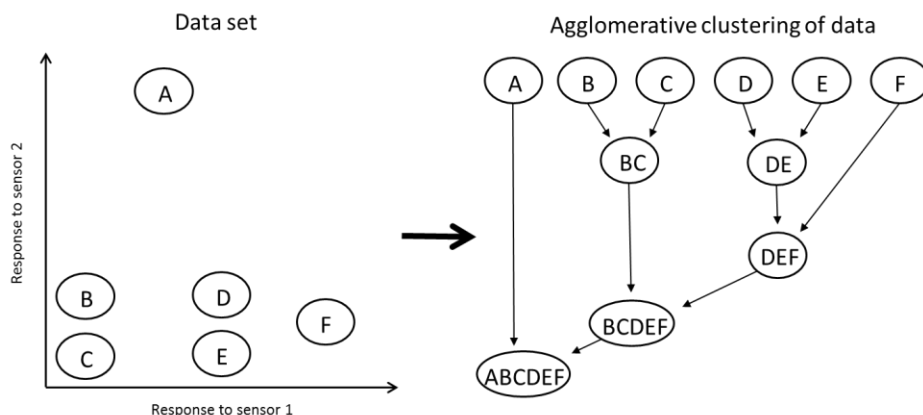


Figure 1.4 *Agglomerative clustering.* Illustration of how agglomerative clustering uses the Euclidean distance between data points as the basis for grouping the data points. (Reproduced from Ref. 42: © Royal Society of Chemistry, 2013)

1.1.1.4 Artificial Neural Networks

An ANN is a system of layer functions that can be “trained” to give a desired response pattern.^{8, 12} It consists of layers including the input, “hidden,” and output layers. The input layer is usually the data from the array (e.g. fluorescence counts or RGB color values), while the output layer is the desired analysis of the system (e.g. analytes). The “hidden” layers are user determined and can be adjusted to suit the system. The training process adjusts the hidden layers in order to maximize the number of times the desired output is achieved from the given inputs (**Figure 1.5**). Then, the system can be used to predict the identity of an analyte given the data set for that unknown. The hidden layers need not be linear functions, so this method is useful when trends in the data are expected to be non-linear.¹⁰ However, ANNs can easily overfit data and generate an unstable and overly narrow model.

A related method to ANNs that has found increasing popularity in chemometrics is a support-vector machine. Like an ANN, this method can also model non-linear trends in data. It does so by performing a non-linear transformation of the data into a higher dimensional space where it is possible to determine the best linear separators of the classes in the data.

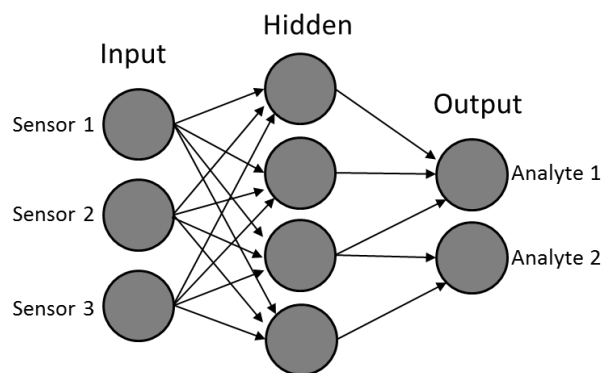


Figure 1.5 *The layers of ANN.* The inputs are the response data from the array, while the outputs are the different analyte identities. The hidden layers are adjusted during training to maximize the accuracy of the analyte identification. (Reproduced from Ref. 42: © Royal Society of Chemistry, 2013)

1.1.2 Optical Detection in Chemical Sensing

Another major component of a chemosensor array is the mechanism by which the interaction of the receptor and the analyte generates a measurable signal. Each receptor and analyte interact with one another through covalent, non-covalent, and/or reversible-covalent interactions. This recognition event must be transduced by some means, such as spectroscopic, electrochemical, or calorimetric methods. The signal transduction might be inherent to the recognition event between the two molecules, such as a chemical shift change in a $^1\text{H-NMR}$ spectrum upon binding, or it might be an indirect consequence of the specific receptor/analyte interaction, such as the displacement of a fluorescent dye from

the receptor by the analyte.¹³ The detection method should be sufficiently selective that it can be attributed to the binding event.

Optical methods in general are desirable for many sensing applications because of the potential for high sensitivity, good selectivity, rapidity of analysis, portability of instrumentation, and overall cost-effectiveness.¹⁴ Spectroscopic detection is possible with vapor-, solution-, and solid-phase sensing systems. Optical sensing assays are most commonly colorimetric or fluorescence-based.

1.1.2.1 Colorimetric Detection

Colorimetric sensors detect the change in color of a pigment upon interaction with an analyte vapor or solution. The dyes most commonly used are solvatochromic, Lewis base sensitive, pH sensitive, or redox reactive. These chromophores exhibit a change in their absorption of light of ultraviolet or visible wavelengths. This color change can sometimes be detected visually for rapid qualitative analysis.

For quantitative detection in the solution phase, a spectrophotometer is used to measure changes in the absorbance of the chromophore at one or more wavelengths compared to a blank sample. For the array format, the solutions are usually placed in microtiter plates rather than in cuvettes to improve the efficiency of the data collection process.

With printed arrays, quantitative colorimetric detection can be accomplished by reflectance detection.¹⁵ In this method, a desktop scanner or digital camera is used to capture the light reflected off the surface of the colored sensor regions on the printed array. The red, green, blue (RGB) values of the images of the spots are determined by standard digital image analysis software. This detection method is particularly desirable because

cameras and scanners are widely available and familiar technologies outside the context of scientific instrumentation.

1.1.2.2 Fluorescence Detection

Measuring fluorescence is perhaps the more popular optical detection method for array sensing. Most commonly, fluorescence quenching of an intrinsically fluorescent molecule in the presence of an analyte is measured. Quenching arises either from a static process (e.g. the quenching of a bound complex) or a dynamic process (e.g. bimolecular quenching of the excited state) and is described by the Stern-Volmer equation.¹⁶ In fluorescence measurements, the output is light emitted, which yields a better signal-to-noise ratio than absorbance measurements, where the output is the ratio of light absorption. The detection limits of fluorescence methods are usually one to three orders of magnitude lower than absorption methods.¹⁷ However, in the case of an analyte quenching the fluorescence of the sensor, a subtraction from the large background signal of the fluorescent molecule in the absence of analyte can translate to low sensitivity, especially if the effect of quenching is small. In array methods, this large background is generally less problematic. Usually the comparison being made is between the fluorescence responses of different sensors to an analyte and not to the emission of the sensor alone.

Nevertheless, a turn-on fluorescence approach is an increasingly popular signaling method that involves the use of a fluorescent molecule that increases the intensity of its emission in the presence of an analyte. Because the signal measured is relative to a dark background (i.e. the sensor alone exhibits relatively low to no emission), this method has good sensitivity.¹⁸ Changes in fluorescence lifetime, emission wavelength, and spectral shape can also be used as sensor responses.

Fluorescence detection can have some drawbacks, including photodegradation of the fluorophores, slow response times, and a lack of selectivity for the analyte of interest. However, in the case of successful differential arrays, this lack of specificity of fluorescence response is overcome by employing multiple different receptors in combination with pattern extraction techniques (LDA, PCA, etc.).

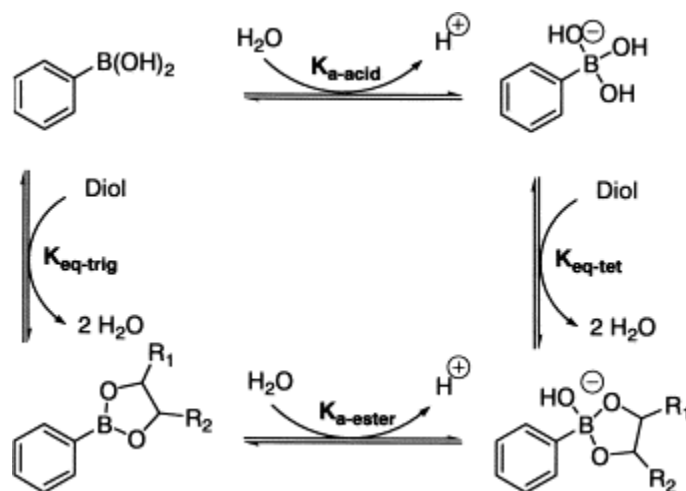
For solution-phase assays, microtiter plates with 96 or more wells per plate used in combination with a spectrophotometric plate reader are used for high-throughput signal quantitation. Plate readers consist of a series of optical components that provide excitation light of a wavelength or wavelengths to each sample well and detect emission light of a wavelength or wavelengths from each sample well. For vapor-phase assays, the sensors are immobilized on a surface like a glass slide or the distal ends of optical fibers. In optical fibers, the excitation light travels up to the immobilized sensors on the distal end containing the sensors from the other end of the fiber, and the emitted light from the sensors is sent back through the fiber to a detector.¹⁹ This convenient setup allows for easier remote sensing of analytes because the sensors need not be placed in close proximity to the excitation source and emission detector, as is necessary for interrogating sensors on glass slides.

1.2 ARRAYS FOR SMALL BIOMOLECULES

Cross-reactive arrays have been designed that target a wide range of small molecules. Biologically relevant small molecules such as sugars, amino acids, and nucleotides have been of particular interest as analytes. This section details some examples of systems that were developed to target such molecules.

1.2.1 Sugars

Since sugars contain hydroxyl groups as the only functional handles, the detection of carbohydrates by chemosensors in aqueous solution has been a challenging task. Differential sensing of sugars has primarily employed boronic acid-based receptors, as the following examples illustrate. Boronic acids reversibly form cyclic boronate esters with diols, including the diols found in carbohydrates (**Scheme 1.1**).^{20,21}



Scheme 1.1 *Boronic acid-diol binding.* Relationship between phenylboronic acid and its boronate ester with a diol. (Adapted from Ref. 20: © Elsevier, 2002)

For example, Chang and coworkers used an array consisting of boric acid and phenylboronic acid mixed with various pH indicators to discriminate 23 mono-, di-, and trisaccharides in aqueous solution.²² The sugars have different affinities for boric acid and phenylboronic acid, and thus modulate the pH of the solution differently based on these affinities (**Scheme 1.1**). The pH indicators report on this pH change by measurable changes in their absorption. PCA of the data from their array showed good differentiation of the sugars (**Figure 1.6**). However, nearly all of this discrimination is displayed along

F1 (95%), indicating that the receptors interact similarly to one another with the sugars. In a low-dimensionality case such as this one, an array format is likely unnecessary because all of the receptors are more or less providing the same information about the analyte. Nevertheless, this work demonstrates how commercially available components can be readily combined to generate arrays with good discriminatory power.

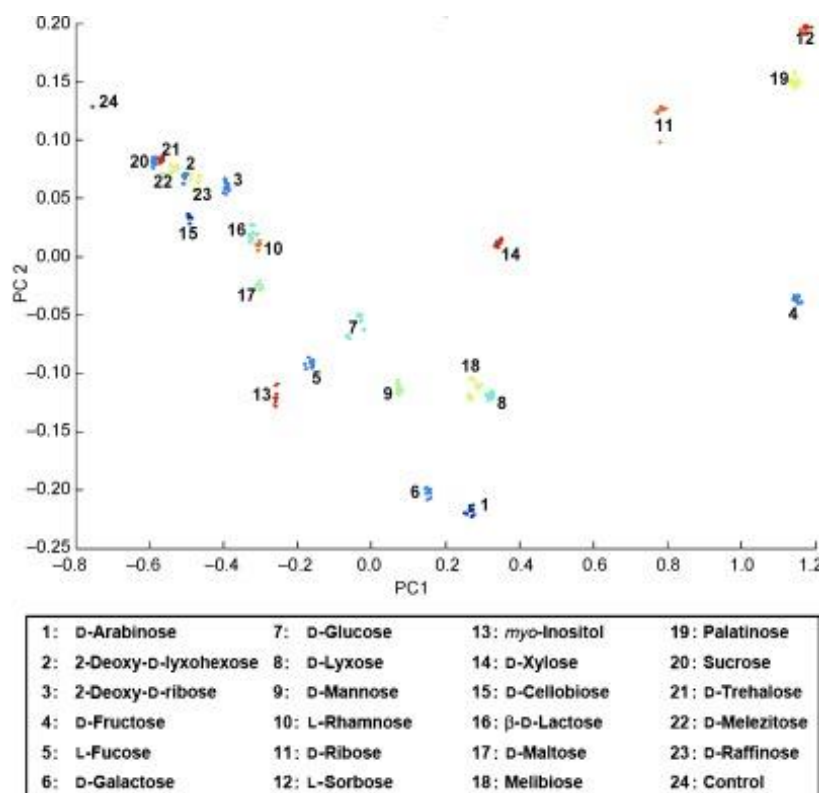


Figure 1.6 *Discrimination of saccharides by boronic acids/pH indicators by Chang et al.* PCA plot showing PC1 (95.4%) and PC2 (3.4%) for the identification of 23 carbohydrates based on six combinations of boric acid and phenylboronic acid each with three pH indicators. (Reproduced from Ref. 22: © Wiley-VCH, 2006)

Suslick and coworkers developed a similar array consisting of 3-nitrophenylboronic acid and several pH indicators printed as sol gel formulations to

distinguish a panel of fifteen sugars and sweeteners.²³ The printed arrays were exposed to the solutions of sugar and boronic acid for five minutes to allow for equilibration, resulting in visible color changes of the pH indicators (**Figure 1.7**). The authors asserted that the modulation of the responses by the dyes was due not just to pH changes, but also to Lewis acid-base and hydrogen bonding interactions of the sugars and sugar-boronic acid adducts with the dyes. Their assertion that multiple different types of interactions were occurring between the saccharides and receptors was evidenced by the LDA classification of the sugars requiring F1-F6 to achieve complete accuracy.

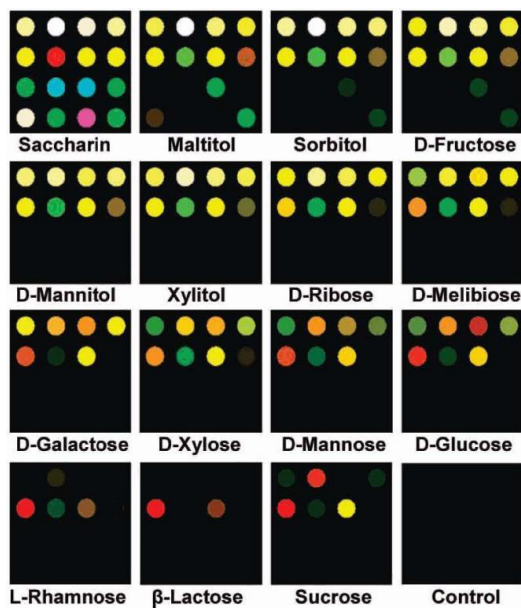


Figure 1.7 *Color difference maps of the array by Suslick and coworkers.* The sugars were discriminated by the responses of the pH indicators in the presence of 3-nitrophenylboronic acid. (Reproduced from Ref. 23: © American Chemical Society, 2008)

In order to improve the discriminatory power of boronic acids for carbohydrates, researchers have appended other functionalities to boronic acids. For instance, Singaram and coworkers attached boronic acids to cationic benzyl viologens (BBVs) in order to

incorporate fluorescent indicator displacement signaling in their array.²⁴ The cationic BBV interacts with an anionic fluorescent indicator. Upon formation of the negatively-charged boronate ester with a sugar, the affinity of the anionic indicator for the BBV diminishes and turns on its fluorescence (**Figure 1.8**). Six different BBV receptors and an anionic fluorophore were exposed to each of twelve saccharides in aqueous solution, and the fluorescence data from each solution was collected to generate an LDA plot (**Figure 1.9**). The sugars were successfully differentiated at a concentration of 2 mM, which is an order of magnitude lower than the concentrations used by Chang and Suslick in their systems. The lower detection limit could be attributed, at least in part, to the use of fluorescence rather than colorimetric detection.

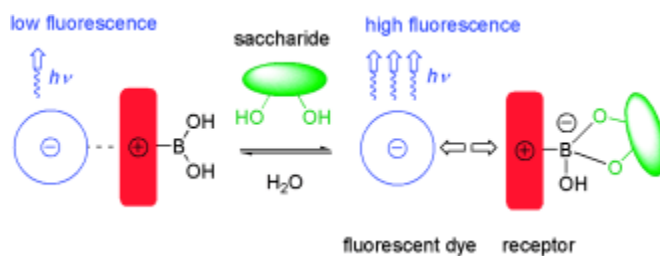


Figure 1.8 *Proposed indicator displacement mechanism for the BBV receptor.*
(Reproduced from Ref. 24: © Wiley-VCH, 2007)

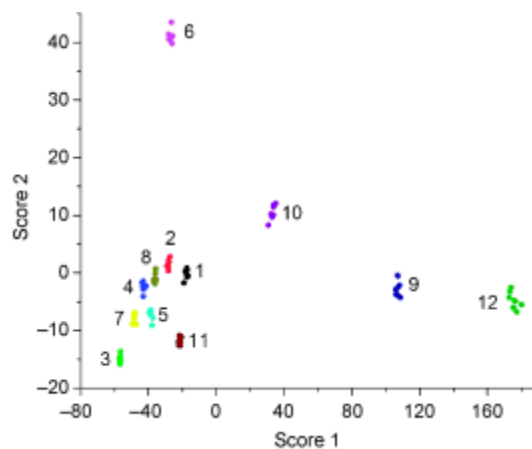
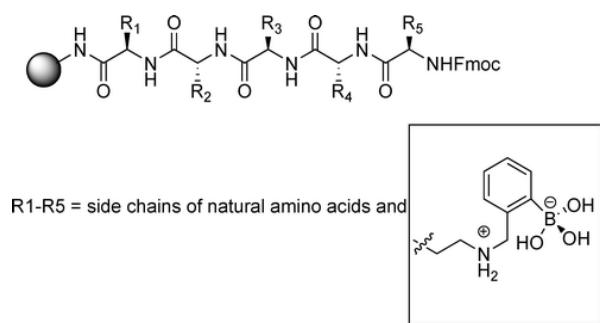


Figure 1.9 LDA plot for the discrimination of twelve sugars by the BBV- based system of Singaram *et al.* Key: 1) D-ribose, 2) D-arabinose, 3) L-rhamnose, 4) D-xylose, 5) D-lyxose, 6) D-glucose, 7) D-mannose, 8) D-galactose, 9) D-fructose, 10) L-sorbose, 11) melibiose, and 12) lactulose. (Reproduced from Ref. 24: © Wiley-VCH, 2007)

Similarly, Anslyn and coworkers appended boronic acids to resin-bound peptides to build immobilized sugar receptors that generate a signal through an indicator displacement event.²⁵ In addition to the boronate ester formation, the peptides themselves interact with the carbohydrates through hydrogen bonding and hydrophobic interactions, which was hypothesized to improve differentiation of the sugars compared to boronic acids alone. The boronic acid-peptide receptors were generated in a combinatorial fashion on beads (**Scheme 1.2**), and these resin-bound receptors were exposed to solutions of saccharide (80 μM), washed, and then exposed to a solution of colorimetric indicator. As the indicator solution was applied, the beads were interrogated with a CCD camera to measure the change over time of the indicator color. The data collected was analyzed by LDA (**Figure 1.10**) and had a cross validation of 100%, indicating complete discrimination of each of the sugars or sugar-derivatives. To further demonstrate the robustness and generality of the system, the researchers went on to use this system and the training set data from the analysis of the pure saccharide samples to identify sucralose in a real-world

beverage sample. Overall, this work demonstrates the advantages of incorporating additional interaction opportunities between a receptor and analyte as well as incorporating indicator displacement for colorimetric signaling.



Scheme 1.2 *General structure of pentapeptides from combinatorial library.*
(Reproduced from Ref. 25: © American Chemical Society, 2007)

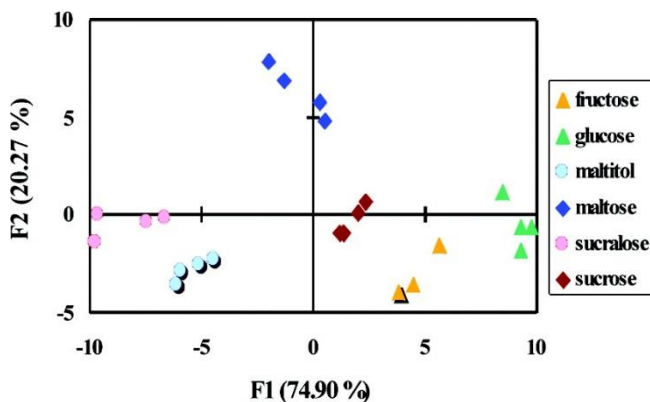
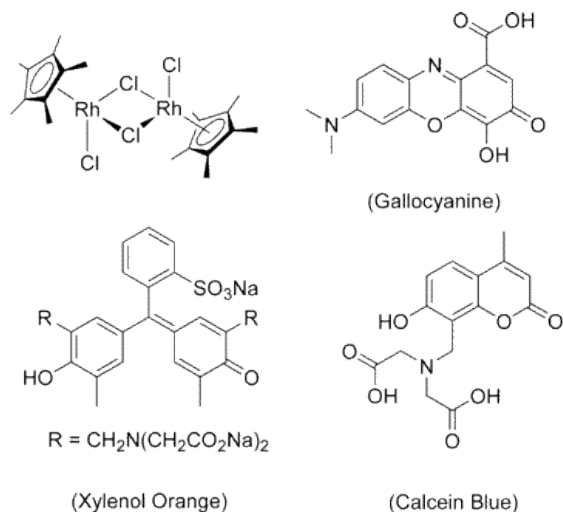


Figure 1.10 *Discrimination of sugars with a boronic acid-peptide receptor array.* LDA plot for the array by Anslyn *et al.* Cross validation: 100%. (Reproduced from Ref. 25: © American Chemical Society, 2007)

1.2.2 Amino Acids

In contrast to sugars, amino acids consist of a variety of functional groups, including hydrophobic, positively-charged, negatively-charged, and aromatic side chains.

These functionalities allow for perhaps simpler discrimination by chemosensors compared to sugars. For example, Severin and coworkers developed a system for differentiating twenty natural amino acids using indicator displacement-based sensors from commercially available components.²⁶ A rhodium cyclopentadiene species was employed as the receptor to which the amino acids bind. Three dyes were used as colorimetric indicators that compete with the amino acid for the rhodium species, which modulates the absorbance of the indicators (**Scheme 1.3**).



Scheme 1.3 *Rhodium species and dyes.* (Reproduced from Ref. 26: © American Chemical Society, 2005)

Different pH solutions of each amino acid/dye combination were also used. Since the binding affinity of dye and amino acid changes with pH, this provided additional discriminating information about each amino acid. The absorbance data from the array was subjected to LDA and PCA. In the LDA, the cross validation showed good classification of each amino acid with some misclassifications occurring between valine and isoleucine. These amino acids both contain alkyl side chains, so it is unsurprising that

they would be difficult to distinguish. In the PCA plot (**Figure 1.11**), good visual separation of the amino acids was obtained. Furthermore, the authors noted that amino acids with similar classes of side chain were placed near one another on the plot. For example, the aromatic amino acids phenylalanine, tryptophan, and tyrosine are close to one another on the scoreplot as are serine and threonine. This work highlights how simple arrays can be constructed from commercially available components that have good discriminatory power for structurally similar analytes.

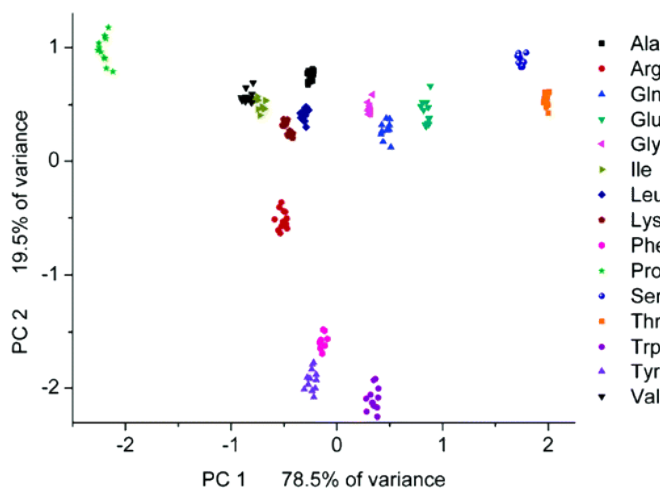


Figure 1.11 *Discrimination of natural amino acids by IDA-based array by Severin et al.* PCA plot from array data. (Reproduced from Ref. 26: © American Chemical Society, 2005)

Anslyn and coworkers also used an IDA-based array for amino acid differentiation that was further able to distinguish enantiomers.²⁷ The array was constructed from Cu(II) complexes derived from bidentate N-donor ligands and salicylate-derived chromophores that are modulated in their absorbance upon metal coordination. Since the Cu(II) complexes are themselves chiral, the complexes exhibit a preference for L or D amino acids, allowing for enantioselective determination of the amino acids. The absorbance data

from the array was subjected to PCA, which showed complete separation of each amino acid, primarily along PC1, as well as the L and D forms of each amino acid, primarily along PC2.

1.2.3 Nucleotides

Nucleotides, the components that make up DNA and RNA, are perhaps more challenging analytes for chemosensor targeting compared to amino acids. Nucleotides consist of aromatic cores and hydrogen bond donor and acceptor functionalities. Hence, the chemical and structural differences between nucleotides are subtle and difficult to target.

One example of an approach to nucleotide sensing was developed by Mirsky and coworkers using molecularly imprinted polymers (MIPs) to synthesize semi-selective receptors for nucleotides.²⁸ The receptors are synthesized by adsorbing an alkanethiol matrix onto a gold surface. The image of the analyte is imprinted within the alkanethiol matrix using a thiol derivative of the analyte. The resulting template is an MIP that is selective for detection of adenine, cytosine, thymine, and uracil. The binding of the nucleotides to the MIPs was monitored electrochemically, and the data was analyzed by PCA. The PCA plot (**Figure 1.12**) shows good separation of each nucleotide. This work demonstrates a powerful approach to generate semi-selective receptors for subtly different analytes.

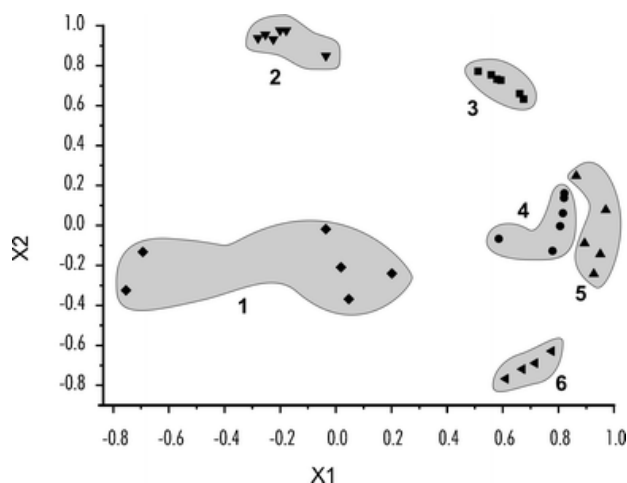


Figure 1.12 *Discrimination of nucleotides by an array of MIPs by Mirsky et al.* The PCA plot distinguishes 1) caffeine, 2) uracil, 3) adenine, 4) cytosine, 5) thymine, and 6) uric acid with X1 and X2 comprising 75% of the total variance. (Reproduced from Ref. 28: © Royal Society of Chemistry, 2003)

Another macromolecule-based strategy for nucleotide detection was investigated by Prins and coworkers.²⁹ In this approach, a gold nanoparticle was mixed with several different anionic fluorescent dyes that bind to the nanoparticle's surface monolayer and can be displaced differentially by nucleotides. The cationic counterion for the gold nanoparticle was also varied between Zn(II) and Cu(II) to obtain additional discriminatory information. By subjecting eight nucleotides to the array at two different concentrations, the PCA plot showed complete separation of all nucleotides and concentrations (**Figure 1.13**). This study demonstrates how the nonspecific and complex interactions that occur on the nanoparticle monolayer surface can be exploited for distinguishing structurally similar analytes.

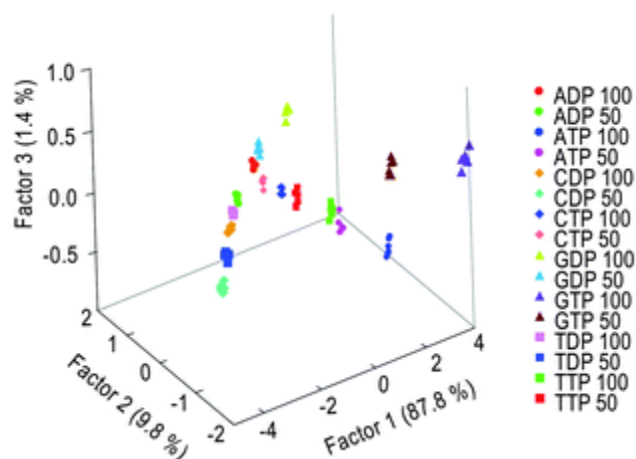
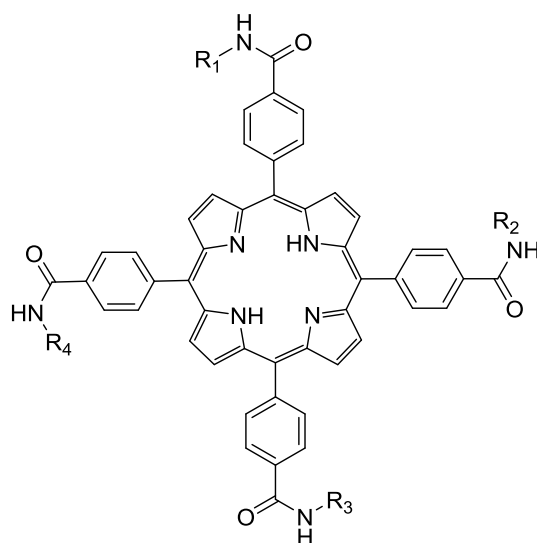


Figure 1.13 Differentiation of nucleotides by the gold nanoparticle/fluorescent dye array of Prins *et al.* The PCA plot discriminates eight nucleotides at 50 μM and 100 μM concentrations. (Reproduced from Ref. 29: © Royal Society of Chemistry, 2013)

1.3 ARRAYS FOR PROTEINS

Receptors for proteins must be able to target the relatively large, solvent-exposed surface of these molecules. Hence, larger supramolecular scaffolds or macromolecular structures have often been employed for protein detection in cross-reactive arrays.

For instance, Hamilton and coworkers synthesized a library of fluorescent porphyrin receptors with four peptidic arms in a combinatorial fashion to target proteins.³⁰ The peptide arms contained a variety of sequences with aromatic, acidic, and basic amino acids, resulting in a library of 35 porphyrin receptors with varying hydrophobicity, charge, size, and symmetry (**Scheme 1.4**). From the 35-member library, eight porphyrin receptors were chosen for an array to target cytochrome c, ferredoxin, cytochrome c551, and myoglobin.



Scheme 1.4 *Structure of porphyrin scaffold.* R groups were TyrAsp, Lys, TyrLys, Asp, or Lys1,5-pentane-diamine.

The fluorescence of the porphyrin is differentially quenched in the presence of the proteins. The authors found that the proteins were bound most strongly by receptors with complementary characteristics such as charge. The fluorescence responses from the array clearly distinguished each protein (**Figure 1.14**). The authors went on to refine this array to differentiate metal and metal-free proteins.³¹ In this study, they employed PCA to process the array data and obtained good discrimination of ten protein analytes.

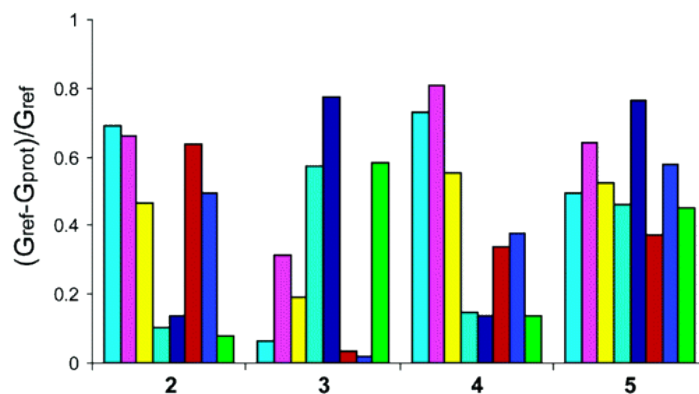
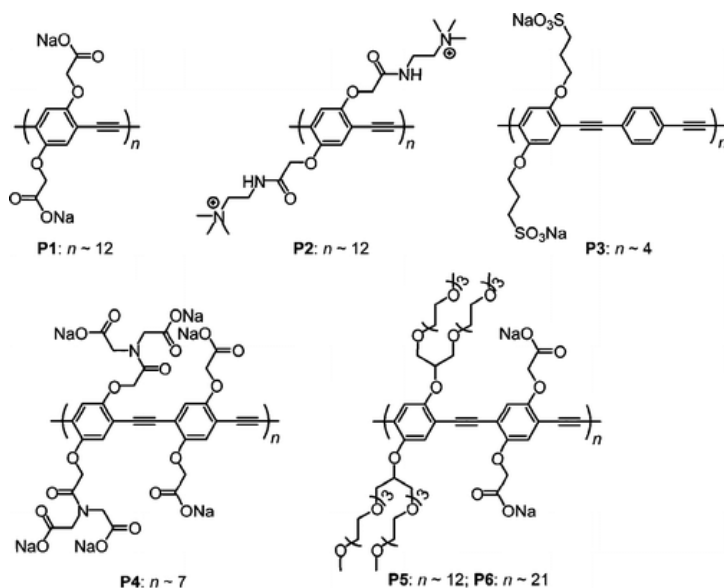


Figure 1.14 *Protein fingerprints from porphyrin receptor array by Hamilton et al.* Key: 2) cytochrome c, 3) ferredoxin, 4) cytochrome c551, and 5) myoglobin. The bars represent the fluorescence modulation of the solutions with protein (G_{prot}) compared to solutions without protein (G_{ref}). (Reproduced from Ref. 30: © American Chemical Society, 2004)

Rotello and coworkers developed an array of water-soluble, conjugated polymers with terminally-charged residues to discriminate a panel of 17 proteins.³² The polymers were poly(*p*-phenylene ethynylene)s (PPEs) and were chosen because their emission is highly sensitive to environmental changes and because these large polymers have the potential for multivalent interactions with the protein analytes. They chose six polymers with different charge characteristics and molecular weights (**Scheme 1.5**). In most cases, the proteins quenched the fluorescence of the polymers upon binding, implying that the protein binding stimulates a modulation of the electronic state of the polymer. The array was able to discriminate the 17 proteins (**Figure 1.15**), demonstrating the generalizability of these receptors for targeting a variety and large number of analytes.



Scheme 1.5 Structures of the PPEs used by Rotello *et al.* (Reproduced from Ref. 32: © American Chemical Society, 2007)

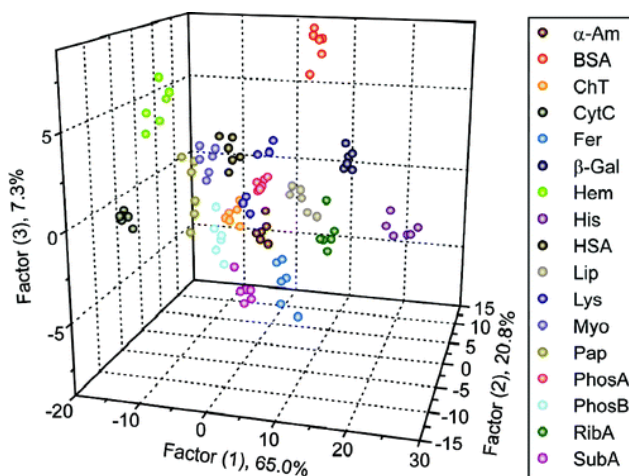
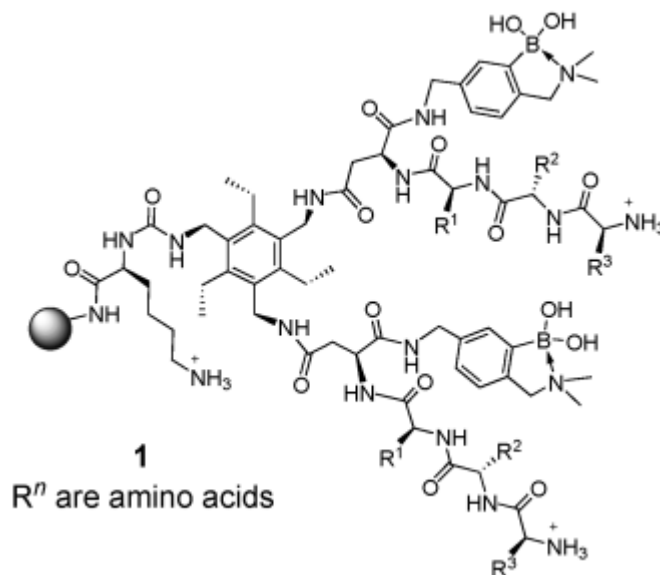


Figure 1.15 Differentiation of 17 proteins by a fluorescent polymer array by Rotello *et al.* PCA plot generated from fluorescence data from the array. (Reproduced from Ref. 32: © American Chemical Society, 2007)

Based on a similar strategy to Hamilton *et al.* of using peptide-based receptors, Anslын and coworkers generated a combinatorial library of receptors with different

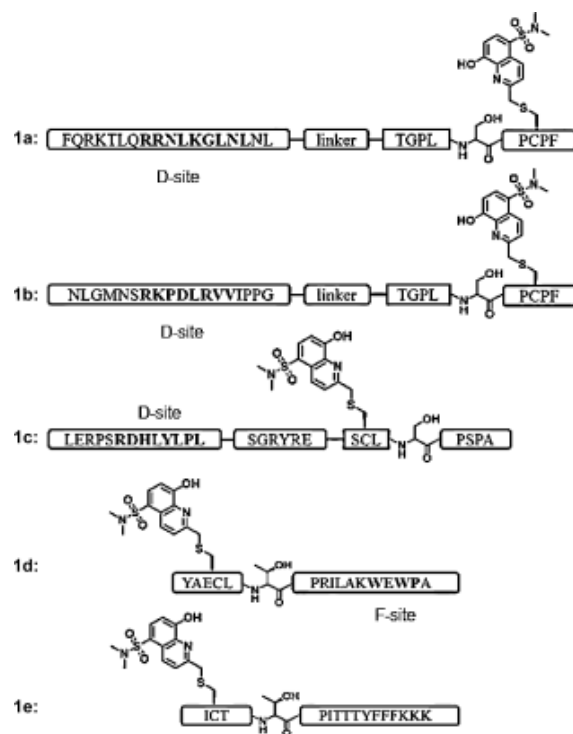
peptidic arms on a substituted benzene scaffold (**Scheme 1.6**).³³ The peptides contained boronic acid moieties to target glycoproteins. From the library, 29 receptors were chosen at random to use in the array in conjunction with the bromopyrogallol red chromophore as a reporter based on indicator displacement (see 1.2.1). Five proteins, including two glycoproteins, were exposed to the array, and the data was analyzed using PCA, which showed discrimination of the proteins. The 29 receptors were then sequenced to determine their identities. Based on the PCA data and the structures of the receptors, the authors determined that subtly different receptors were able to distinguish the proteins.



Scheme 1.6 *Boronic acid-functionalized peptides.* (Reproduced from Ref. 33: © Wiley-VCH, 2005)

Perhaps a more challenging task than targeting a wide range of proteins and protein classes is targeting several proteins within a single class that are structurally similar to one another. For instance, Anslyn and coworkers patterned mitogen-activated protein kinases (MAPKs) using an array of SOX-peptides.³⁴ The peptide sequences were selected, based

on previous literature, to have docking sites for the various kinases in the analyte panel (**Scheme 1.7**) so that the peptide will be phosphorylated by the kinase. The adjacent SOX fluorophore increases in its fluorescence upon phosphorylation of the peptide, acting as the reporter. The proteins studied were nine isoforms of MAPK, and they were applied to the solutions of SOX-peptide, and the fluorescence of the solution was monitored over time. The fluorescence data was analyzed by LDA (**Figure 1.16**) and PCA, both of which showed good separation of the kinases. The PCA plot revealed cross-reactivity between the SOX-peptides and kinases, which was surprising considering that some of the peptide sequences were chosen to be selective for a particular kinase. Additionally, the authors used their array for quantitative analysis of kinases using SVM analysis. Overall, the approach used in this study demonstrates that employing receptors with more selectivity, rather than simply targeting general structural characteristics such as charge or hydrophobicity, can be used to detect subtle differences in a class of very similar and large analytes.



Scheme 1.7 *SOX-peptides with the kinase docking sites indicated. Linker = three 6-aminohexanoic acid moieties. (Reproduced from Ref. 34: © Wiley-VCH, 2014)*

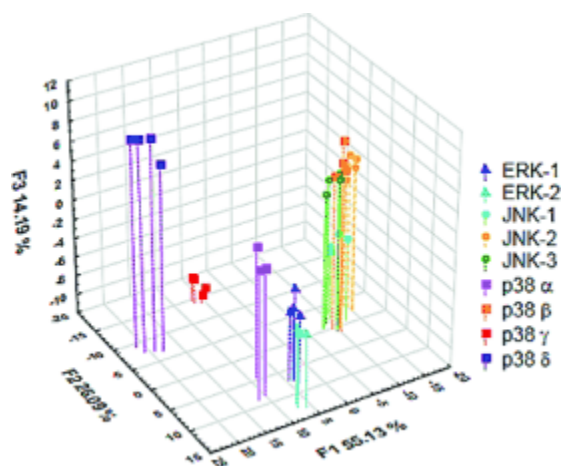


Figure 1.16 *Discrimination of MAPK isoforms by SOX-peptide receptors by Anslyn et al. LDA plot with cross validation of 94.4%. (Reproduced from Ref. 34: © Wiley-VCH, 2014)*

1.4 ARRAYS FOR WHOLE CELLS

Perhaps the ultimate challenge in the area of array-based biological sensing is discrimination of whole cells based on their surface features or on the volatile organic compounds (VOCs) they emit. Cell surfaces are covered in a variety of lipids, carbohydrates, and proteins. The number and type of these molecules can differ based on the cell type or on the health state of the cell,^{35,36} so there is an interest in detecting these subtle differences. Furthermore, bacteria produce a variety of gaseous metabolites that could be used to identify a particular species or strain. Since the surfaces and emitted VOCs of cells consist of many different compounds, a major advantage of using an array-based approach to cell surface sensing is that high specificity is not necessary. Additionally, array detection does not require that the constituent compounds be fully characterized.

1.4.1 Prokaryotic cells

Rapid identification of different species and strains of human pathogenic bacteria remains an important sensing challenge.³⁷ Differential sensing methods have been applied to bacteria sensing, mostly utilizing headspace analysis of mature cultures and “electronic nose” array technologies that make use of electrochemical signals. Some efforts toward array-based methods with optical detection have been undertaken. The colorimetric sensor array used by Suslick and coworkers has been applied to headspace analysis of bacterial cultures.³⁷ Several species and strains of bacteria were cultured and monitored over time by a 36-dye array and flatbed scanner to digitize the array image. Based on the changing color of the array, it was concluded that the nature of the volatile compounds emitted by the culture changes over time. Data collected from monitoring of the bacteria was analyzed using several pattern recognition methods. Using simple Euclidean distances, which compare only the overall response of the array, correct classification of 162 out of 164 instances was achieved. The only exception was misclassification between *E. faecalis*

and vancomycin resistant mutant (VRE). HCA had a similar success rate and misclassified only three instances, all of which were between *E. faecalis* and vancomycin resistant mutant (VRE) as well. PCA showed high dimensionality, with the first three dimensions accounting for only 79% of total variance. Finally, the PCA space (**Figure 1.17**) was used to classify new inputs (i.e. unknown cultures of one of the 11 types studied) based on their array responses. Using the first seven PCs, 99.2% correct classification was achieved. Since the full extent of these volatile metabolites are not known and would be difficult to fully characterize for every species of bacteria at every time point, it is clear that this detection challenge is well-addressed by array-based methods compared to developing specific sensors for each volatile metabolite. Versions of Suslick's colorimetric sensor array have been commercialized by Metabolomx as disposable breath analyzers for cancer and bacterial metabolite detection.

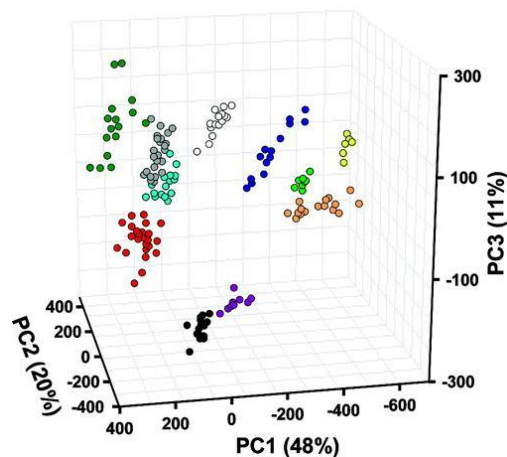


Figure 1.17 PCA plot from response of array by Suslick *et al.* to ten bacterial strains. Key: black, *S. aureus*; red, MRSA; dark green, *S. epidermidis*; purple, *S. sciuri*; orange, *P. aeruginosa*; white, *E. faecium*; light blue, *E. faecalis*; gray, *E. faecalis* VRE; yellow, *E. coli* 25922; light green, *E. coli* 53502; dark blue, control. (Reproduced from Ref. 37: © American Chemical Society, 2011)

An approach to bacteria detection that does not involve headspace analysis of cultures instead targets the lipopolysaccharides (LPS) of Gram negative bacteria.³⁸ Lipopolysaccharides are glycolipids found in the outer membrane of Gram negative bacteria. The sugar portion differs for different bacterial species and strains, so it provides a target analyte for differentiation of bacteria. The sensors employed in this assay were diacetylene liposomes. The functionalization of these liposomes (e.g. with tryptophan) was chosen to be generic for interactions with sugars instead of highly specific, as is normally the case when liposomes are used as sensors. These diacetylene liposomes are known to change color from blue to red upon analyte binding, providing an assay signal (**Figure 1.18**). The LPS harvested from five strains of bacteria were exposed to three different liposomes, and the color change of the solution was quantified from absorbance spectra of the solutions before and after exposure to analyte. To add more variables to the array, measurements were taken in the presence of a detergent (SDS) and a metal chelating agent (EDTA). The resulting response patterns were found to be unique for each bacterial species upon visual inspection. While the need to harvest the LPS from bacteria samples probably limits the use of this method as a rapid detection system, the authors indicate that further avenues of study will include development of a version of this approach that performs detection of intact bacteria, which would greatly expand the method's applicability.

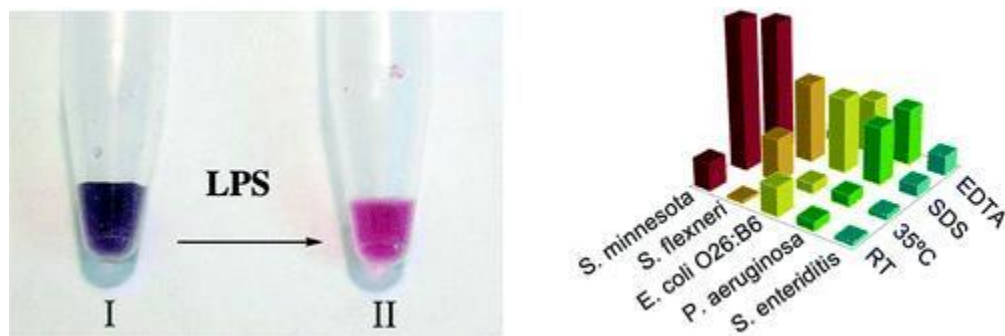


Figure 1.18 Color response of liposome solution in the presence of LPS. Left: visible color change of liposome upon addition of LPS. Right: Graph of liposome sensor conditions to different bacteria. (Reproduced from Ref. 38: © American Chemical Society, 2004)

Finally, Bunz and coworkers have targeted the bacterial cell surface with their gold nanoparticle-conjugated polymer sensors.³⁹ Initially, anionic polymers are associated with cationic gold nanoparticles and thus are quenched in their fluorescence. Upon exposure to a negatively-charged bacterial cell surface, the gold nanoparticles associate with the cell surface, releasing the polymers and turning on the fluorescence of the polymers (**Figure 1.19**). Three hydrophobic ammonium-functionalized nanoparticles were employed in the array as well as one anionic fluorescent polymer. Twelve different kinds of bacteria were exposed to the array, with each exhibiting a unique fluorescence response to the array. The twelve bacteria were successfully discriminated from one another, including different strains of *E. coli*. Identification of randomly selected solutions of one of the twelve bacteria was accomplished with 95% accuracy. Since it targets whole bacteria, this method has good potential applicability for screening of bacteria. Furthermore, this assay also benefits from the high sensitivity characteristic of turn-on fluorescence signals.

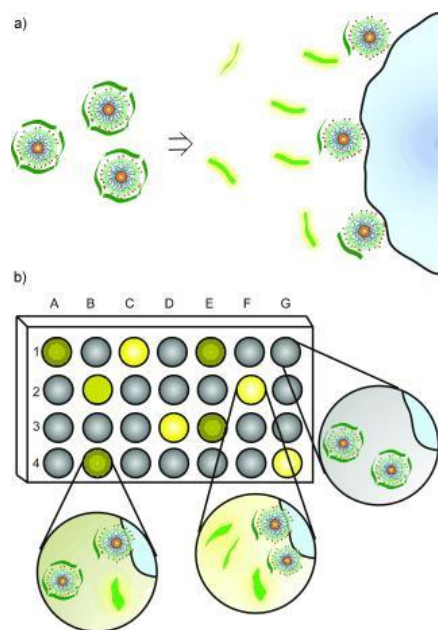


Figure 1.19 *Schematic of fluorescent polymer/gold nanoparticle array for bacterial cell surfaces.* a) Fluorescence turn on occurs upon displacement of polymers from the nanoparticle surface, b) Differential displacement of polymers by differing affinity of the nanoparticle for the cell surface leads to fluorescence modulation. (Reproduced from Ref. 39: © American Chemical Society, 2011)

1.4.2 Eukaryotic cells

Mammalian cell surfaces are of particular interest as sensing targets because these surfaces are known to be indicative of disease states, such as cancer. An example of an array-based approach by Rotello and coworkers differentiated cell lines and health states of genetically identical cells.⁴⁰ The array consisted of PPE-derived polymers like those used in the arrays described in section 1.3 for proteins and in section 1.4.1 for bacterial cells. In this study, eight polymers were chosen with varying charge characteristics and degrees of polymerization in order to provide diversity in the receptors. In the presence of cells, the polymers aggregated differentially, resulting in modulation of their intrinsic fluorescence. Of the eight polymers, four were found by LDA analysis to contribute most

to the differentiation of the cell lines. From the fluorescence responses of these four polymers to four cancer cell lines: HeLa (cervical), MCF7 (breast), HEPG2 (liver), and NT2 (testicular), an LDA plot was constructed (**Figure 1.20**) that showed good differentiation of the cell lines. Most of the discrimination is along F1 (96.7%), indicating that the receptors are providing similar information to the analysis. Three of the four polymers selected for the array were cationic, which suggests that F1 could represent varying degrees of negative charge on the cell surfaces. The authors went on to study the ability of the array to distinguish healthy from unhealthy cells. Three genetically identical cell lines that were healthy (CDBgeo), cancerous (TD), and metastatic (V14) were subjected to the array. The LDA analysis of the data from this assay was able to separate the three health states, although the dimensionality of the data was also low (F1 96.3%).

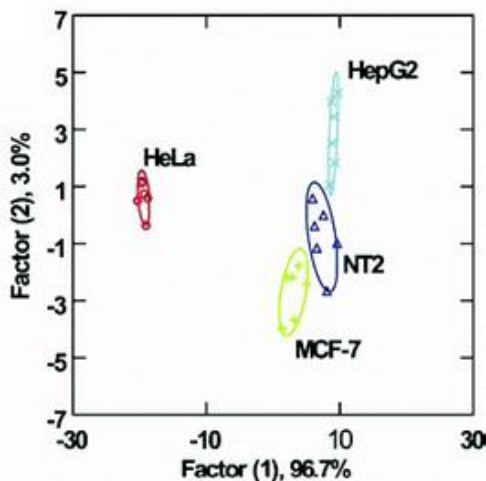



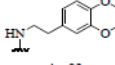
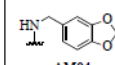
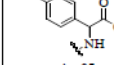
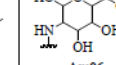







Figure 1.20 *Discrimination of cancer cell lines with the array from Rotello et al.* LDA plot discriminating the four cell lines. (Reproduced from Ref. 40: © American Chemical Society, 2010)

Another recent example of the application of array sensing to cancer cell surfaces was executed by Yan and coworkers. In this study, dual-ligand nanoparticles were used in an array to differentiate three cancer cell lines that all overexpress folate receptors.⁴¹ The gold nanoparticles were coated with two different recognition units in order to achieve multivalent binding interactions with the cell surfaces. One of the ligands on all of the nanoparticles was folic acid, which they expected to interact with all of the cell lines' folate receptors. The second ligand was varied as shown in **Table 1.1** in a combinatorial fashion to yield 30 unique nanoparticle receptors. The cells were exposed to the array of nanoparticles and then were washed to remove unbound material. The binding of the particles to the cells was determined by quantifying the gold content of the samples by inductively-coupled plasma spectrometry (ICP-MS). The authors expressed their data as a heat map (**Figure 1.21**). The heat map shows a wide range of interactions between the dual-ligand nanoparticles and the cell lines. Significantly, the binding of the dual-ligand particles was enhanced compared to mono-ligand particles and blank control samples. This finding suggested that the authors were correct in their hypothesis that the dual-ligand feature would allow for multivalent binding to the surfaces.

Table 1.1 Structures of the 30 dual-ligand nanoparticles. FA= folic acid, TA = 1,2-dithiolan-3-yl pentanamide moiety, Tyr = tyrosine. (Reproduced from Ref. 41: © American Chemical Society, 2011)

 TA-PEG-FA TA-Tyr-R ₁ (1-6)R ₂ (1-5)		R ₁					
		 Am01	 Am02	 Am03	 Am04	 Am05	 Am06
R ₂	 Ac01	D11	D12	D13	D14	D15	D16
	 Ac02	D21	D22	D23	D24	D25	D26
	 Ac03	D31	D32	D33	D34	D35	D36
	 Ac04	D41	D42	D43	D44	D45	D46
	 Ac05	D51	D52	D53	D54	D55	D56

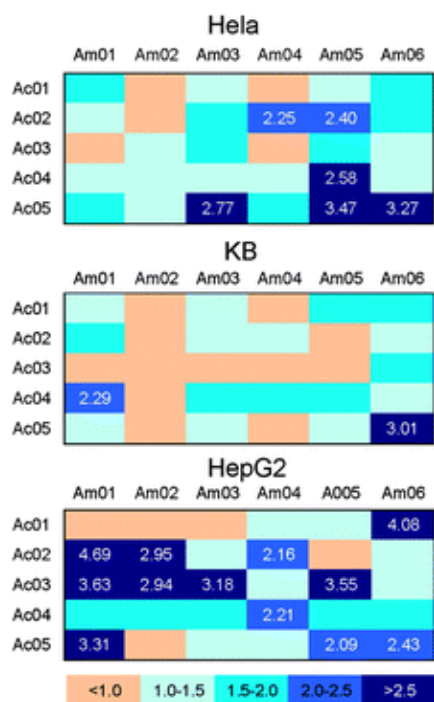


Figure 1.21 *Discrimination of cell lines with array by Yan et al.* Heat maps calculated from gold content of cells. Top: HeLa cells, Middle: KB cells, Bottom: HepG2 cells. The darker colors indicate a higher gold content and stronger binding compared to lighter colors. (Reproduced from Ref. 41: © American Chemical Society, 2011)

1.5 SUMMARY AND OUTLOOK

Array-based methods have demonstrated their usefulness in addressing many sensing challenges. By using collections of semi-selective receptors along with chemometric analysis, arrays can be designed to discriminate various structurally similar analytes and analyte mixtures. The approaches described in this chapter focused on assays that utilize optical signals because, in general, colorimetric and fluorescence sensors have good sensitivity and tunability as well as great practical potential in terms of ease of use and price of instrumentation. While this approach to sensing has had success in highly controlled environments, there are still challenges to overcome before many of these methods could be implemented outside the research laboratory. The ultimate usefulness

of array sensing for practical applications lies in its predictive power. The predictive ability of an assay in a controlled laboratory environment can differ significantly from diagnostic capability in the field for a number of reasons.

Olfactory and gustatory systems in mammals provide mostly qualitative information about odors and tastes. The response to a particular odor or taste is apparently unique enough that the concentration of the odorant or tastant is not very important to correctly identifying it; rather we perceive at most that the odorant or tastant is present in a high, medium, low, or trace amount. However, many artificial arrays have a significant concentration dependence, which could lead to misidentification of analytes if the system is not properly trained at multiple concentrations. Also, if the response pattern is not unique enough for each analyte over the desired range of possible concentrations, then the assay is likely to confuse analytes. In addition, the identification of complex mixtures is an important advantage of differential sensing over the use of highly specific receptors designed for particular molecules. The mixture need not be fully characterized in order to be detected, which is particularly useful when dealing with analyte mixtures that are impossible or difficult to characterize extensively such as explosive mixtures or bacterial metabolites. However, when the array is presented with the complex mixture, it can be trained to identify that mixture but not the components of the mixtures. Changing the mixture slightly (e.g. different relative concentrations of the same components or adding new components) may also confuse the system. In general, interferents can be problematic for analyte identification if they are not anticipated and included in the training process.

Data from the responses of the analytes to the array sensors and the analyte identities or classes are presented to the system as a training set. When the system is presented with responses from an unknown analyte, it has only the information it was given to base its prediction on. In order to create array-based assays that can accurately predict

the identity of the desired analytes under conditions that are reasonable for a specific application, several conditions must be achieved. The response of analytes to the array should be reproducible. The system has to be able to match responses from unknowns to responses from a training set of analytes. Hence, these responses have to be similar enough from exposure to exposure of the same analyte to the array in order to do this successfully. The sensor array should have sufficient semi-selectivity for the range of analytes of interest. If the sensors are less selective, then low dimensionality in the data is achieved, and the system will have fewer bases on which to discriminate analytes correctly. Each analyte must have a distinct enough pattern of response to be identified from other analytes that the system is trained to identify. Appropriate training of the system is vital, as the assay will only be capable of doing what it has been trained to do. The specific chemometric method used may also have to be optimized for a particular system. A balance between accuracy and time/cost of the overall sensing system will still have to be struck based on the application. Overall, the studies described in this chapter represent progress toward the application of array-based sensing methodologies to important sensing challenges.

1.6 ADDITIONAL INFORMATION

Portions of this chapter were adapted from a review published in *Chemical Society Reviews*.⁴²

1.7 REFERENCES

- (1) Stitzel, S. E.; Aernecke, M. J.; Walt, D. R. *Annu. Rev. Biomed. Eng.* **2011**, *13*, 1–25.
- (2) Buck, L. B. *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 6128–6140.
- (3) Doty, R. L. *Handbook of Olfaction and Gustation*; 2nd ed.; Marcel Dekker: New York, 2003.

- (4) Axel, R. *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 6110–6127.
- (5) Malnic, B.; Hirono, J.; Sato, T.; Buck, L. B. *Cell* **1999**, *96*, 713–723.
- (6) Lavigne, J. J.; Anslyn, E. V. *Angew. Chemie Int. Ed.* **2001**, *40*, 3118–3130.
- (7) Nakamoto, T. *Handbook of Machine Olfaction: Electronic Nose Technology*; Pearce, T. C.; Schiffman, S. S.; Nagle, H. T.; Gardner, J. W., Eds.; Wiley-VCH: Weinheim, Germany, 2006.
- (8) Jurs, P. C.; Bakken, G. A.; McClelland, H. E. *Chem. Rev.* **2000**, *100*, 2649–2678.
- (9) Jolliffe, I. T. *Principal Component Analysis*; 2nd ed.; Springer: New York, 2002.
- (10) Varmuza, K.; Filzmoser, P. *Introduction to Multivariate Statistical Analysis in Chemometrics*; CRC Press: Boca Raton, FL, 2009.
- (11) Brereton, R. G. *Chemometrics Data Analysis for the Laboratory and Chemical Plant*; Wiley: West Sussex, UK, 2003.
- (12) Burns, J. A.; Whitesides, G. M. *Chem. Rev.* **1993**, *93*, 2583–2601.
- (13) Nguyen, B. T.; Anslyn, E. V. *Coord. Chem. Rev.* **2006**, *250*, 3118–3127.
- (14) Agbaria, R. A.; Oldham, P. B.; McCarroll, M.; McGown, L. B.; Warner, I. M. *Anal. Chem.* **2002**, *74*, 3952–3962.
- (15) Martinez, A. W.; Phillips, S. T.; Whitesides, G. M.; Carrilho, E. *Anal. Chem.* **2010**, *82*, 3–10.
- (16) Lacowicz, J. R. *Principles of Fluorescence Spectroscopy*; Plenum Press: New York, 1986.
- (17) Skoog, D. A.; Holler, J. F.; Crouch, S. R. *Principles of Instrumental Analysis*; 6th ed.; Thomson Brooks/Cole: Belmont, CA, 2007.
- (18) Germain, M. E.; Knapp, M. J. *Inorg. Chem.* **2008**, *47*, 9748–9750.
- (19) Aernecke, M. J.; Walt, D. R. *Sensors Actuators B Chem.* **2009**, *142*, 464–469.
- (20) Springsteen, G.; Wang, B. *Tetrahedron* **2002**, *58*, 5291–5300.

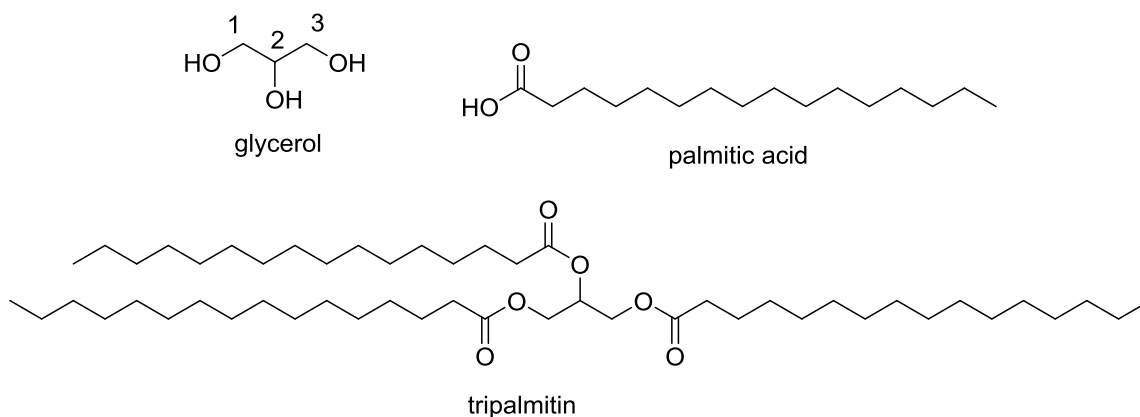
- (21) James, T. D.; Sandanayake, K. R. A. S.; Shinkai, S. *Angew. Chemie Int. Ed. English* **1996**, *35*, 1910–1922.
- (22) Lee, J. W.; Lee, J.-S.; Chang, Y.-T. *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 6485–6487.
- (23) Lim, S. H.; Musto, C. J.; Park, E.; Zhong, W.; Suslick, K. S. *Org. Lett.* **2008**, *10*, 4405–4408.
- (24) Schiller, A.; Wessling, R. A.; Singaram, B. *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 6457–6459.
- (25) Edwards, N. Y.; Sager, T. W.; McDevitt, J. T.; Anslyn, E. V. *J. Am. Chem. Soc.* **2007**, *129*, 13575–13583.
- (26) Buryak, A.; Severin, K. *J. Am. Chem. Soc.* **2005**, *127*, 3700–3701.
- (27) Folmer-Andersen, J. F.; Kitamura, M.; Anslyn, E. V. *J. Am. Chem. Soc.* **2006**, *128*, 5652–5653.
- (28) Hirsch, T.; Kettenberger, H.; Wolfbeis, O. S.; Mirsky, V. M. *Chem. Commun.* **2003**, 432–433.
- (29) Pezzato, C.; Lee, B.; Severin, K.; Prins, L. J. *Chem. Commun. (Camb)*. **2013**, *49*, 469–471.
- (30) Baldini, L.; Wilson, A. J.; Hong, J.; Hamilton, A. D. *J. Am. Chem. Soc.* **2004**, *126*, 5656–5657.
- (31) Zhou, H.; Baldini, L.; Hong, J.; Wilson, A. J.; Hamilton, A. D. *J. Am. Chem. Soc.* **2006**, *128*, 2421–2425.
- (32) Miranda, O. R.; You, C.-C.; Phillips, R.; Kim, I.-B.; Ghosh, P. S.; Bunz, U. H. F.; Rotello, V. M. *J. Am. Chem. Soc.* **2007**, *129*, 9856–9857.
- (33) Wright, A. T.; Griffin, M. J.; Zhong, Z.; McCleskey, S. C.; Anslyn, E. V.; McDevitt, J. T. *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 6375–6378.
- (34) Zamora-Olivares, D.; Kaoud, T. S.; Jose, J.; Ellington, A.; Dalby, K. N.; Anslyn, E. V. *Angew. Chemie* **2014**, *126*, 14288–14292.
- (35) Marth, J. D.; Grewal, P. K. *Nat. Rev. Immunol.* **2008**, *8*, 874–887.

- (36) An, H. J.; Kronewitter, S. R.; de Leoz, M. L. A.; Lebrilla, C. B. *Curr. Opin. Chem. Biol.* **2009**, *13*, 601–607.
- (37) Carey, J. R.; Suslick, K. S.; Hulkower, K. I.; Imlay, J. A.; Imlay, K. R. C.; Ingison, C. K.; Ponder, J. B.; Sen, A.; Wittrig, A. E. *J. Am. Chem. Soc.* **2011**, *133*, 7571–7576.
- (38) Rangin, M.; Basu, A. *J. Am. Chem. Soc.* **2004**, *126*, 5038–5039.
- (39) Miranda, O. R.; Li, X.; Garcia-Gonzalez, L.; Zhu, Z. J.; Yan, B.; Bunz, U. H. F.; Rotello, V. M. *J. Am. Chem. Soc.* **2011**, *133*, 9650–9653.
- (40) Bajaj, A.; Miranda, O. R.; Phillips, R.; Kim, I. B.; Jerry, D. J.; Bunz, U. H. F.; Rotello, V. M. *J. Am. Chem. Soc.* **2010**, *132*, 1018–1022.
- (41) Zhou, H.; Jiao, P.; Yang, L.; Li, X.; Yan, B. *J. Am. Chem. Soc.* **2011**, *133*, 680–682.
- (42) Diehl, K. L.; Anslyn, E. V. *Chem. Soc. Rev.* **2013**, *42*, 8596–8611.

Chapter 2: Differential Sensing For the Regio- and Stereoselective Identification of Glycerides: A Qualitative Fingerprinting and Quantitative Analysis Study

2.1 GLYCERIDES

Glycerides are the primary component of animal fats and vegetable oils.¹ They consist of one, two, or three fatty acids esterified on glycerol, and hence are referred to as mono-, di-, and triglycerides, respectively (**Scheme 2.1**). The structural diversity of glycerides derives in part from their fatty acid alkyl groups, which can differ in carbon number (i.e. chain length), the degree of unsaturation, the position of olefins, and the configuration of the olefins (i.e. *cis/trans*). Furthermore, these fatty acid alkyl groups can be connected to the *sn*-1, -2, or -3 carbons on glycerol. Hence, a variety of regio- and stereoisomers can exist for glycerides, posing a challenge for mass spectrometry.² Further, because the differences in chain length primarily result from the presence of greater or fewer methylene groups, NMR spectroscopy can be ambiguous.³



Scheme 2.1 Structure of glycerol, a fatty acid, and a triglyceride. The *sn*-1, -2, and -3 positions of glycerol are labeled.

2.1.1 Importance of Glyceride Determination

The analysis of glycerides is primarily important to the food and nutrition industries for tasks such as authenticating edible oils,⁴ designing foods with certain physical properties,⁵ and studying how fats are digested and absorbed.⁶ In particular, classifying all the various kinds of regio- and stereoisomers of glycerides is biologically important because lipases, enzymes that catalyze the hydrolysis of glycerides into fatty acids and glycerol, exhibit selectivity based on these features of the glyceride substrates. For instance, the position and configuration of olefins, the identity of fatty acid alkyl groups, as well as their position on glycerol (i.e. *sn*-1,3 versus *sn*-2), all contribute to differing biological activity.^{7,8} For example, hormone-sensitive lipase (HSL), which plays a role in mobilizing free fatty acids from acylglycerols in stored fat, prefers 2,3-*sn*-diacylglycerols over 1,3-*sn*-diacylglycerols as substrates.⁹ Studying the selectivity of these lipases has important applications in understanding diseases, including fat malabsorption disorders, hypercholesterolemia (i.e. high blood cholesterol levels), atherosclerosis, and diabetes.^{10,}

11

2.1.2 Glyceride Determination Methods

The most common method of glyceride identification is mass spectrometry (MS).^{2,}
¹² However, as alluded to above, this approach has drawbacks. Because glycerides are neutral molecules, they must be ionized to be analyzed by mass spectrometry. Saponification can be used to obtain the fatty acids, which are both volatile and charged, thereby facilitating MS analysis, but information about the glyceride structure is lost in this process.¹² Electrospray ionization (ESI) and atmospheric pressure chemical ionization (APCI) can be used to ionize glycerides directly; however, the ion yields are low compared to preionized lipid fatty acids.^{13, 14} Furthermore, the ability of a glyceride to be ionized using these methods often varies. For example, ion abundance generally increases with

increasing number of double bonds in the fatty acid alkyl chain and can also depend on fatty acid alkyl chain length.¹⁵ These significant variations in ion abundance means that ionization methodologies must be developed and tailored to a specific application to satisfactorily detect each glyceride of interest.¹³ Finally, these variations render the quantification of glycerides, particularly in a complex mixture, quite challenging when using mass spectrometry.¹⁶

Regio- and stereoisomers further confound the discrimination of glycerides by mass spectrometry, since isomers share the same mass. Other techniques such as chemical derivitization of the glycerides, ion fragmentation, and specialized HPLC must be coupled with mass spectrometry to differentiate isomeric species. For example, ozonolysis has been used to cleave the double bonds in unsaturated glycerides prior to ionization in order to deduce the positions of double bonds.¹⁷ Non-aqueous reverse phase (NARP)-HPLC can resolve *cis/trans* isomers of triacylglycerols and double bond positional isomers after treatment of the olefins with bromine.¹⁸ Silver ion chromatography has been used to separate triacylglycerol positional isomers under specifically-developed solvent and column temperature conditions.¹⁹ Silver cationization as a post-column treatment in conjunction with NARP-HPLC and ion fragmentation has also been employed for triglyceride positional isomer determination (**Figure 2.1**).^{20, 21}

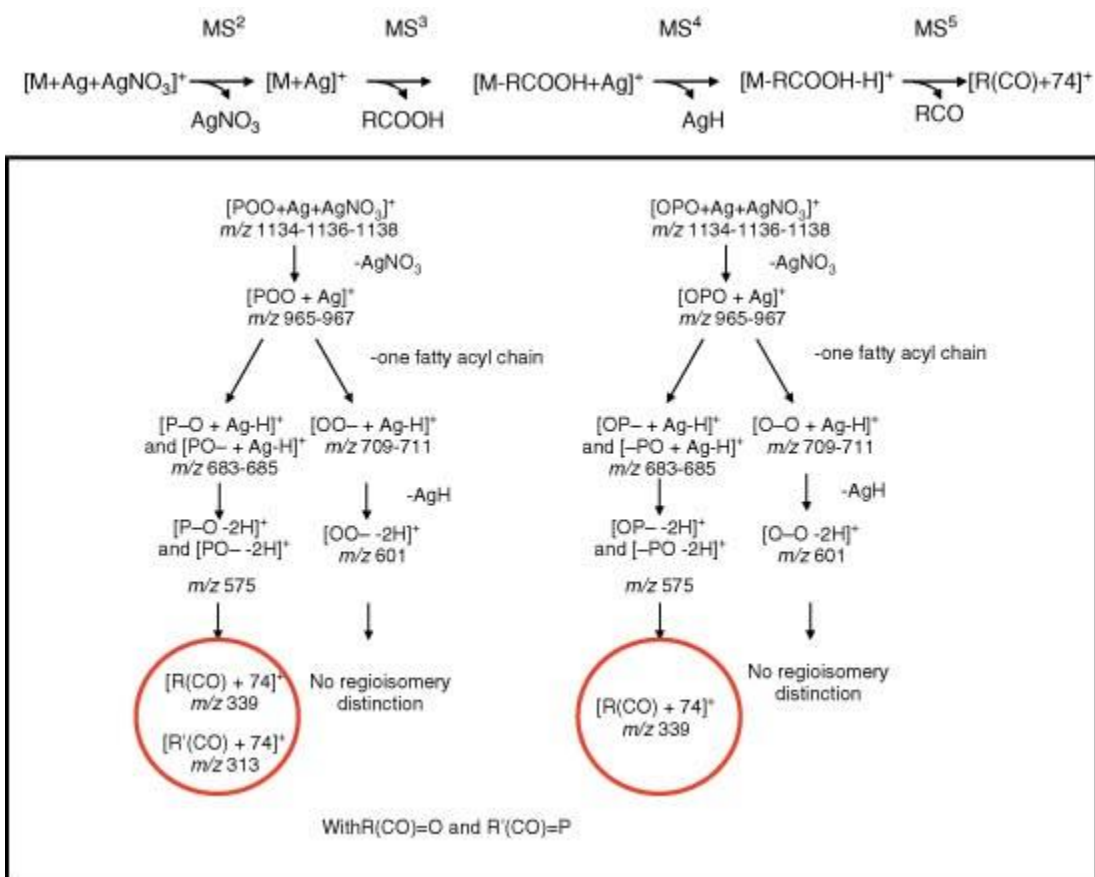
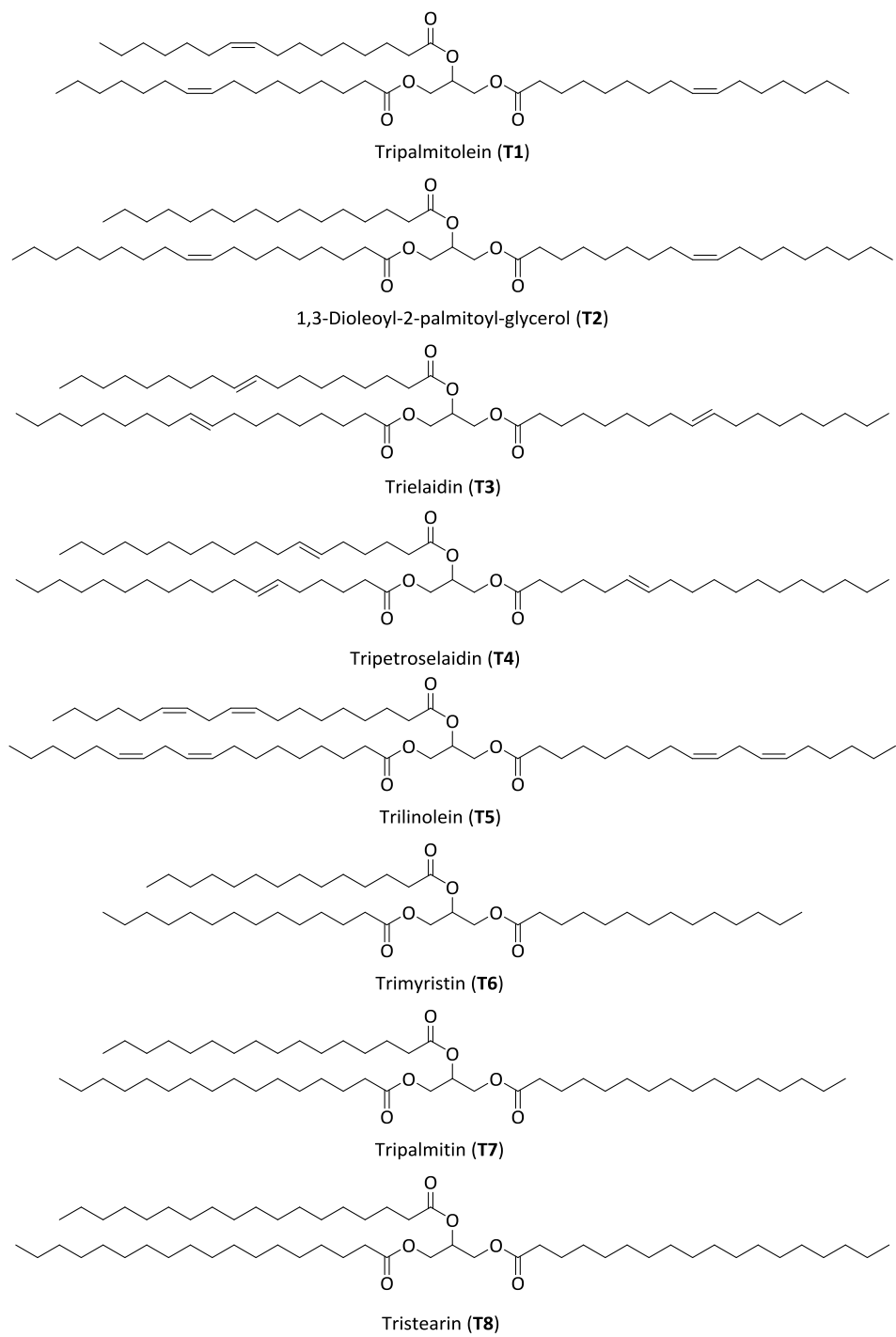


Figure 2.1 *MS⁵ experiment with silver cationization to discriminate regioisomers.* Triglyceride isomers 1(3),2-dioleoyl-3(1)-palmitoyl-*sn*-glycerol (POO) and 1,3-dioleoyl-2-palmitoyl-*sn*-glycerol (OPO) were distinguished by repeated fragmentation that was able to yield a unique ion for POO (m/z 313) that was not observed with OPO. (Reproduced from Ref. 20: © Wiley-VCH 2010)

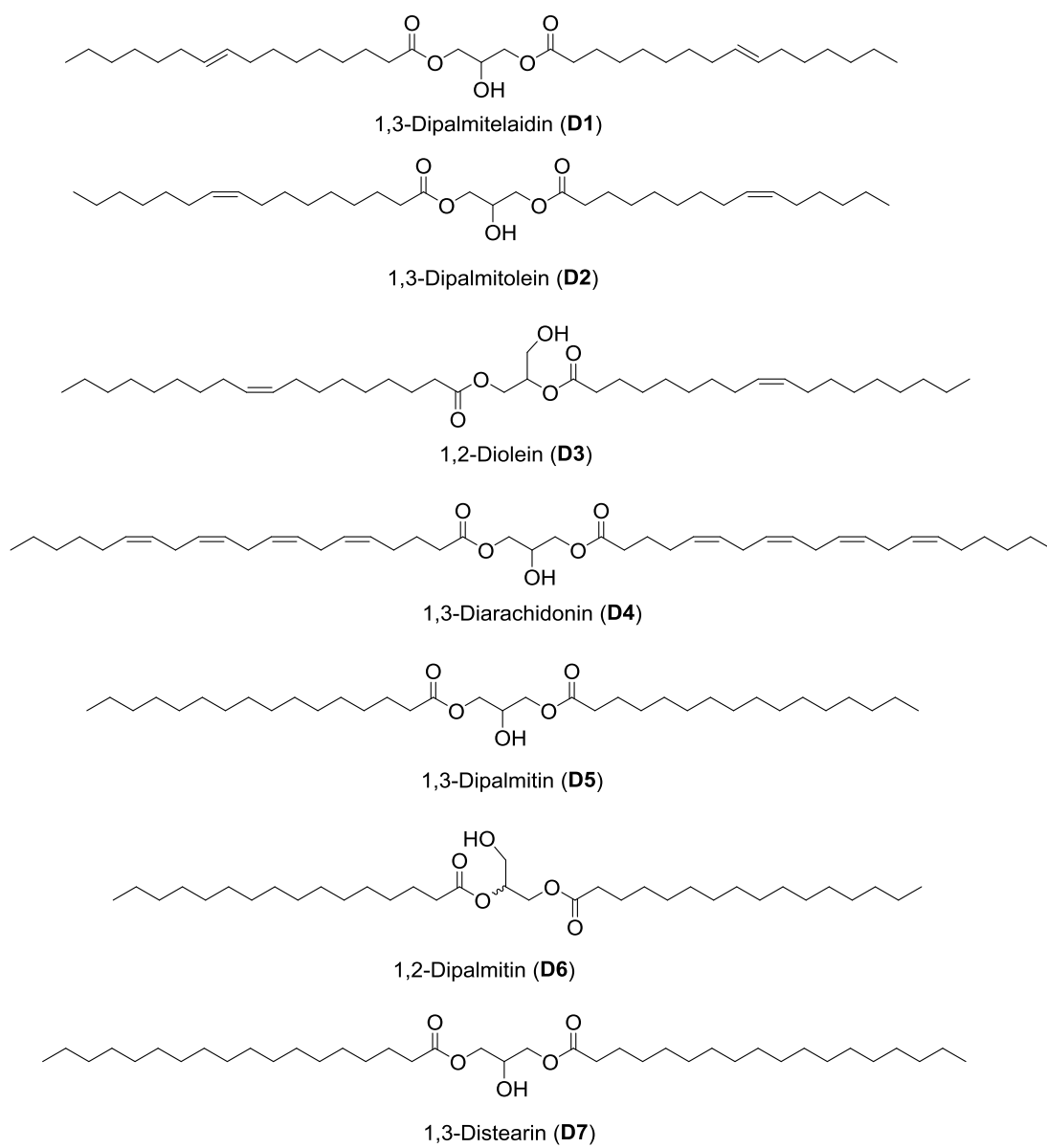
When analyzing complex, real world fat mixtures such as vegetable oils by chromatography, researchers have employed statistical analysis protocols like principal component analysis (PCA) in order to process the large amount of data generated and to build predictive models.^{22,23} Thus, while these current approaches to glyceride isomer analysis are successful, they are complicated, labor intensive, time-consuming, and often inconsistent in their results.²⁰

2.2 AN ARRAY-BASED APPROACH TO GLYCERIDE DETECTION

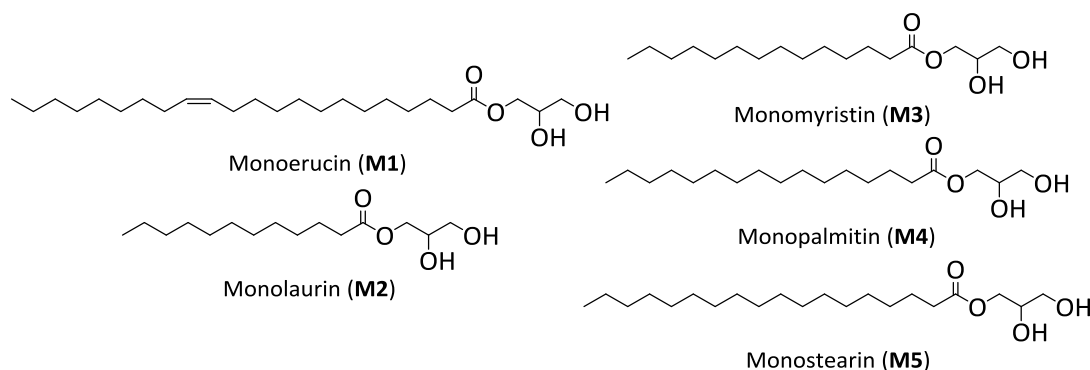
Because glycerides are very structurally similar to one another and lack variety in terms of functional groups, we believed that a differential sensing array-based approach would be most suitable for their classification. Our hypothesis was that if a cross-reactive array could be created that is responsive to the subtle structural differences inherent in glycerides, it could be used to pattern individual glycerides, identify structural features of unknown glycerides, and potentially quantitate glycerides in a mixture. Cross-reactive arrays have been successfully employed in a number of sensing applications, such as the examples described in Chapter 1.²⁴⁻²⁹ In the laboratory, chemometric routines such as principle component analysis (PCA) and linear discriminant analysis (LDA) are used to extract the relevant information from the array (see Chapter 1). Therefore, the goal of this project was to develop an array of cross-reactive receptors that could discriminate glycerides. The glycerides selected for the development of this array are shown in **Schemes 2.2, 2.3, and 2.4**. The panel includes commercially available mono-, di-, and triacylglycerols with fatty acid alkyl groups that are relevant to mammalian biology.³⁰ Moreover, the panel consists of examples of each of the following stereo- and regioisomers: 1) *cis/trans* olefins (**D1** and **D2**; **T2** and **T3**), 2) differing position of the olefin (**T3** and **T4**), and 3) differing position of the fatty acid alkyl groups on glycerol (**D5** and **D6**). Clearly, it would be extremely challenging to create highly-selective receptors for each individual glyceride, and thus a differential sensing method seems the only reasonable approach to creating an optical sensing routine to identify and classify such structures.



Scheme 2.2 *Triglyceride structures.*



Scheme 2.3 *Diglyceride structures.*



Scheme 2.4 *Monoglyceride structures.*

Thus, herein we describe a method using serum albumins to fingerprint glycerides, which classifies them as mono-, di-, or triglycerides. The glycerides were further classified based on fatty acid chain length, ester positions on glycerol, and olefin regio- and stereochemistry. For the unsaturated glycerides in the panel, differentiation based on olefin position and stereochemistry was achieved by the use of a pretreatment olefin metathesis. Using the protocols described herein, structural features of unknown glycerides were identified. Furthermore, the quantitation of trilinolein in a mixture of triglycerides was achieved by application of the standard addition method using a net analyte signal (SANAS) technique presented herein.

2.3 SERUM ALBUMIN

Because glycerides are extremely hydrophobic analytes, we postulated that serum albumins (SAs) would be suitable cross-reactive receptors with which to test our hypothesis. Serum albumin is a common plasma protein that performs several important functions in the bodies of vertebrates.³¹ The protein plays a major role in stabilizing the osmotic pressure of plasma. Additionally, serum albumin binds hydrophobic molecules in order to transport them through the hydrophilic environment of blood. The protein shuttles long-chain fatty acids from the intestines to the liver, from the liver to muscle, and to and

from adipose tissue. In addition to fatty acids, serum albumin transports other endogenous compounds like copper, bile acids, hormones, and vitamins. Albumin also binds toxins and delivers them to the liver for excretion. As part of its role as transporter, serum albumin has also been linked to metabolic effects, such as stimulating lipoprotein lipase activity in adipose tissue.

2.3.1 General Properties

Serum albumin is an approximately 66 kDa protein consisting of about 580 amino acids. Other than cobra albumin, all albumins have 35 cysteine residues that participate in 17 disulfide bridges, which contribute to the superior stability of the protein. The primary sequences of bovine (BSA) and human serum albumin (HSA) have been determined by Edman degradation, cDNA analysis, and mass spectrometry. The crystal structures of bovine³² (**Figure 2.2**) and human albumin³³ (**Figure 2.3**) have been solved as well, showing a heart-shaped structure for both species' albumins. In solution, BSA is ellipsoidal in shape with dimensions of approximately 40 x 140 Å.³⁴ BSA and HSA are both highly helical molecule (67% in HSA and 68% in BSA), with a low content of β -turns (10% in HSA and 17% in BSA).^{33,35} Albumin consists of three domains (I, II, and III), which are each further divided into two subdomains (IA, IB, IIA, IIB, IIIA, and IIIB) as shown in Figure 2.3.

The primary sequence of albumin differs depending on the organism it comes from. For instance, BSA and HSA have a 24% difference in sequence.³¹ Since albumins have different sequences and thus can exhibit different binding properties, multiple species' serum albumins can be employed in a differential array to obtain cross-reactive interactions with analytes.

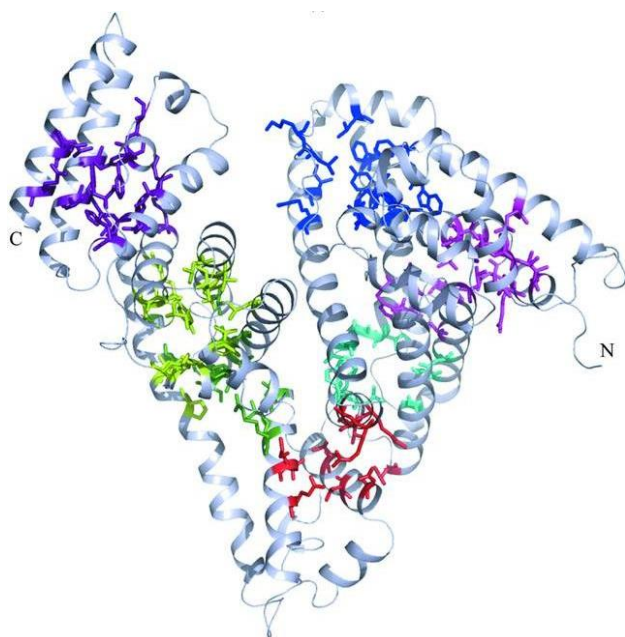


Figure 2.2 *Crystal structure of bovine serum albumin.* The amino acids in the fatty acid binding pockets are highlighted. (Reproduced from Ref. 32: © Wiley-VCH, 2012)

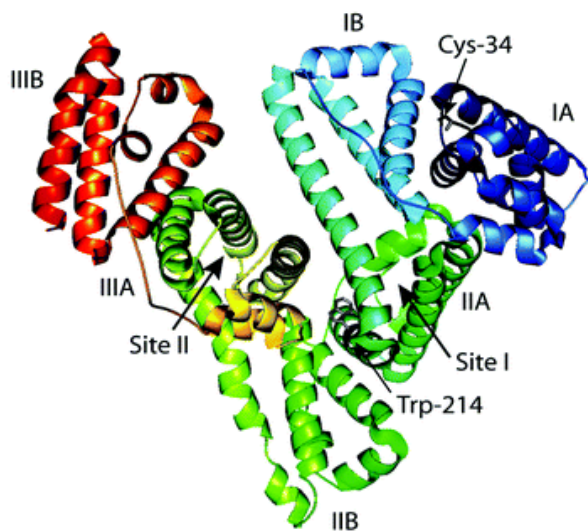


Figure 2.3 *Crystal structure of human serum albumin.* The subdomains and drug binding sites (I and II) are labeled.³⁶ (Reproduced from Ref. 36: © Royal Society of Chemistry, 2009)

2.3.2 Ligand Binding

Serum albumins bind a number of endogenous compounds, including long-chain fatty acids ($K_a = 10^6$ - 10^7 M⁻¹),³⁷ bile acids ($K_a = 10^3$ - 10^5 M⁻¹),³⁸ steroids ($K_a = 10^3$ - 10^5 M⁻¹),³⁹⁻⁴¹ and bilirubin ($K_a = 10^7$ M⁻¹).⁴² In addition, a variety of exogenous fluorophores and drugs are known to bind to the protein. These types of ligands typically bind in Sudlow Site I or II, also known as drug binding sites I and II.^{43,44} Sudlow Site I is located in the subdomain IIA, and Sudlow Site II is located in subdomain IIIA (**Figure 2.3**). Some ligands show clear specificity for one site over the other, and some ligands bind in both to some extent.^{31,45} In general, Site I ligands are heterocyclic anions with centrally-distributed charge, while Site II ligands are usually aromatic and neutral or anionic with a peripherally-located charge. Examples of Site I ligands include warfarin, digitoxin, salicylate, bromocresol green, methyl orange, and dansyl amide. Examples of Site II ligands include diazepam, ibuprofen, naproxen, medium-chain fatty acids, and dansyl proline. Despite being composed of fatty acid alkyl groups, glycerides bind less tightly to SAs and in a different location than their fatty acid counterparts.⁴⁶

2.4 PREVIOUS WORK USING SERUM ALBUMIN AS A CROSS-REACTIVE RECEPTOR

Our group has used arrays of serum albumins (SAs) for the differentiation of hydrophobic analytes including fatty acids,⁴⁷ terpenes,⁴⁸ and plasticizers.⁴⁹ In each study, two to three different serum albumins, BSA, HSA, and rabbit serum albumin (RSA), were used in conjunction with fluorescent indicators such as fluorescein, PRODAN, 2-anthracene carboxylate (2-AC), and dansyl amide (DNSA) to build an array. In the terpene and plasticizer studies, the use of non-fluorescent additives like deoxycholate and ascorbic acid in the array was investigated to determine if the additive could improve discrimination of the analytes (**Figure 2.4**). It was determined that such dopants can improve the cross-reactivity of the serum albumin/fluorophore receptors and aid in distinguishing structurally

similar analytes. In each case, complex mixtures that contain the targeted analytes were also tested with the arrays. For the fatty acid array, samples of edible vegetable oils were discriminated. The terpene array was used to discriminate sample of Masaki perfume that had been spiked with various terpenes. Finally, the plasticizer array was tested with simulated plastic explosive mixtures that contained RDX and PETN simulants mixed with the plasticizers. The array was even able to differentiate the explosive mixtures after they had been contaminated with soil (**Figure 2.5**).

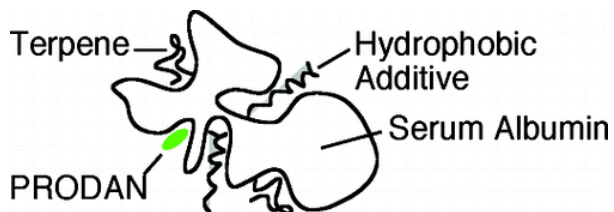


Figure 2.4 *Schematic of the sensing ensemble of serum albumin, PRODAN, and a hydrophobic additive in the presence of terpene.*

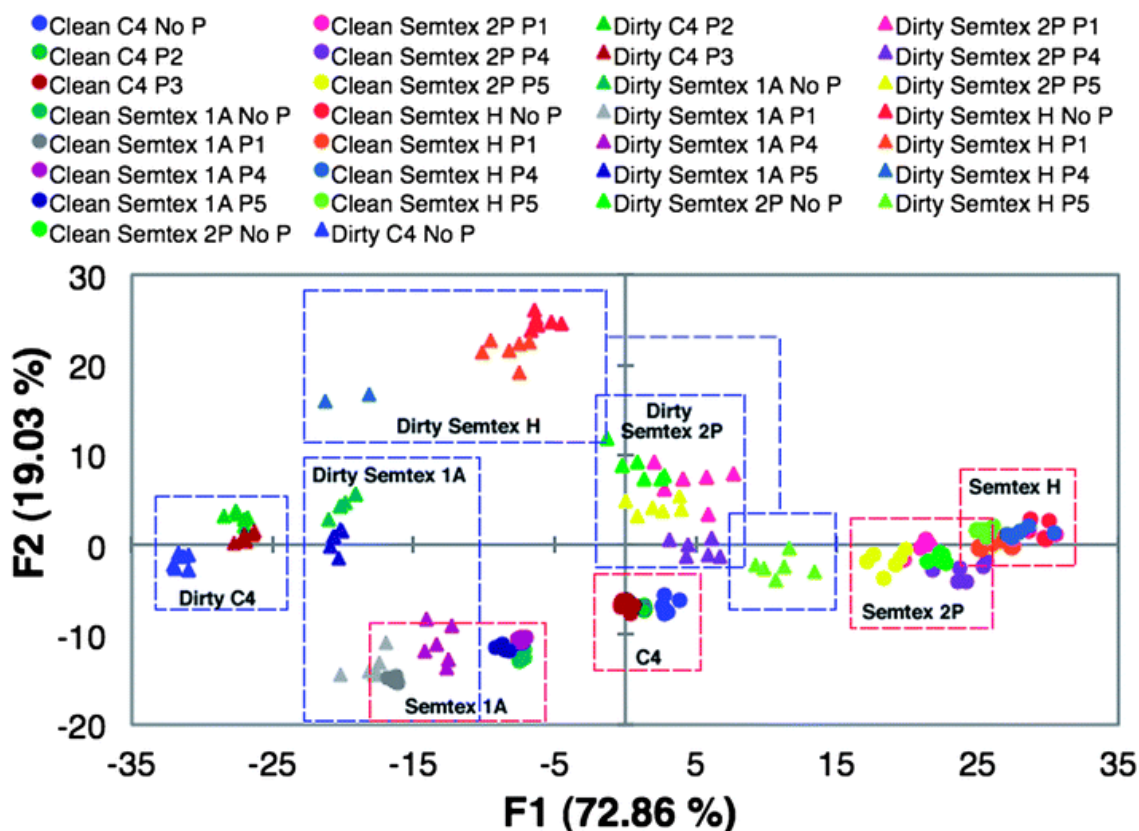


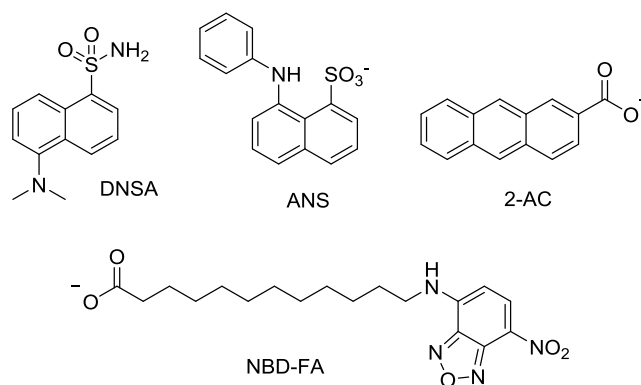
Figure 2.5 *Discrimination of contaminated explosive mixtures by the serum albumin array.* LDA plot generated from the array data with a cross validation of 96.5%. P1-5 are the different plasticizers. Semtex 1A, H, and 2P are different compositions of Semtex. (Reproduced from Ref. 49: © Royal Society of Chemistry, 2012).

Nevertheless, these previous studies did not involve differences between the analyte structures as subtle as with glycerides, nor had we challenged our methods to identify structural aspects of an unknown. We had also never implemented a quantitation assay in a complex mixture. Nevertheless, because the binding of ligands to SAs is known to depend on subtle differences in their structure,³¹ we anticipated that success could be achieved with glyceride discrimination but would be highly dependent upon the signaling modality and potentially also on analyte prederivitization.

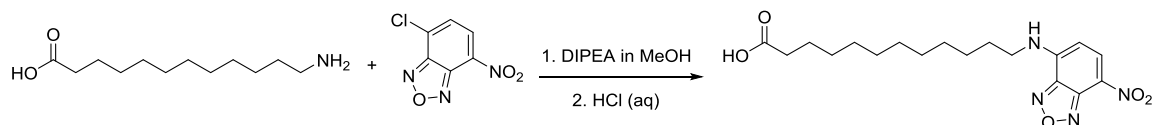
2.5 DEVELOPMENT OF AN ARRAY FOR GLYCERIDES

2.5.1 Indicator Uptake Experiments with Serum Albumin and Triglyceride

The optical signaling approach we envisioned required that fluorophores bind to the SAs, and signal the addition of glycerides upon binding. In order to identify appropriate fluorophores for the sensing ensemble, several candidates were screened for their fluorescence modulation both in the presence of bovine serum albumin and triglycerides. The triglycerides are insoluble in water but dissolve in the presence of serum albumin, presumably due to binding in the hydrophobic sites of the protein. Indicator uptake experiments were undertaken in which a serum albumin/triglyceride solution was titrated with a fluorescent indicator.⁵⁰ The change in emission of the indicator in the presence of the triglyceride was compared with the indicator's change in emission when it was added to serum albumin alone. This protocol revealed which fluorophores bind the SAs and modulate in the presence of glycerides. The fluorophores screened (**Scheme 2.5**) included dansyl amide (DNSA), 2-anthracenecarboxylate (2-AC), 1-anilinonaphthalene sulfonic acid (ANS), and a nitrobenzoxadiazole fatty acid (NBD-FA). The first three compounds were known previously to bind to serum albumin.³¹ NBD-FA was synthesized by a literature procedure (**Scheme 2.6**),⁵¹ and the binding of this fluorophore to SA was studied.⁵⁰



Scheme 2.5 Structures of the fluorophores screened for use in the array.



Scheme 2.6 Synthesis of NBD-FA.

An example of these indicator uptake experiments is shown in **Figure 2.6** for NBD-FA, while the titrations for the other indicators can be found in reference 50. When NBD-FA is added to serum albumin, its emission increases (“BSA” in **Figure 2.6**); however, when trimyrustin (**T6**) is present with the BSA, this increase in emission is smaller (“BSA/**T6**” in **Figure 2.6**), indicating that the binding of **T6** to serum albumin is inhibiting the binding of NBD-FA. It should be noted that the decrease in emission at concentrations above 100 μM NBD-FA is attributed to self-quenching as multiple NBD-FAs bind to the SA.

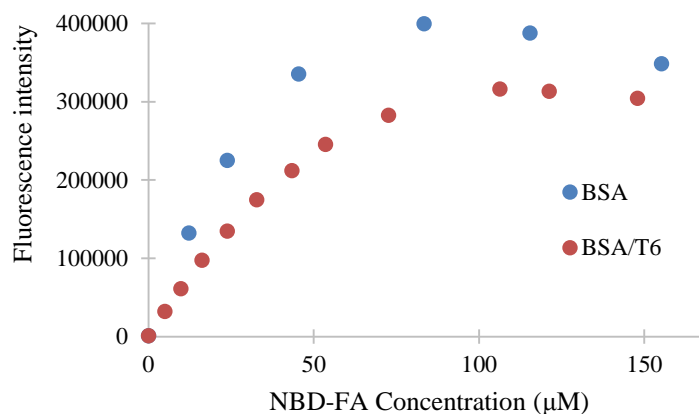


Figure 2.6 *NBD-FA uptake.* Addition of NBD-FA (0-155 μM) to BSA (100 μM) and to BSA (100 μM)/T6, (90 μM) in 10 mM phosphate buffer, pH 7.0, 0.02% NaN_3 , alone; < 0.3% THF; $\lambda_{\text{ex}} = 470 \text{ nm}$, $\lambda_{\text{em}} = 540 \text{ nm}$. (Reproduced from Ref. 50: © University of Texas at Austin, 2013)

From these experiments, DNSA and NBD-FA were inhibited in their binding to serum albumin in the presence of triglycerides and therefore were suitable for use in the array. The other indicators, 2-AC and ANS, did not demonstrate significant inhibition in binding to serum albumin in the presence of trimyristin. However, we chose to include ANS in the cross-reactive array because it is commonly used to probe the folding of SAs since it binds in the crevices between domains,³¹ and we anticipated it would respond to other glycerides even if our initial tests with trimyristin did not reveal emission modulations.

2.5.2 First Iteration of the Well Plate Array

The cross-reactive array was constructed using 96-well plates. As a preliminary test of the array's ability to differentiate the glycerides, an assortment of eight mono-, di-, and triglycerides from the total panel were selected. Each glyceride was mixed with individual combinations of the three indicators (DNSA, NBD-FA, and ANS) and bovine and human

serum albumin (six combinations) in a 96-well plate with eight replicates for each glyceride, and the emission of each well was recorded.

An LDA score plot of the fluorescence data collected from this array is shown in **Figure 2.7**, which has a 98% cross-validation analysis. While the eight glycerides were differentiated by this array with high accuracy, there was significant visual overlap along the F1 and F2 axes of three of the glycerides. **D1** and **D2**, which are *cis/trans* isomers of one another, were not well discriminated, and **T5** was also poorly separated. The fact that our approach demonstrated overlap of unsaturated glycerides with only eight of the twenty total targets in the panel caused us to reconsider the approach. Because differences between the unsaturated glycerides were the hardest to discriminate, we anticipated even further problems when attempting to classify the position of the double bonds, their stereochemistry, and the numbers of double bonds.

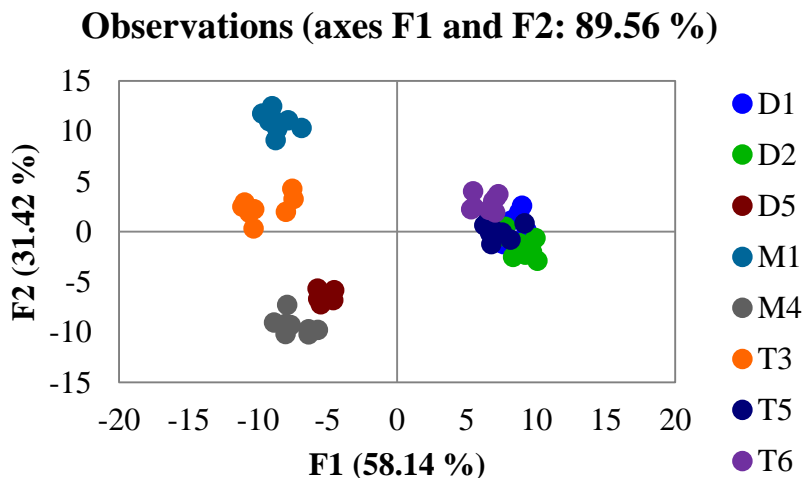


Figure 2.7 *Discrimination of eight glycerides by the six-receptor array.* LDA plots of data collected from 96-well plates: BSA and HSA (100 μM), glyceride (90 μM), DNSA (60 μM), ANS (60 μM), and NBD-FA (60 μM) in phosphate buffer with < 2% THF (see Table 2.7 for read parameters). Cross validation = 98%.

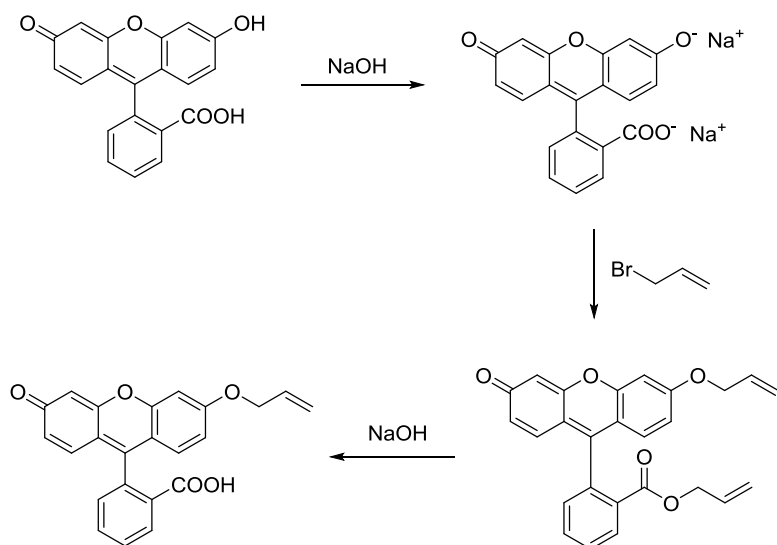
2.6 OLEFIN METATHESIS OF GLYCERIDES

One strategy to discriminate the differences in the olefins would be through derivitization. We contemplated the use of bromination and dihydroxylation, but ultimately chose olefin metathesis. Metathesis reactions of glycerides and fatty acids have been previously explored with the intent of using these compounds to make chemical products more sustainably.^{52,53} Such a prior derivitization approach is similar, in principle, to the use of other olefin reactions, such as halogen addition¹⁸ and silver cationization, that are able to resolve stereo- and regioisomers of glycerides when combined with NARP-HPLC-MS methods.²⁰

Olefin metathesis was chosen for several reasons. First, using a metathesis reaction results in products of differing length depending upon the positions of the olefins in the glycerides, thus potentially making the products from such glycerides unique. Second, olefin metathesis catalysts have differing reactivity toward *cis* and *trans* stereoisomers,⁵⁴ thus potentially leading to different extents of metathesis depending upon the stereochemistry of the starting fatty acid chains. Third, the reaction conditions are relatively mild, and the reaction mixture can be used directly in the SA array without any purification. This factor allows the cross metathesis reactions of multiple glycerides to be performed in parallel in a polypropylene well-plate for efficient workflow. Lastly, this reaction could be used to introduce an additional fluorophore for optical analysis. Our strategy therefore utilized fluorescein conjugated to an olefin, resulting in mixed olefin products.

2.6.1 Synthesis of Allyl Fluorescein

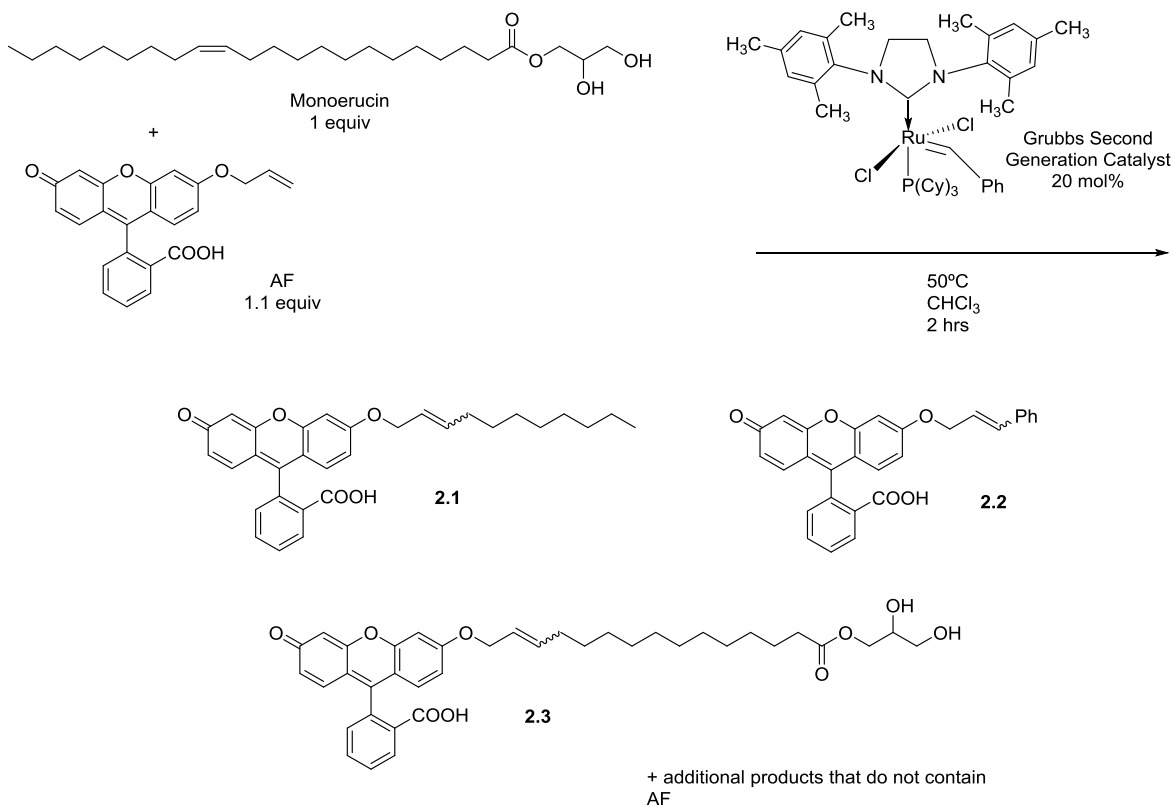
With these goals in mind, the allyl fluorescein derivative **AF** was synthesized according to a literature procedure (**Scheme 2.7**).^{55,56}



Scheme 2.7 *Synthesis of allyl fluorescein (AF).*

2.6.2 Optimizing the Metathesis Reaction with Monoerucin

As a model, we then screened several different reaction conditions for cross metathesis between AF and monoerucin (**M1**) (**Scheme 2.8**) in order to optimize the reaction conditions. LC-MS analysis of the reaction mixtures confirmed the conversion of AF to mixed glyceride- and AF-containing compounds.



Scheme 2.8 *Olefin metathesis of monoerucin.*

The conditions that were screened are summarized in **Table 2.1**. The application of vacuum to the reaction was found to be unnecessary (reaction #1), so the reactions were run at atmospheric pressure to simplify the reaction setup. Although the reaction also worked in dichloroethane (DCE), chloroform (CHCl₃) was preferred as the solvent because it is non-flammable and hence can be heated more safely in a plastic well plate. The reaction did not proceed when tetrahydrofuran (THF) was used as the solvent. The presence of even very small amounts (<10%) of THF appeared to hinder or entirely halt the reaction (e.g. reaction #12). The addition of copper⁵⁷ provided no significant improvement in the reaction, so it was ultimately excluded.

The origin of reaction #2's failure is unclear. THF was not used as the solvent during the reaction, but the reagents were first dissolved in THF, and then the THF was removed by rotary evaporation. Perhaps the presence of a trace amount of THF was enough to affect the reaction. Using more **M1** compared to AF did improve the yield of the **M1**-AF mixed products, but we did not want to use more material than necessary, so an excess of glyceride was not used in the array. Pre-reacting **M1** with the catalyst before adding AF hindered the formation of any **M1**-AF mixed products (reaction #11). Moreover, inconsistent results were obtained when repeating the metathesis reaction with an old batch of catalyst. New batches of catalyst purchased from Sigma gave reproducible results. The conditions that were selected to be optimal for our purposes are shown in **Scheme 2.8**.

Table 2.1 *Reaction conditions for the metathesis reaction and AF and M1.*

	Solvent	Temperature (°C)	Reaction time	M1 (equiv)	AF (equiv)	G2*** (mol%)	CuI (mol%)	Reaction success
1*	CHCl ₃	55, vacuum	24 hr	1.1	1.0	0.1	NA	Yes
2	CHCl ₃	55	24 hr	1.1	1.0	0.1	NA	No
3	THF	50	24 hr	1.1	1.0	0.1	NA	No
4	DCE	70	2 hr	1.0	15	1.5	NA	Yes
5	DCE	75	24 hr	1.0	15	2.0	3.0	Yes
6	DCE	75	24 hr	1.0	5.0	20	25	Yes
7	CHCl ₃	50	2 hr	1.0	2.0	20	25	Yes
8	CHCl ₃	50	2 hr	1.0	1.1	20	30	Yes
9	CHCl ₃	50	6 hr	2.0	1.1	50**	NA	Yes
10	CHCl ₃	50	2 hrs	3.0	1.0	20**	NA	Yes
11	CHCl ₃	50	24 hrs	1.0	1.1, added after 2 hrs	20**	NA	No
12	CHCl ₃	50	2 hr	1.0, prepared from a THF stock of M1	1.1	20**	NA	No
13	CHCl ₃	50	2 hr	1.0	1.1	20**	NA	Yes

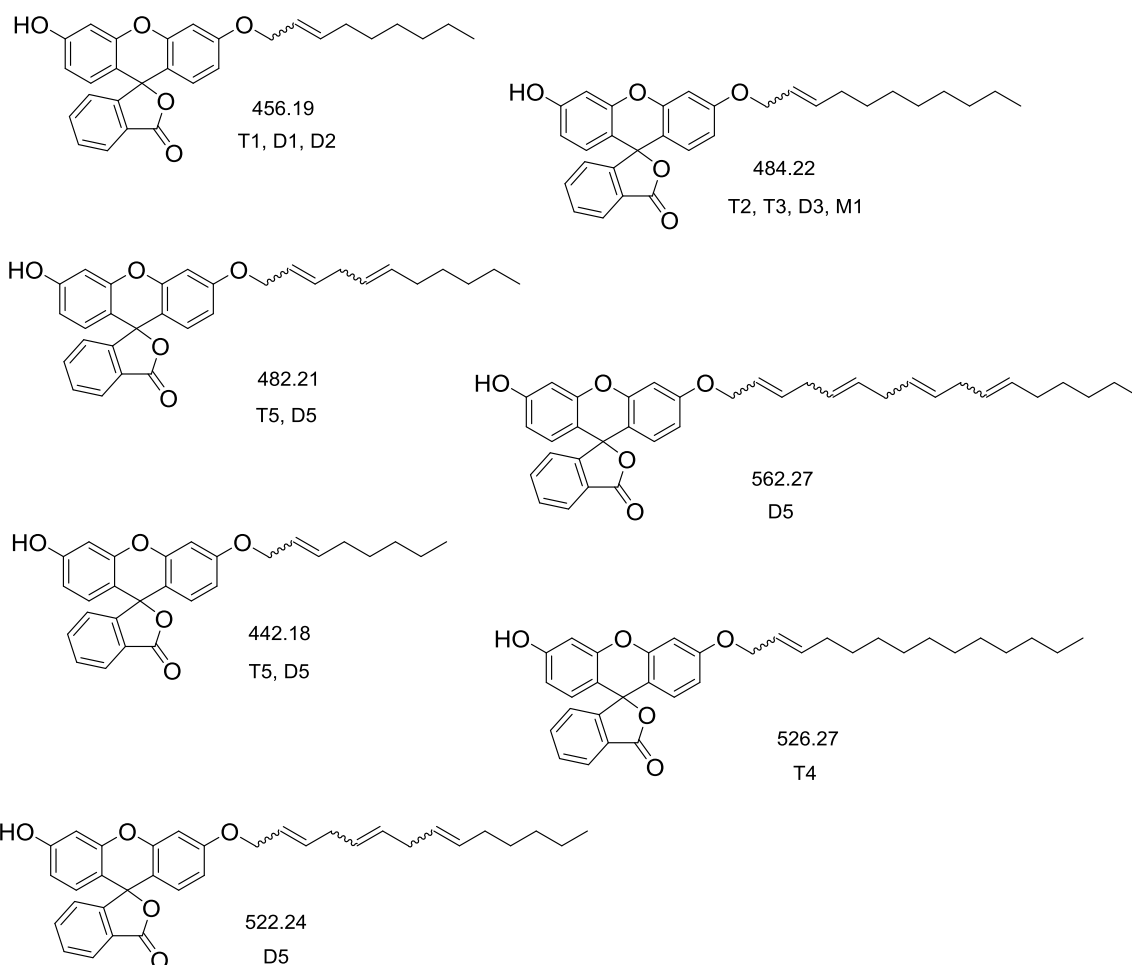
* Reactions 1-8 were performed by Michelle Adams Ivy and published in ref 50

**New batch of catalyst was used from Sigma

***G2 = Grubbs 2nd Generation Catalyst

2.6.3 Evaluating the Olefin Metathesis with All of the Unsaturated Glycerides in the Panel

Once the optimal conditions were identified for **M1**, all of the unsaturated glycerides in the panel were subjected to the cross metathesis reaction in parallel in a well plate, and the reaction mixtures were analyzed by LC-MS to determine that AF was metathesizing with the glycerides and that the conversion was the same for each repetition of the reaction with each glyceride. In these experiments, the AF-glyceride products observed in the LC-MS analysis contained the alkyl portion of the glyceride, as in **2.1** (**Schemes 2.8** and **2.9**), and smaller amounts of products such as **2.3**. **Table 2.2** summarizes the percent conversion of AF to the corresponding mixed product (**2.1** type) for each unsaturated glyceride. The chromatogram for absorption at 280 nm was used to quantitate the conversion, and we assumed that the extinction coefficient at that wavelength for AF and AF-glyceride species did not differ significantly. In each reaction, some unreacted AF remained, and some of product **2.2**, which results from the catalyst loading used, was also observed. Finally, it is highly likely that olefin combinations that did not involve AF (i.e. self-metathesis of the glyceride) are also generated in the reaction. However, these species were not observable using conventional LC-MS analysis because they neither absorb light at any unique wavelength nor do they ionize readily. Although the conversion differed for each glyceride, all of the glycerides were found to metathesize with AF.



Scheme 2.9 *Metathesis products detected by mass spectrometry from reaction of AF with each of the unsaturated glycerides. The exact mass of each compound is given along with the corresponding glyceride from which each product originates.*

Table 2.2 Summary of the olefin metathesis reaction for the unsaturated glycerides and AF.

Glyceride	<i>Cis</i> or <i>Trans</i> *	m/z (M+1) of mixed metathesis product	Percent conversion of AF to mixed product (%)
T1- Tripalmitolein	<i>cis</i>	457.2	68
T2- 1,3-Dioleoyl-2-palmitoyl glycerol	<i>cis</i>	485.2	52
T3- Trielaidin	<i>trans</i>	485.2	65
T4- Tripetroselaidin	<i>trans</i>	527.3	54
T5- Trilinolein	<i>cis</i>	443.2	49
	<i>cis</i>	483.2	23
D1- 1,3-Dipalmitelaidin	<i>trans</i>	457.2	55
D2- 1,3-Dipalmitolein	<i>cis</i>	457.2	56
D3- 1,2-Diolein	<i>cis</i>	485.2	59
D5- 1,3-Diarachidonin	<i>cis</i>	443.2	33
	<i>cis</i>	483.2	16
	<i>cis</i>	523.2	9
	<i>cis</i>	563.3	10
M1- Monoerucin	<i>cis</i>	485.2	8

*Configuration of the double bond in the original glyceride

2.6.4 Binding Studies of AF and 2.1 to Serum Albumin

In order to investigate the binding of **2.1** to serum albumin, the compound was isolated by HPLC to obtain a pure sample. BSA was titrated into AF and **2.1**, separately.⁵⁰ AF is fluorescent alone in solution, and this fluorescence is quenched in the presence of serum albumin. On the other hand, **2.1** has relatively lower emission alone but increases in its emission upon addition of serum albumin. The lower emission in solution indicates that the alkyl tail quenches the fluorescence of the fluorescein intramolecularly.

2.6.5 Exploring FRET with Dansyl Amide and Allyl Fluorescein

Dansyl and fluorescein moieties have been reported to act as a Förster Resonance Energy Transfer (FRET) pair,^{58,59} and therefore we investigated this property with the metathesized glycerides. FRET would be advantageous to the chemometric patterning because it would add an additional facet of cross-reactivity to the array. To test whether FRET occurred between DNSA and **2.1** in the presence of BSA, we measured the emission spectra of solutions of DNSA/BSA, **2.1**/BSA, and DNSA/**2.1**/BSA by exciting at 350 nm (**Figure 2.8**). This data supports FRET because the DNSA/**2.1**/BSA sample exhibits emission at the λ_{max} of 520 nm for fluorescein that was much higher in intensity than the **2.1**/BSA sample. We also performed a titration in which DNSA was titrated into BSA/**2.1** while exciting at 350 nm (for dansyl) and observed an increase in emission of fluorescein with a λ_{max} at 520 nm,⁵⁰ which supports FRET between dansyl amide and **2.1** when they are both bound to BSA.

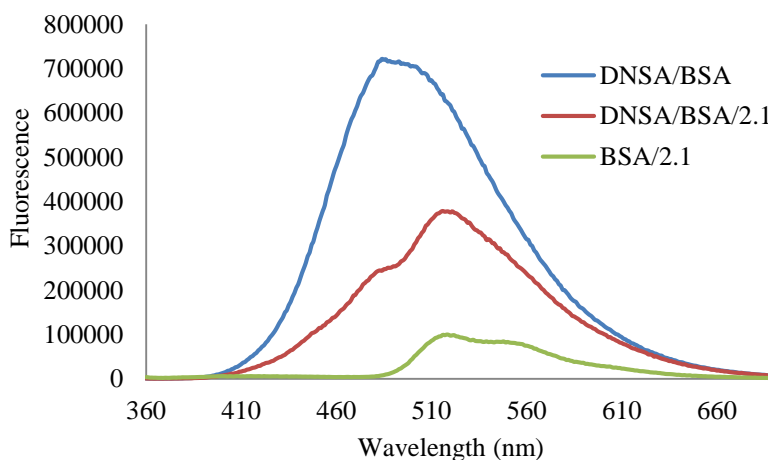


Figure 2.8 *Evidence of FRET between DNSA and 2.1.* Fluorescence spectra for DNSA/BSA (250 μM /100 μM), **2.1**/BSA (90 μM /100 μM), and DNSA/**2.1**/BSA (250 μM / 90 μM / 100 μM) in 10 mM phosphate buffer, pH 7.00, 0.02% NaN_3 ; $\lambda_{\text{ex}} = 350$ nm. (Reproduced from Ref. 50: © University of Texas at Austin, 2013)

However, we knew from the analysis of the metathesis reaction mixtures that a significant amount of AF remained in the solution. AF has a broad excitation spectrum compared to **2.1** (**Figure 2.39**) and can be excited directly at 350 nm. Furthermore, we titrated DNSA into AF/BSA and observed a small decrease in emission (**Figure 2.40**), which is not indicative of FRET. Therefore, we expected to observe a mixture of these two effects for the FRET parameter in the array, since both AF and species like **2.1** could both be present in the well.

2.6.6 Second Iteration of the Well-Plate Array

Once we had established that the olefin metathesis reaction was successfully mixing AF with fragments of the unsaturated glycerides in our panel, we were ready to incorporate this reaction into the well plate array with SA. To do so, we performed the metathesis reaction on the glycerides in a well plate in chloroform. However, because chloroform is immiscible with buffered SA solutions, after metathesis the chloroform was allowed to evaporate and the residues were taken up in THF. The THF solutions of the glyceride metathesis reactions were then exposed to BSA and HSA, separately, in a 96-well plate. Dansyl amide (DNSA) was added to one set of BSA and HSA plates and buffer was added to another set of BSA and HSA plates. For the plates that contained DNSA, the emission of fluorescein, DNSA, and the FRET pair between the two fluorophores was measured. For the plates without DNSA, only the emission of the fluorescein was measured. We wanted to measure the fluorescein emission both in the absence and presence of DNSA to see if there was a significant difference in this signal and thus an additional variable in our array. The first iteration of the well plate array was also performed exactly as described above for each glyceride. Both sets of fluorescence data were used in the pattern recognition algorithm for a total of fourteen variables.

The LDA plot generated for eight glycerides is shown in **Figure 2.9**. The cross-validation was 100%. The clustering within replicates of the same glyceride in this plot (B) is much tighter than in the plot generated using only the first iteration of the well plate array (A), and the separation between different glycerides is more marked. In particular, compared to **Figure 2.7**, the *cis* isomer **D2** is discriminated from its *trans* isomer **D1**, and **T5** no longer overlaps with **D1** and **D2**.

From the factor-loading plot (**Figure 2.10**), we can conclude that the metathesis parameters contributed significantly to the discrimination. A factor-loading plot shows the contribution of each of the original input variables to each factor axis in the reduced variable space. The loading plot in **Figure 2.10** shows that the metathesis variables, which are marked in red, contribute significantly to F1 and F2. This result supports our hypothesis that the metathesis reaction would improve the differentiation of a panel of glycerides containing unsaturated species.

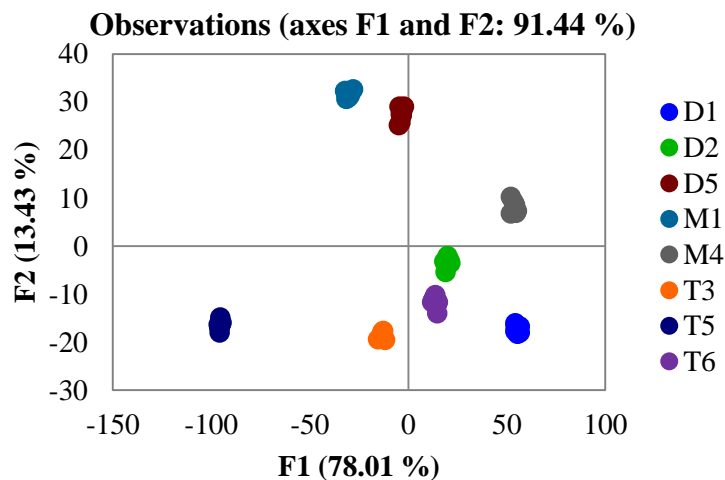


Figure 2.9 *Discrimination of eight glycerides by the fourteen-receptor array.* LDA plots of data collected from 96-well plates with olefin metathesis. BSA and HSA (100 μM), glyceride (90 μM), DNSA (60 μM), ANS (60 μM), and NBD-FA (60 μM), metathesized glyceride (90 μM), AF (100 μM), and DNSA (60 μM) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters). Cross validation = 100%.

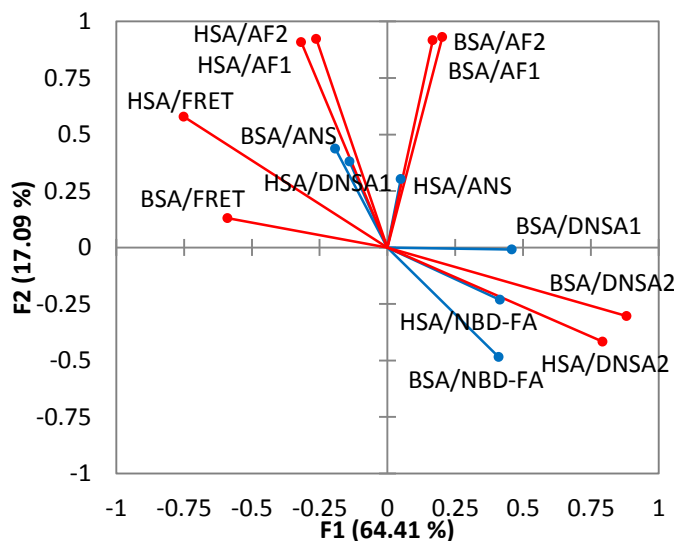


Figure 2.10 *Loading plot corresponding to the LDA in Figure 2.9.* DNSA1 refers to the non-metathesis part of the assay, while DNSA2 refers to the measurement of the DNSA emission in the presence of the metathesized glycerides. AF1 refers to the fluorescein emission in the absence of any DNSA, while AF2 refers to the fluorescein emission in the presence of DNSA.

2.7 ARRAY REPRODUCIBILITY

Next, we wanted to be sure that we could replicate the results of array. In other words, we wanted to know if we would get essentially the same pattern for the glycerides each time we performed an independent repetition of the entire experiment, including repeating the metathesis reaction and starting from newly prepared stock solutions of all of the array components. Such a complete reproduction has rarely, if ever, been discussed in papers describing differential sensing routines. Hence, we performed an independent repetition of the array on those same eight glycerides two more times for a total of three data sets (**Figure 2.11**). The relative position of all eight glycerides was essentially the same for each plot, and we concluded that the array is reproducible.

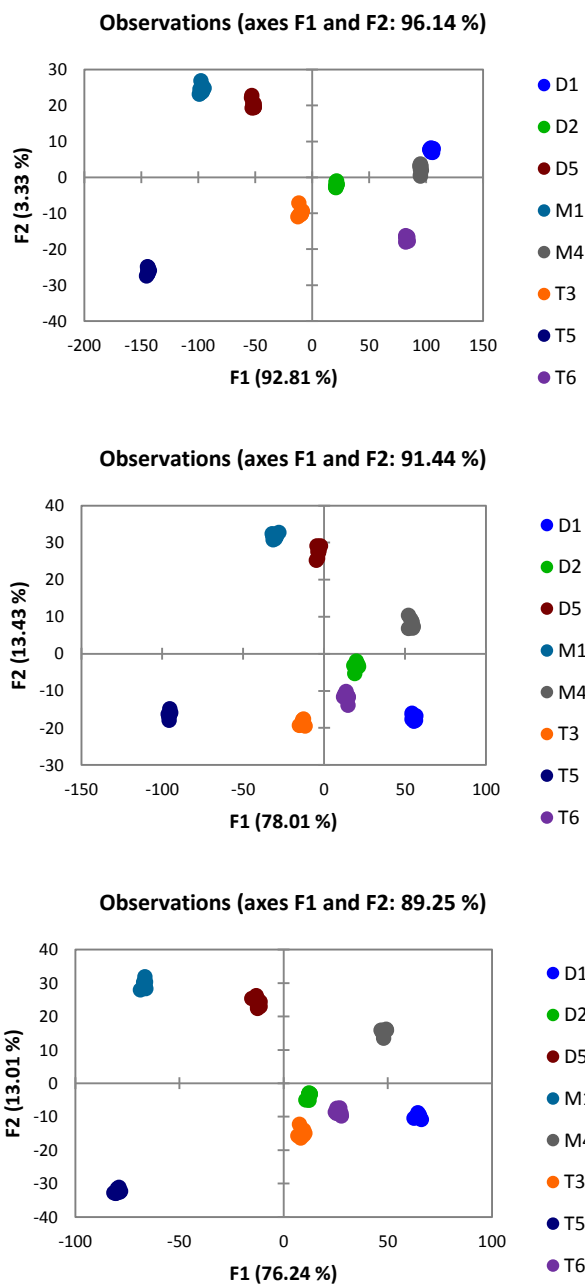


Figure 2.11 LDA plots of data collected from 96-well plates in three independent repetitions. BSA and HSA (100 μM), glyceride (90 μM), DNSA (60 μM), ANS (60 μM), and NBD-FA (60 μM), metathesized glyceride (90 μM), AF (100 μM), and DNSA (60 μM) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters). The fluorescence counts were express as a percentage of the total fluorescence for each receptor/variable.

In order to further test the consistency of the array, we used LDA as a predictive tool. Since LDA entails that the classes be given as inputs, one can take two sets of data, in our case the independent repetitions of the array on those eight glycerides, and assign one as a training set and the other as a prediction set. The training set teaches the algorithm which responses from the array correspond to which glyceride, and then the algorithm can assign the glyceride identity to the sets of array responses in the prediction set. Hence, we alternately treated each of the three repetitions as the training set or the prediction set. For 64 data points in each set (8 replicates of 8 glycerides), the average accuracy was 87% for the six combinations of the three independent replications (**Table 2.3**).

Table 2.3 *Summary of prediction accuracy for the six combinations of the three repetitions in Figure 2.11.*

Training Set	Prediction Set	% Correct Classification (of 64 cases)
Rep 1	Rep 2	75
Rep 1	Rep 3	84
Rep 2	Rep 1	80
Rep 2	Rep 3	100
Rep 3	Rep 1	84
Rep3	Rep 2	100

2.8 ANALYSIS OF THE FULL GLYCERIDE PANEL

Finally, we performed the full fourteen-variable array on all twenty glycerides in the panel. An LDA plot was obtained with a cross-validation of 98% (**Figure 2.12**). The only error in classification in the cross validation was between one replicate of the saturated triglyceride **T7**, which was misclassified as another saturated triglyceride **T8**. In general, the unsaturated di- and triglycerides are on the right side of the plot, while the saturated glycerides are on the left side. Within the unsaturated glycerides, there is very clear separation between all of the mono-, di-, and triglycerides. The unsaturated monoglyceride

M1 is associated more closely with the saturated glycerides on the left side of the plot. However, this result is reasonable because **M1** metathesizes poorly (see **Table 2.2**). The two glycerides that consist of fatty acid alkyl groups with multiple double bonds (**T5**, **D4**) are found near one another at the bottom center of the plot.

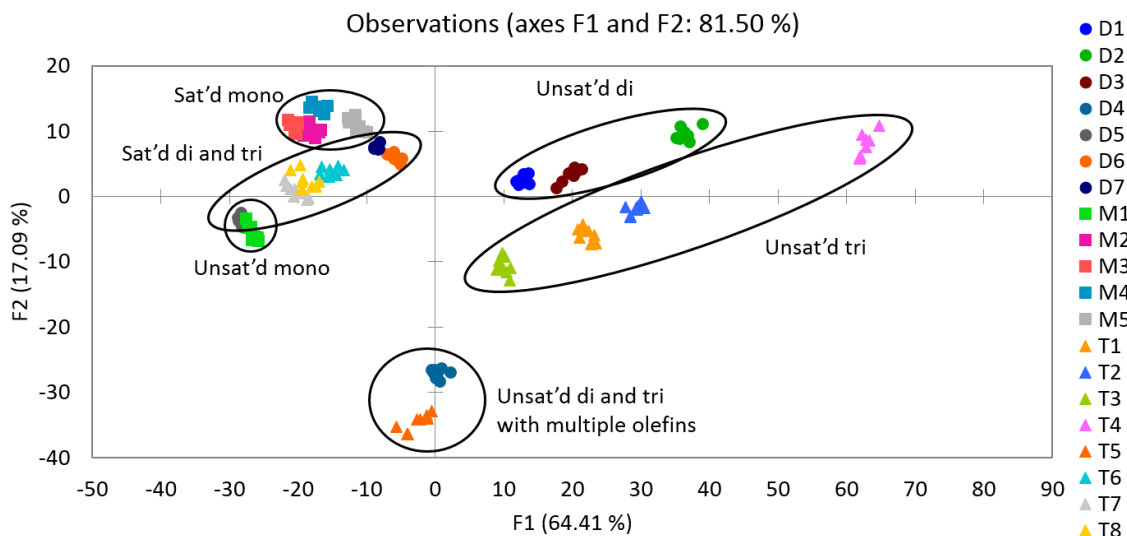


Figure 2.12 *Discrimination of twenty glycerides by the array.* LDA plot of data collected from 96-well plates. The array components consisted of BSA and HSA (100 μ M), glyceride (90 μ M), DNSA (60 μ M), ANS (60 μ M), NBD-FA (60 μ M), metathesized glyceride (90 μ M), AF (100 μ M), and DNSA (60 μ M) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters). Cross validation: 98%.

Moreover, the stereo- and regioisomers of the unsaturated di- and triglycerides are successfully distinguished by the array. Within the unsaturated diglycerides, the *cis/trans* isomers **D1** and **D2** are discriminated from one another. Within the unsaturated triglycerides, the olefin positional isomers **T3** (olefin between carbons 9 and 10) and **T4** (olefin between carbons 6 and 7) are also clearly separated. Furthermore, **T2** contains fatty

acid alkyl groups that are stereoisomeric (*cis*) to the fatty acid alkyl groups in **T3** (*trans*), and these species are also clearly differentiated from one another.

All the saturated glycerides are also well discriminated, although the separation is less marked than between the unsaturated glycerides. This finding is unsurprising since the metathesis reaction does not occur with the saturated glycerides, and hence their discrimination would not be expected to be as significantly improved by the pretreatment step. Within the saturated mono-, di-, and triglycerides, the difference between the compounds is simply the number of carbons in the fatty acid alkyl chains; however, the panel does contain one regioisomer pair, **D5** and **D6**. **D5** is esterified at the *sn*-1 and *sn*-3 positions of glycerol, while **D6** is esterified at the *sn*-1 and *sn*-2 positions of glycerol. The array is able to differentiate these regioisomers as well.

The PCA plot (**Figure 2.13**) also shows good discrimination of the glycerides and similar groupings of glycerides based on their structural features to the LDA plot. In PCA, the identities of the glycerides are not given as inputs, so the analysis only takes the variance in the variables (i.e. emission signals for each receptor) into account without knowing what the classes (i.e. identities of the glycerides) are supposed to be. It is apparent in the PCA plot that the serum albumin-based receptors respond to the glycerides in a differential manner. In other words, the receptors behave differently from one another with each analyte. When receptors in an array behave similarly to one another, the PCA shows low dimensionality (i.e. most of the variance is described by one or two factor axes). For example, if the receptors all display their highest emission values in the presence of analyte 1 and the lowest emission values in the presence of analyte 2, the PCA would be able to describe the bulk of the variance in the data along a single factor axis. In our case, the first five factor axes (F1-F5) describe 90% of the variance in the data. The high dimensionality in the data is indicative of receptors that respond differently from one

another to the analytes in the panel. High dimensionality in the data is indicative of a truly cross-reactive array of receptors.

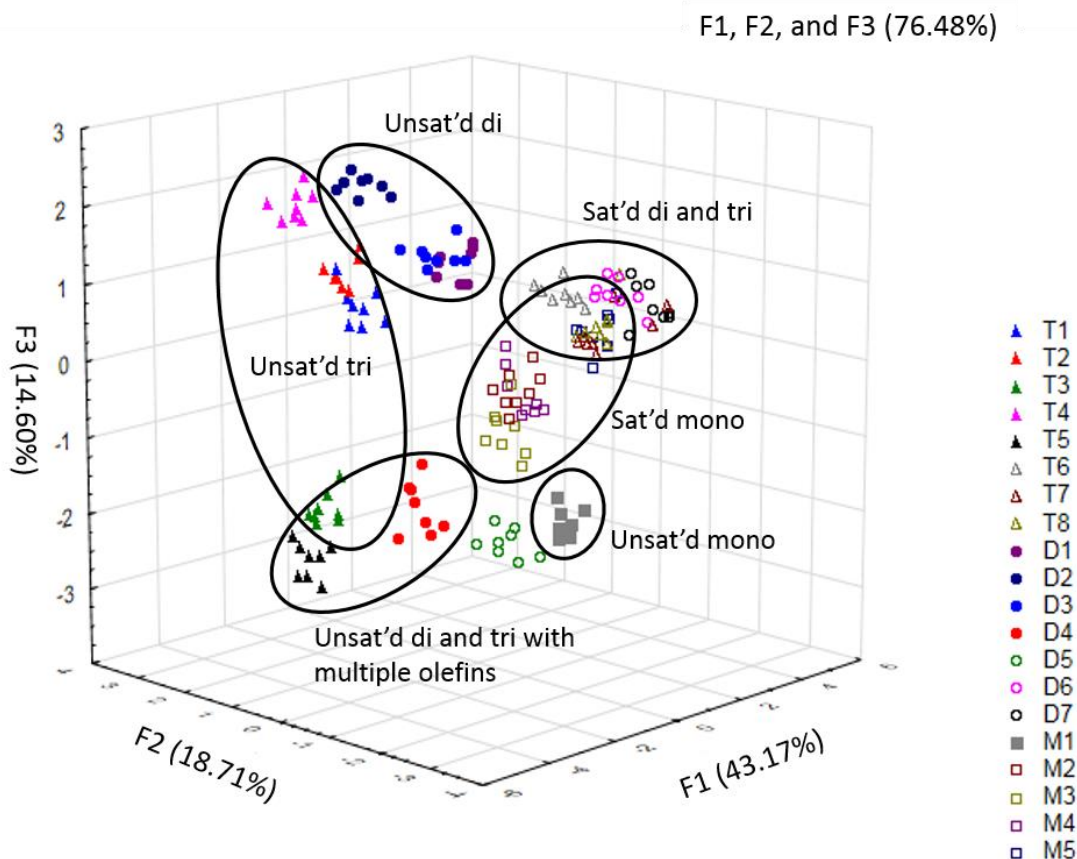


Figure 2.13 *Discrimination of twenty glycerides using PCA.* PCA plot of data collected from 96-well plates. The array components consisted of BSA and HSA (100 μM), glyceride (90 μM), DNSA (60 μM), ANS (60 μM), NBD-FA (60 μM), metathesized glyceride (90 μM), AF (100 μM), and DNSA (60 μM) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters).

2.9 PREDICTION OF GLYCERIDE STRUCTURAL FEATURES

Since glycerides that share structural features are grouped in the LDA and PCA plots, we hypothesized that our array could be used to identify these features in unknown glycerides. As a way of testing this hypothesis, an independent reproduction of the array

was performed on all twenty glycerides (**Figure 2.14**), and the predictive feature of LDA was used. The data set from one experiment was used as the training set, and the other data set was used as the prediction set. The analyte classes that were provided to the algorithm were structural features of the glycerides: unsaturated triglyceride, saturated triglyceride, polyunsaturated triglyceride, unsaturated diglyceride, saturated diglyceride, polyunsaturated diglyceride, unsaturated monoglyceride, and saturated monoglyceride (**Figure 2.15**).

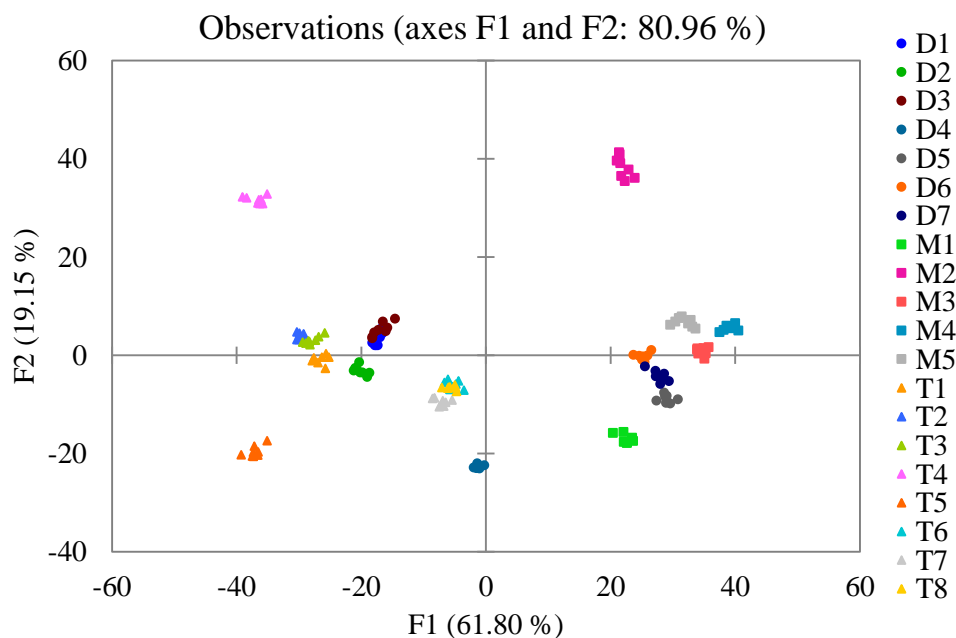


Figure 2.14 *Independent repetition of array discrimination of twenty glycerides.* LDA plot of data collected from 96-well plates in an independent repetition of the array on all twenty glycerides. The array components consisted of BSA and HSA (100 μ M), glyceride (90 μ M), DNSA (60 μ M), ANS (60 μ M), NBD-FA (60 μ M), metathesized glyceride (90 μ M), AF (100 μ M), and DNSA (60 μ M) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters). Cross validation: 98%

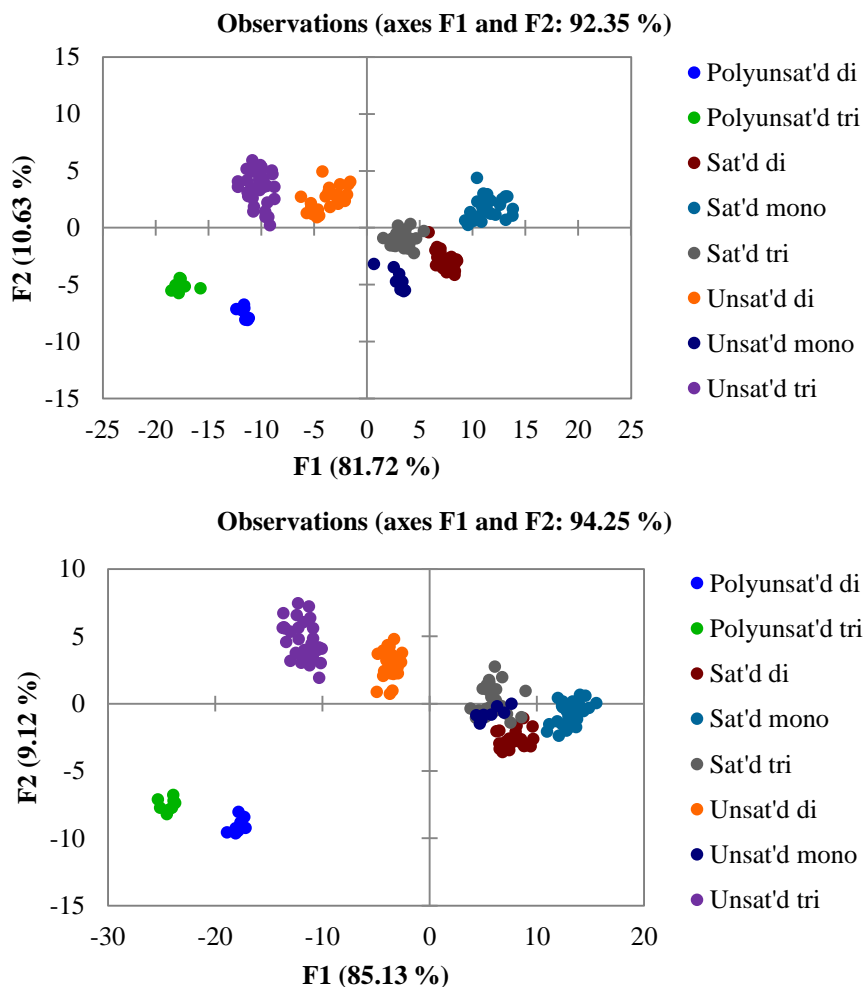


Figure 2.15 *LDA plots of training sets with glyceride classes designated as structural features.* The array components consisted of BSA and HSA (100 μM), glyceride (90 μM), DNSA (60 μM), ANS (60 μM), NBD-FA (60 μM), metathesized glyceride (90 μM), AF (100 μM), and DNSA (60 μM) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters). The fluorescence counts were expressed as a percentage of the total fluorescence for each receptor/variable. Cross validation: 100% (top) and 98% (bottom).

LDA correctly assigned the glycerides in the training set to its structural feature class in an average of 89% of 160 cases. Since our array is primarily designed to target differences between unsaturated glycerides through the metathesis pretreatment, we

postulated that separating the saturated and unsaturated glycerides from one another first would improve the accuracy of the predictions. The analyte classes were assigned more generally to either saturated, unsaturated, or polyunsaturated (**Figure 2.16**). LDA correctly assigned the glycerides in the training set according to these classes in an average of 96% of 160 cases.

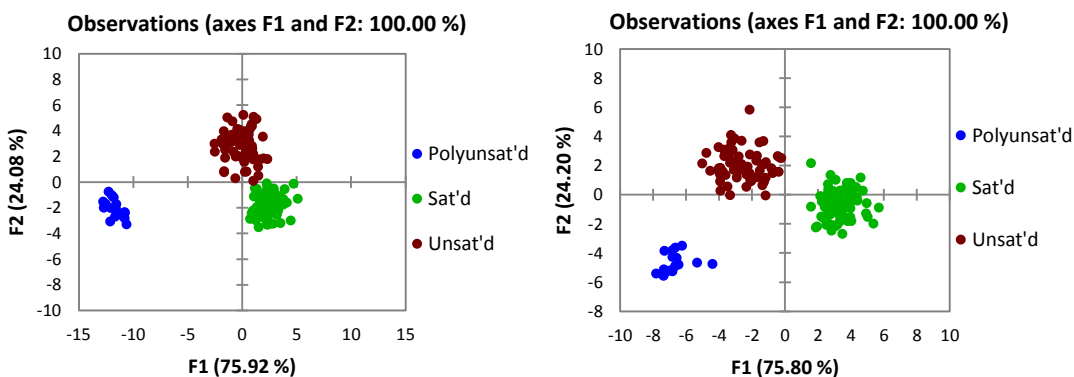


Figure 2.16 *LDA plots of the training sets with glyceride classes designated as saturated, unsaturated, and polyunsaturated.* The array components consisted of BSA and HSA (100 μ M), glyceride (90 μ M), DNSA (60 μ M), ANS (60 μ M), NBD-FA (60 μ M), metathesized glyceride (90 μ M), AF (100 μ M), and DNSA (60 μ M) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters). The fluorescence counts were expressed as a percentage of the total fluorescence for each receptor/variable. Cross validation: 99% (left) and 100% (right).

Once it was known whether the glycerides were saturated or unsaturated, then whether they were mono-, di-, or triglycerides could be determined. The training and prediction sets were split into saturated and unsaturated. Within the unsaturated set, the analyte classes were triglyceride, diglyceride, monoglyceride, polyunsaturated diglyceride, and polyunsaturated triglyceride (**Figure 2.17**).

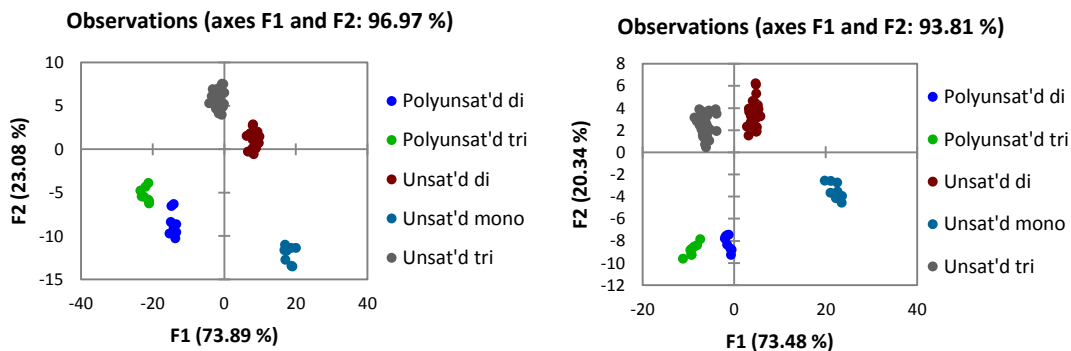


Figure 2.17 *LDA plots of the training sets of unsaturated glycerides with classes designated as shown in the key. The array components consisted of BSA and HSA (100 μ M), glyceride (90 μ M), DNSA (60 μ M), ANS (60 μ M), NBD-FA (60 μ M), metathesized glyceride (90 μ M), AF (100 μ M), and DNSA (60 μ M) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters). The fluorescence counts were expressed as a percentage of the total fluorescence for each receptor/variable. Cross validation: 100% (left) and 100% (right).*

With the saturated set, the analyte classes were triglyceride, diglyceride, and monoglyceride (**Figure 2.18**). For the unsaturated set, the glycerides were assigned correctly in an average of 95% of the 80 cases, while for the saturated set, the glycerides were assigned correctly in an average of 73% of the 80 cases. The lower accuracy of the prediction for the saturated glycerides is unsurprising since it is apparent in the LDA plot that the saturated glycerides are less clearly separated from one another compared to the unsaturated glycerides. However, overall, the array shows a good ability to predict the structural features of glycerides using LDA.

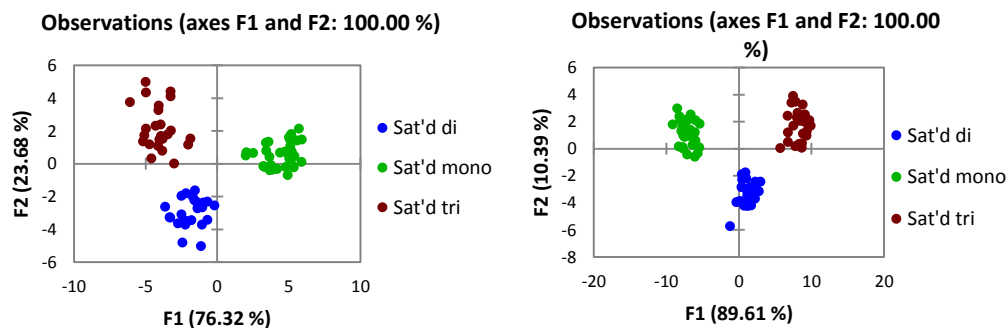


Figure 2.18 *LDA plots of the training sets of saturated glycerides with analyte classes designated as shown in the key. The array components consisted of BSA and HSA (100 μ M), glyceride (90 μ M), DNSA (60 μ M), ANS (60 μ M), NBD-FA (60 μ M), metathesized glyceride (90 μ M), AF (100 μ M), and DNSA (60 μ M) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters). The fluorescence counts were expressed as a percentage of the total fluorescence for each receptor/variable. Cross validation: 100% (left) and 100% (right).*

2.10 ADIPOCYTES

Adipocytes store dietary fat in the form of glycerides.¹ Triglycerides are ingested and converted into free fatty acids by pancreatic lipases. These free fatty acids are absorbed by the body and transported to adipocytes, hepatocytes, and muscle cells. The adipocytes reesterify the fatty acids to obtain triglycerides that remain in the cells until the glycerides are needed for energy production. When energy is needed, glycerides are mobilized by the process of lipolysis, in which lipases hydrolyze the glycerides to yield free fatty acids that can be catabolized (**Figure 2.19**).

There is a recent interest in investigating the lipid patterns of diabetic individuals compared to non-diabetic individuals as a potential diagnostic tool.⁶⁰⁻⁶³ Several enzymes, including adipocyte triglyceride lipase (ATGL), hormone sensitive lipase (HSL), and monoglyceride lipase (MGL), carry out lipolysis by hydrolyzing the ester bonds in tri-, di-, and monoglycerides in adipocytes, respectively (**Figure 2.19**).⁶⁴ These enzymes are in

turn regulated by hormones like insulin and epinephrine. Many biochemists and biologists are seeking to illuminate the complex interplay of these enzyme and regulatory pathways in order to better understand metabolic diseases like diabetes.^{51,65,66} For example, the inhibition of HSL by insulin is known to lead to an increase in diglyceride content in adipocytes.⁶⁷ Furthermore, it is known that lipolytic enzymes exhibit substrate preferences,^{11,68,69} such as the preference of HSL for 2,3-diglycerides over 1,3-diglycerides.⁹ The different preferences of the enzymes in conjunction with their regulation by different hormones could lead to changes in glyceride profiles in adipocytes in hormonal disorders.⁶⁴

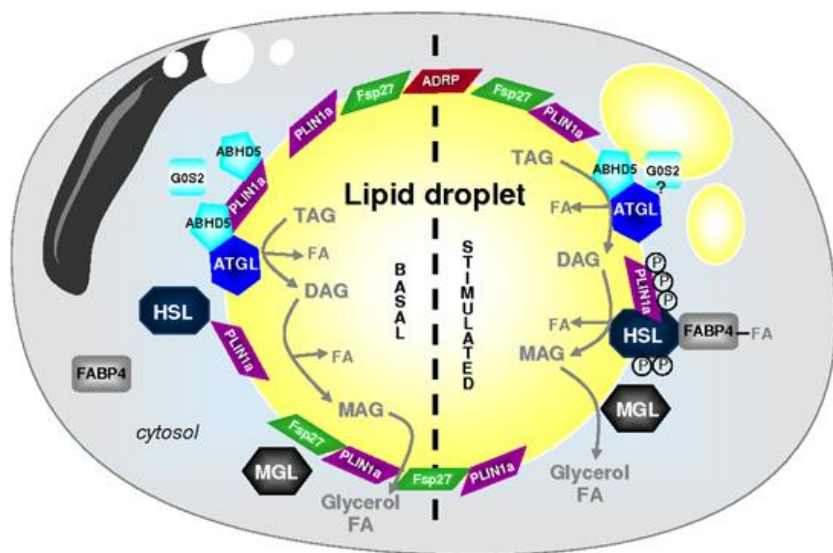
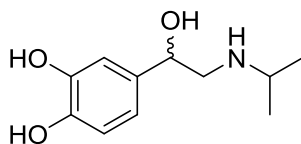


Figure 2.19 *Lipolysis in an adipocyte under basal and stimulated conditions.* ATGL = adipocyte triglyceride lipase, HSL = hormone sensitive lipase, MGL = monoglyceride lipase, TAG = triglyceride, DAG = Diglyceride, MAG = monoglyceride, FA = fatty acid, P = phosphate group. The other components are additional cofactors. (Reproduced from Ref. 64: © Nature Publishing Group, 2012)

2.10.1 Extracts from Human and Rat Adipocytes

We hypothesized that our array could be applied to extracts from mammalian adipocytes to detect differences in these glyceride mixtures. Hence, our collaborators at Sanofi-Aventis and the Helmholtz Institute provided us with extracts from human and rat adipocytes. The fat cells were harvested from the respective organisms and then were exposed to three different types of *in vitro* conditions: no treatment (control), insulin treatment, or isoproterenol treatment (**Table 2.4** and **2.5**, **Scheme 2.10**). Insulin and isoproterenol are known to regulate lipolysis.⁶⁷ Specifically, insulin inhibits HSL, while isoproterenol stimulates HSL activity (**Figure 2.19**). Thus, these exposures were hypothesized to induce glyceride profile changes in the cells: affecting the activity of the substrate-specific HSL while leaving other lipases at their normal level of activity should cause changes in the mobilization of some types of glycerides compared to other types. After treatment, the fat cells were lysed and centrifuged, and the lipid droplet, which is essentially a mixture of glycerides, was collected and provided to us. The fat material was further treated by taking up the samples in ethyl acetate and washing with brine solution. The ethyl acetate portion of each sample was collected, filtered to remove any insoluble material remaining, and placed in a pre-weighed vial to obtain an accurate weight of the material. The solvent was removed by rotary evaporation and then overnight drying on the high vacuum line. The glyceride material was a clear, viscous liquid.



Scheme 2.10 *Isoproterenol*. It is structurally similar to epinephrine.

Table 2.4 *Human fat samples.* HA1-HA8 were provided as a set and HA9-HA14 were provided as a set.

Sample	Treatment	Sample	Treatment
HA1	Control	HA8	Control
HA2	+ 10 nM insulin	HA9	Control
HA3	Control	HA10	+ 10 nM insulin
HA4	Control	HA11	+ 1 μ M isoproterenol
HA5	Control	HA12	Control
HA6	+ 10 nM insulin	HA13	+ 10 nM insulin
HA7	+ 10 nM insulin	HA14	+ 1 μ M isoproterenol

Table 2.5 *Rat fat samples including time courses.*

Sample	Treatment	Sample	Treatment
1	control	16	+ 10 nM insulin
4	+ 1 μ M isoproterenol	37	+ 10 nM insulin, 0 min
28	+ 1 μ M isoproterenol, 0 hr	40	+ 10 nM insulin, 30 min
31	+ 1 μ M isoproterenol, 3 hr	43	+ 10 nM insulin, 6 hr
34	+ 1 μ M isoproterenol, 6 hr		

2.10.2 Application of the Array to the Human Fat

Since the fat samples were mixtures of glycerides with different molecular weights, weight/volume concentrations were used instead of molar concentrations in order to introduce the same amount of material from each sample to the array. Most of the glycerides in the mixture are expected to be triglycerides derived from myristic, palmitic, palmitoleic, oleic, linoleic, and stearic acids.¹ These glycerides range in molecular weight from about 700-900 Da. Thus, an average molecular weight of 800 Da was chosen to loosely correlate a weight/volume concentration to a molar one (e.g. 9 mM \approx 7.2 mg/mL). The purpose of this calculation was to ensure that a similar concentration of glyceride was

being used in the array with these mixtures as was used successfully with the pure samples of glyceride.

The fat samples were subjected to the array in the same manner as for the pure glyceride samples. Samples HA1-HA8 were analyzed together (**Figures 2.20** and **2.21**).

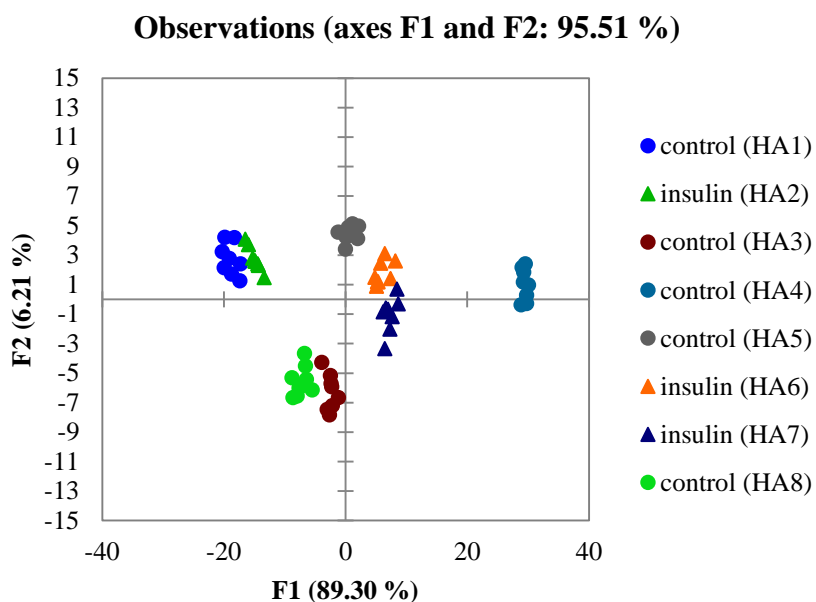


Figure 2.20 *LDA plot of array data from analysis of HA1-HA8.* The insulin-treated samples are marked as triangles. The array components consisted of BSA and HSA (100 μ M), fat (0.072 mg/mL), DNSA (60 μ M), ANS (60 μ M), NBD-FA (60 μ M), metathesized fat (0.072 mg/mL), AF (100 μ M), and DNSA (60 μ M) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters). Cross validation: 98%.

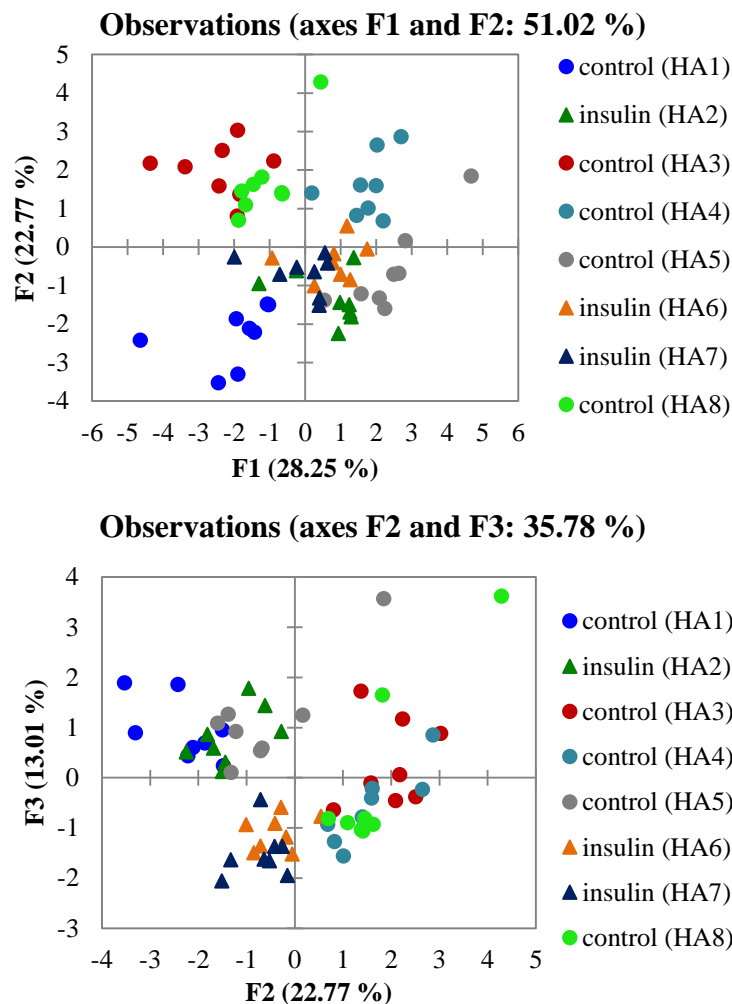


Figure 2.21 *PCA plot of array data from analysis of HA1-HA8. The insulin-treated samples are marked as triangles. Top: F1 versus F2. Bottom: F2 versus F3. The array components consisted of BSA and HSA (100 μ M), fat (0.072 mg/mL), DNSA (60 μ M), ANS (60 μ M), NBD-FA (60 μ M), metathesized fat (0.072 mg/mL), AF (100 μ M), and DNSA (60 μ M) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters).*

The clustering in the LDA plot is quite good; however, the three insulin-treated samples are not grouped together, nor are the control samples. The PCA plot is more revealing in that it shows very poor clustering and separation, which is not resolved even when considering F3. Looking at F1 versus F2, the three insulin-treated samples are

overlapping one another. The control samples are scattered around the insulin-treated samples and are not well separated from them. We repeated the array on samples HA3-HA8 again (**Figures 2.22** and **2.23**). There was insufficient HA1 and HA2 material to analyze them again.

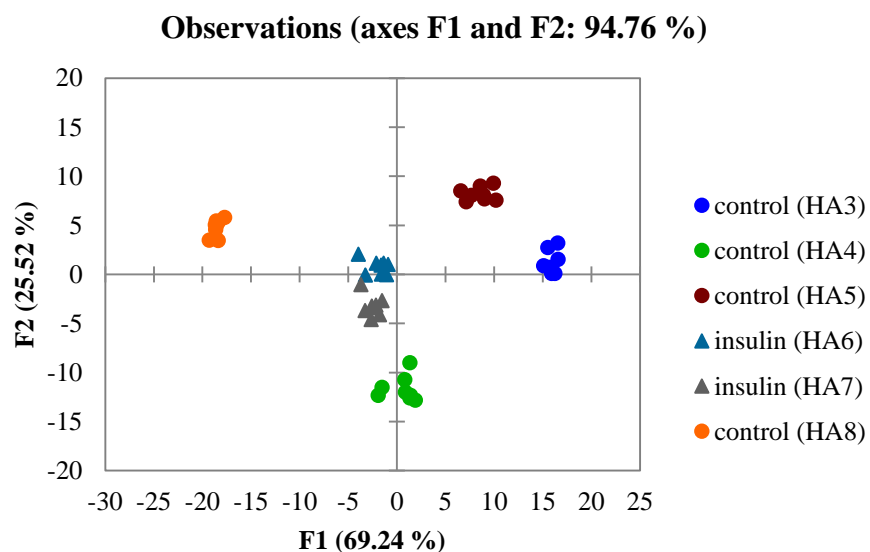


Figure 2.22 *LDA plot of array data from analysis of HA3-HA8.* The insulin-treated samples are marked as triangles. The array components consisted of BSA and HSA (100 μ M), fat (0.072 mg/mL), DNSA (60 μ M), ANS (60 μ M), NBD-FA (60 μ M), metathesized fat (0.072 mg/mL), AF (100 μ M), and DNSA (60 μ M) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters). Cross validation: 98%.

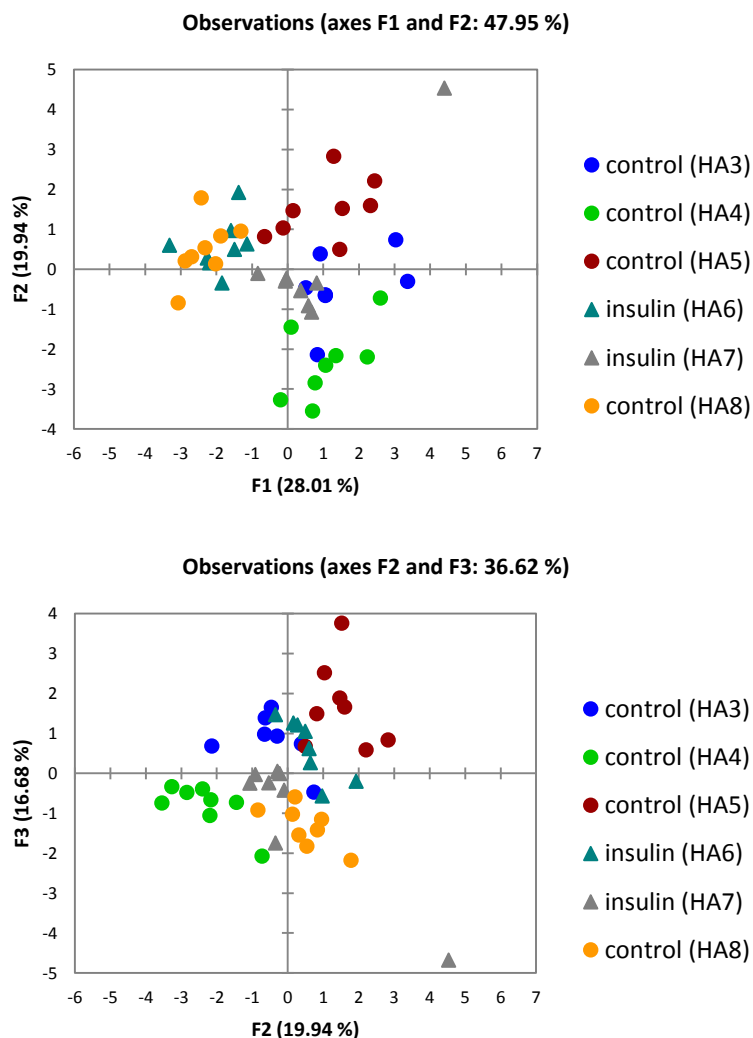


Figure 2.23 *PCA plot of array data from analysis of HA3-HA8.* The insulin-treated samples are marked as triangles. Top: F1 versus F2. Bottom: F2 versus F3. The array components consisted of BSA and HSA (100 μ M), fat (0.072 mg/mL), DNSA (60 μ M), ANS (60 μ M), NBD-FA (60 μ M), metathesized fat (0.072 mg/mL), AF (100 μ M), and DNSA (60 μ M) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters).

Similar results were obtained in the repetition of the analysis of HA3-HA8, which indicates good reproducibility in the method. There is again poor clustering and separation in the PCA plot. This data set represents a good example of why employing both PCA and

LDA is useful: PCA can sometimes reveal the apparently good clustering and separation in an LDA plot to be forced and potentially not meaningful. In both repetitions, the insulin-treated samples do seem to be associated with one another in the LDA and PCA plots, but the control samples are positioned all around the insulin-treated samples. Normally, we would expect to see that two regions of a plot are separable by a boundary that can be drawn as a straight line (e.g. control samples on one side, insulin-treated samples on the other) rather than by a circle. The meaning of this circular boundary between the two sample types is unclear. Putting this inconclusive data aside, we next analyzed HA9-HA14, which also contain isoproterenol-treated samples, with the array (**Figures 2.24 and 2.25**).

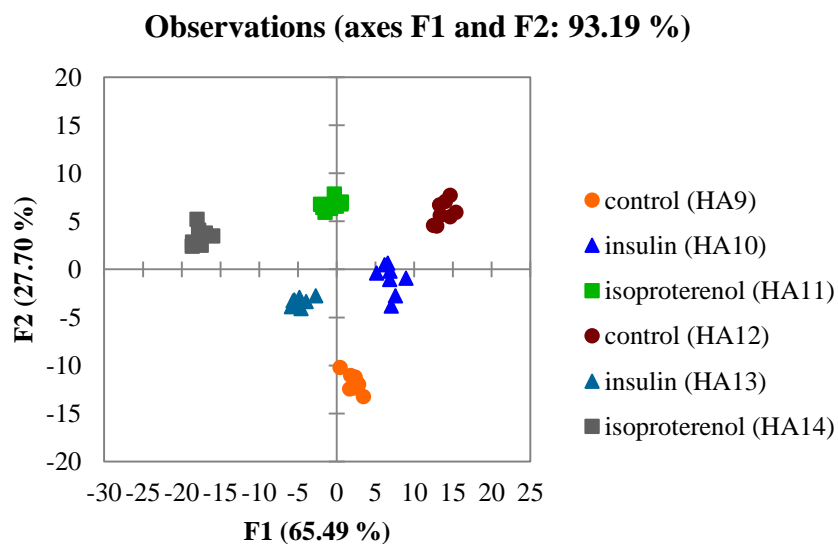


Figure 2.24 *LDA plot of array data from analysis of HA9-HA14.* The array components consisted of BSA and HSA (100 μ M), fat (0.072 mg/mL), DNSA (60 μ M), ANS (60 μ M), NBD-FA (60 μ M), metathesized fat (0.072 mg/mL), AF (100 μ M), and DNSA (60 μ M) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters). Cross validation: 98%.

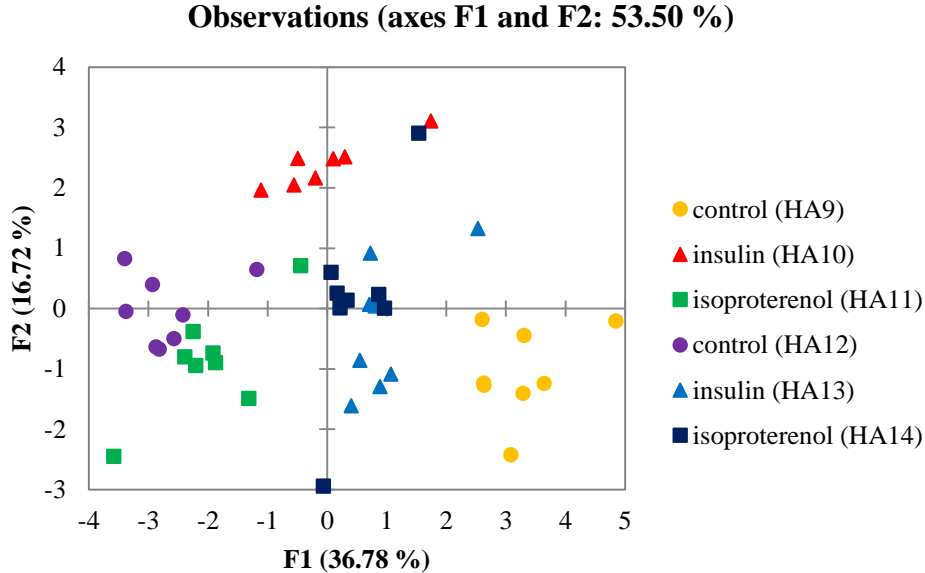


Figure 2.25 *PCA plot of array data from analysis of HA9-HA14.* The array components consisted of BSA and HSA (100 μ M), fat (0.072 mg/mL), DNSA (60 μ M), ANS (60 μ M), NBD-FA (60 μ M), metathesized fat (0.072 mg/mL), AF (100 μ M), and DNSA (60 μ M) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters).

Again, similar results were obtained. The LDA plot looks promising with good separation and clustering, and the treated samples seem to be associated properly (insulin with insulin, isoproterenol with isoproterenol). However, the control samples are not grouped in their own region of the plot. The PCA shows poor clustering and substantial overlap with no obvious groupings with respect to the sample treatment conditions. At this point, we concluded that the array was not discriminating the treatment conditions of the human fat samples in any meaningful or useful way.

2.10.3 Application of the Array to the Rat Fat

The rat fat was applied to the array at the same weight/volume concentration as the human fat. For the first analysis, one sample of each treatment type was selected: 1 (control), 4 (isoproterenol), and 16 (insulin) (**Figures 2.26 and 2.27**).

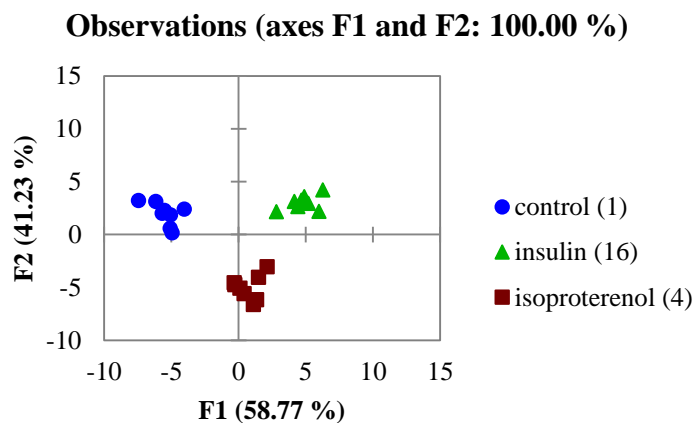


Figure 2.26 *LDA plot of array data from analysis of rat samples 1, 4, and 16.* The array components consisted of BSA and HSA (100 μ M), fat (0.072 mg/mL), DNSA (60 μ M), ANS (60 μ M), NBD-FA (60 μ M), metathesized fat (0.072 mg/mL), AF (100 μ M), and DNSA (60 μ M) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters). Cross validation: 92%.

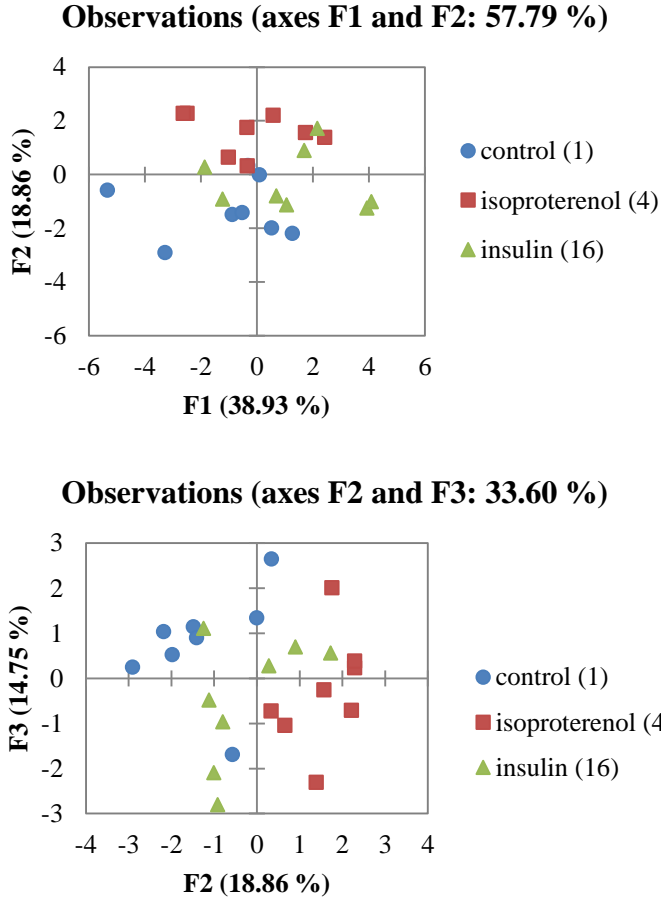


Figure 2.27 *PCA plot of array data from analysis of rat samples 1, 4, and 16. Top: F1 versus F2. Bottom: F2 versus F3. The array components consisted of BSA and HSA (100 μ M), fat (0.072 mg/mL), DNSA (60 μ M), ANS (60 μ M), NBD-FA (60 μ M), metathesized fat (0.072 mg/mL), AF (100 μ M), and DNSA (60 μ M) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters).*

Even though the LDA plot indicates good discrimination of the samples, the PCA plot shows substantial overlap between the three samples, even when considering a third dimension. We decided to try performing the assay at a higher concentration of fat to see if this change would improve the results. All of the rat samples were analyzed at a higher concentration (**Figures 2.28 and 2.29**).

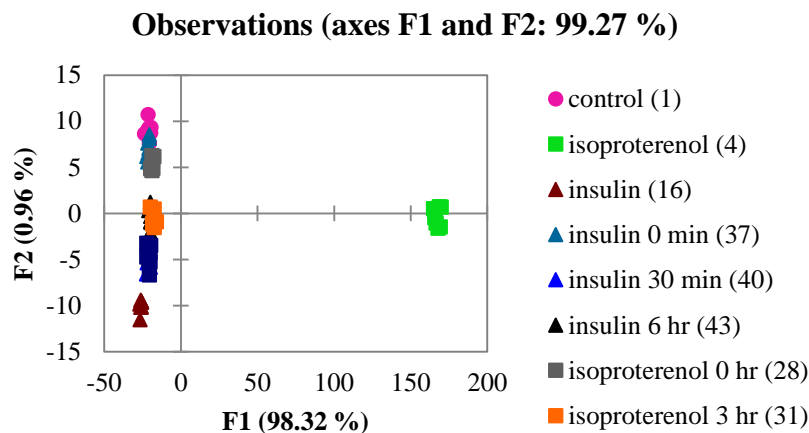


Figure 2.28 *LDA plot of array data from analysis of all of the rat samples.* The array components consisted of BSA and HSA (100 μM), fat (0.6 mg/mL), DNSA (60 μM), ANS (60 μM), NBD-FA (60 μM), metathesized fat (0.6 mg/mL), AF (100 μM), and DNSA (60 μM) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters). Cross validation: 100%.

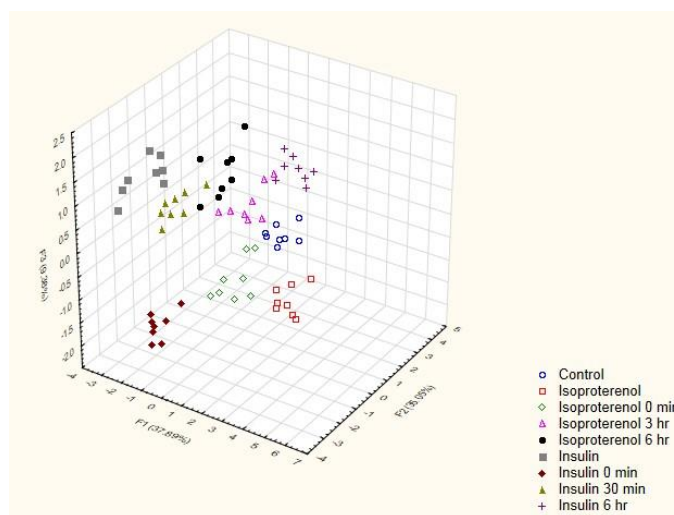


Figure 2.29 *PCA plot of array data from analysis of all of the rat samples.* The array components consisted of BSA and HSA (100 μM), fat (0.6 mg/mL), DNSA (60 μM), ANS (60 μM), NBD-FA (60 μM), metathesized fat (0.6 mg/mL), AF (100 μM), and DNSA (60 μM) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters).

At first, these results appeared more promising than the results from the human samples. The LDA plot is dominated by the separation between isoproterenol (4) and the rest of the samples, but the PCA plot does not include this feature, so it is likely not truly significant. While the PCA plot shows decent clustering within the different sample conditions, it is unclear what the relative position of the different groups could mean. For example, the insulin-treated samples are just as close to one another as they are to isoproterenol-treated samples. Our hypothesis that the insulin and isoproterenol would have very different effects on the glyceride profiles does not seem to be borne out by these results. Hence, we cannot be confident that the discrimination of the rat fat samples in the LDA or PCA is based on real differences in the glyceride content. Rather, the chemometric algorithms could be “discriminating” the samples based on small differences in concentration and other sample preparation imperfections. To probe this question, we repeated the analysis of the rat samples in exactly the same way (**Figures 2.30** and **2.31**). There was insufficient sample 30 (insulin 30 min) to use in the array, so it was excluded.

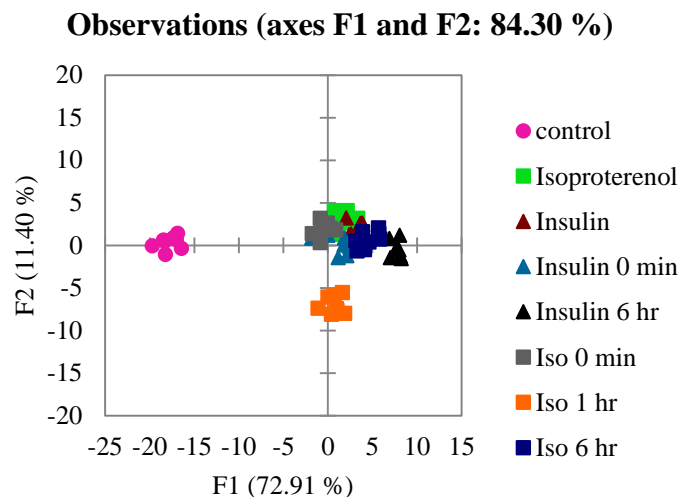


Figure 2.30 *LDA plot of array data from the repetition of the analysis of the rat samples.* The array components consisted of BSA and HSA (100 μ M), fat (0.6 mg/mL), DNSA (60 μ M), ANS (60 μ M), NBD-FA (60 μ M), metathesized fat (0.6 mg/mL), AF (100 μ M), and DNSA (60 μ M) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters). Cross validation: 95%.

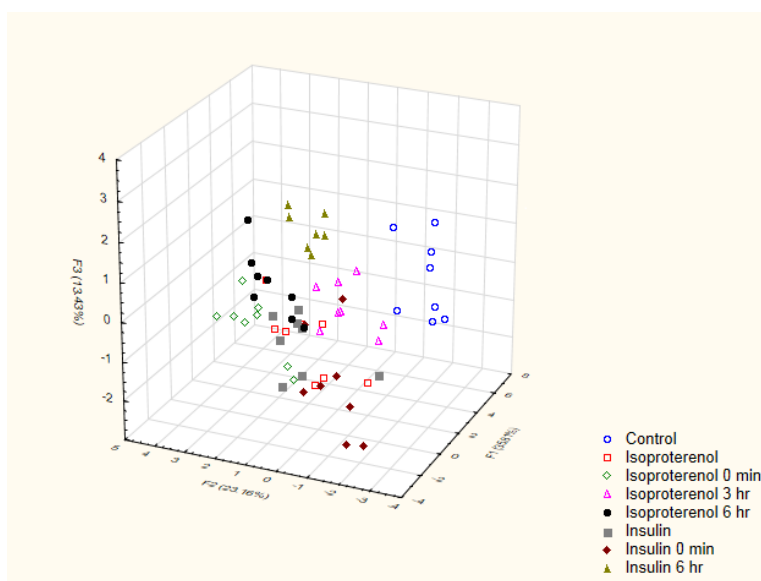


Figure 2.31 *PCA plot of array data from the repetition of the analysis of the rat samples.* The array components consisted of BSA and HSA (100 μ M), fat (0.6 mg/mL), DNSA (60 μ M), ANS (60 μ M), NBD-FA (60 μ M), metathesized fat (0.6 mg/mL), AF (100 μ M), and DNSA (60 μ M) in phosphate buffer with < 5% THF (see **Table 2.8** for read parameters).

The results from the repetition of the array do not resemble the initial results in terms of the relative positioning of the sample types. Furthermore, the scatter and overlap are worse in the repetition. Since the results were not reproduced, the apparent differentiation of the analytes can likely be attributed to small errors in preparing the array (e.g. solution preparation, pipetting, and measuring of the fluorescence). Our conclusion at this point was that the array was not successful in discriminating either the human or rat fat samples based on differences in their glyceride content. Whatever differences in the type and number of glycerides may be present in the samples, this method is not sensitive enough to detect them reproducibly. We next examined mass spectral and $^1\text{H-NMR}$ data from the fat samples in order to determine if any differences could be detected with these methods.

2.10.4 Mass Spectral and $^1\text{H-NMR}$ Data from Fat Samples

Since the glycerides do not ionize readily, it was not possible for us to obtain mass spectral data from the fat samples directly. However, we were able to analyze the metathesis reactions with LC-MS to gain insight into the unsaturated species present in the sample. In fact, this data was obtained routinely for each repetition of the array in order to ensure that the metathesis reaction was working each time.

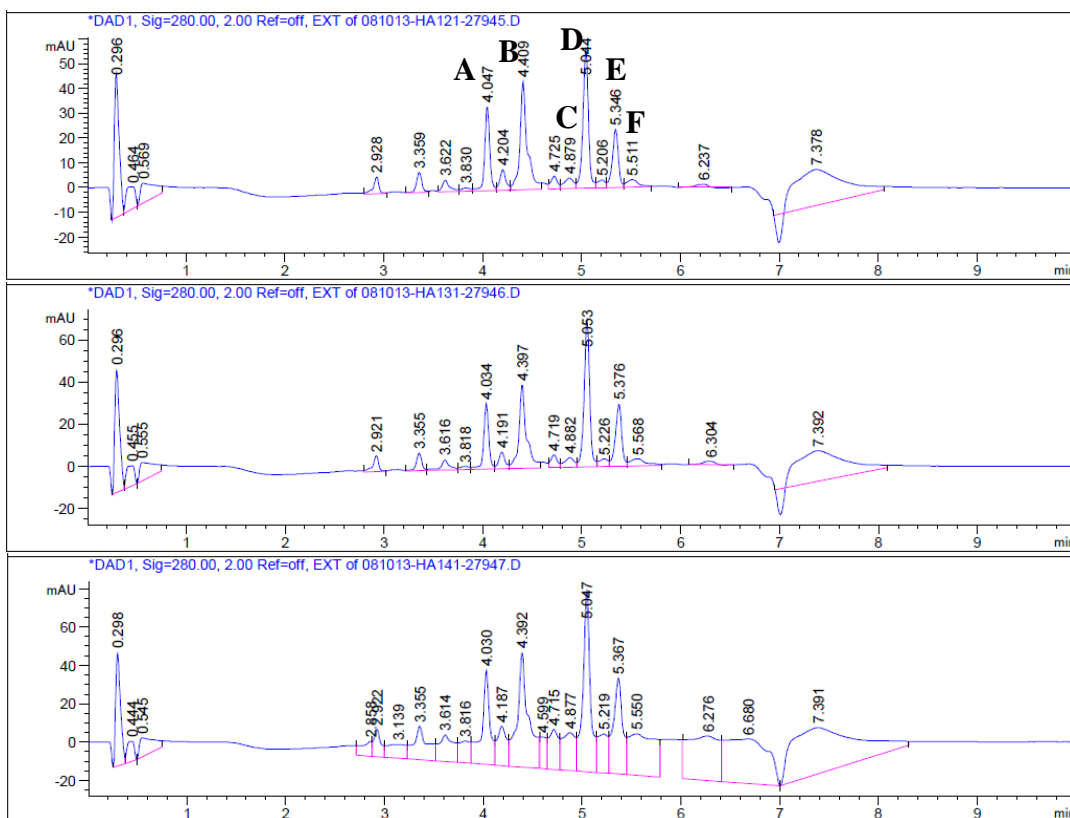


Figure 2.32 *Chromatograms of metathesis reactions of human fat samples.* Top: control (HA12), middle: insulin (HA13), bottom: isoproterenol (HA14). Absorbance detection at 280 nm. From the corresponding ion chromatograms: A- AF, B- 2.2, C- linoleic acid-derived species, D- palmitoleic acid-derived species, E- oleic acid-derived species, and F- petroselaidic acid-derived species.

The chromatograms for the metathesis of human fat samples that had been untreated, insulin-treated, or isoproterenol-treated were nearly identical (**Figure 2.32**). From the corresponding ion chromatograms, some of the species derived from unsaturated glycerides mixed with AF could be identified as indicated in the figure caption (A-F). Based on this data, it was unsurprising that our array could not discriminate the glycerides, since the array works best at differentiating unsaturated glycerides. We performed the

same analysis on the metathesis reactions of the rat fat of the three types from the first repetition of the array with the rat samples (**Figure 2.33**).

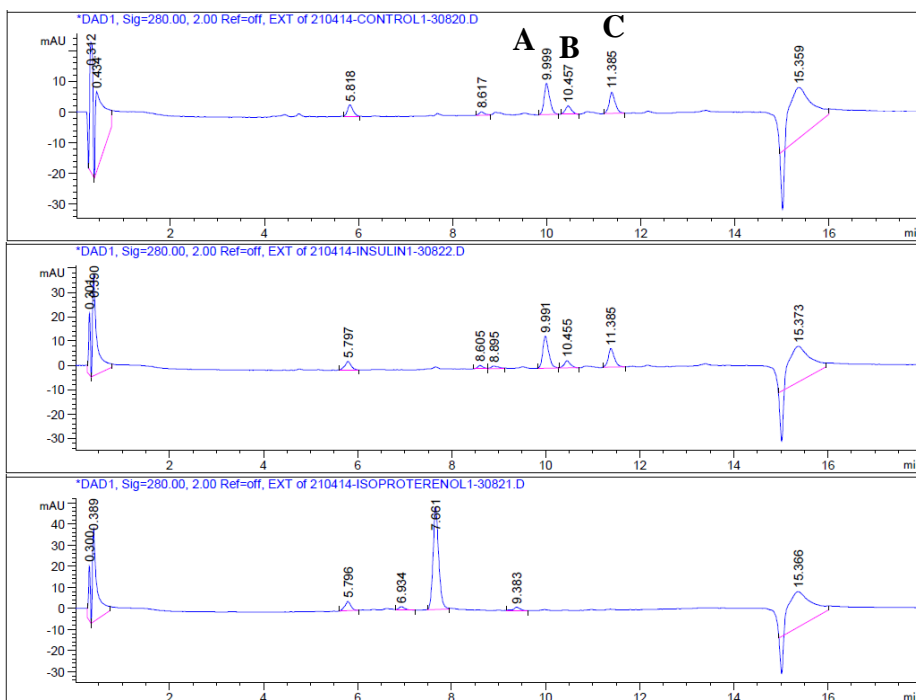


Figure 2.33 *Chromatograms of metathesis reactions of rat fat samples.* Top: control (1), middle: insulin (16), bottom: isoproterenol (4). Absorbance detection at 280 nm. From the corresponding ion chromatograms: A- linoleic acid-derived species, B- palmitoleic acid-derived species, and C- oleic acid-derived species.

The control and insulin samples were identical in their unsaturated glyceride content based on the chromatograms. Unreacted AF was not observed in those chromatograms because the rat fat was introduced at a much higher concentration to the metathesis reaction than with the human fat.

The isoproterenol-treated sample shows only AF (peak at 7.66 min), indicating that there was no reaction. This failure was probably due to some experimental error, like the

presence of THF, rather than the absence of any unsaturated glycerides. This finding explains the LDA plot obtained from the array analysis of these samples (**Figure 2.28**), which had the isoproterenol sample separated markedly from all of the others. This feature was not seen again in the repetition of the experiment (**Figure 2.30**), which further indicates the failure of the metathesis reaction with the isoproterenol-treated sample was due to an error in setting up the reaction rather than an indication that the sample contained no unsaturated glycerides.

Interestingly, while all of the human samples were identical, as were the rat samples, the human and rat samples are different from one another. According to the LC-MS data, the human and rat samples contain different amounts of linoleic, palmitoleic, and oleic acid-derived species.

NMR data was also collected from the rat fat of the three sample types. There was insufficient human fat to obtain NMR data. The ^1H -NMR spectra for the control, insulin, and isoproterenol rat samples were largely the same (**Figure 2.38**). The only potentially significant difference was in the peaks from 4.0-4.5 ppm, which are the glycerol methylene peaks of the glycerides. Additional small peaks in this region could be from mono- and diglyceride species.

Overall, the MS and NMR data support the notion that the fat samples do not differ significantly in their glyceride composition based on sample treatment. Hence, it is unsurprising that our array was unable to discriminate the samples. In fact, this finding supports the validity of our array, since the array correctly detected no differences in the samples when there were none.

2.11 QUANTITATION OF TRILINOLEIN (T5) IN TRIGLYCERIDE MIXTURES USING STANDARD ADDITION

2.11.1 Introduction to standard addition and net analyte signal (NAS)

Since quantitation is one of the major challenges that mass spectrometric analysis faces in regard to glyceride analysis, we sought to demonstrate the use of our assay for quantitation of a glyceride of interest within a mixture of other glycerides. Typically, mixtures of glycerides found in edible oils or adipocyte cell extracts are separated by HPLC or GC before being individually identified by mass spectrometry and quantified by standard calibration methods.² However, we wanted to eliminate the need for an HPLC separation step in order to simplify the analysis by our assay, as this separation itself can be challenging to accomplish.¹⁹ Because the receptors in the array are semi-selective in nature, the presence of glycerides other than the species of interest in the mixture will interfere with the signal elicited from the receptor by that species. Hence, we chose to use the “standard addition” method to minimize the influence of the other glycerides in the mixture on the quantitation of the glyceride of interest.

Standard addition is a classic calibration method that is used when matrix effects are present in a sample that interfere with the signal the compound of interest elicits, confounding accurate quantitation of that compound.⁷⁰ Ordinarily, the experimentalist tries to remove interferents from the sample whenever possible; however, when it is not facile to physically remove the interfering background matrix, standard addition can be employed to minimize the matrix’s effect on the analysis. Standard addition is carried out by measuring the response of the mixture after successive additions of the pure analyte of interest. By plotting the response against the amount of the pure analyte in each addition, the concentration of the analyte of interest in the sample is obtained by a linear fit of the data. The x-intercept of the line is the original analyte concentration. This data analysis

procedure is used when dealing with univariate data (i.e. one signal, y , describes the concentration, x , of the analyte); however, in the multivariate case (i.e. multiple signals, y_1, y_2, \dots, y_n , describe the concentration, x , of the analyte), the data analysis employed to determine the analyte concentration is necessarily more complicated.⁷¹

In the literature, there are numerous examples of multivariate standard addition procedures,⁷²⁻⁷⁷ usually dealing with spectral or electrochemical data obtained from complex sample mixtures. Nevertheless, to our knowledge this is the first report in which a multivariate standard addition has been performed with a cross-reactive chemosensor array. For our purposes, we chose to follow the mathematical procedure developed by Hemmateenejad and Yousefinejad, called standard addition by net analyte signal (SANAS), due to its simplicity and accuracy.⁷⁸ The net analyte signal (NAS) is defined as the portion of the total signal that is directly related to the concentration of the analyte.⁷⁹ For some total signal elicited from a mixture, some portion of the signal is due to the component of interest and the rest is due to the background and other components of the mixture. Lorber has shown that the contribution to the signal due to the component of interest, the NAS, can be computed because it is orthogonal to the contribution to the signal of the background and other components.⁸⁰ Linear algebra is used to determine this orthogonal part of the signal as a vector. In this way, the influence of the matrix as well as background noise can be removed.

In SANAS, the data is first subjected to dimensionality reduction, commonly by principal component analysis (PCA) or partial least squares regression (PLS). As described earlier in this report, these algorithms transform the data into a space described by new orthogonal axes, called factor axes, that are linear combinations of the original variables (i.e. the signals y_1, y_2, \dots, y_n). These factor axes are chosen by the algorithm according to different criteria depending on the exact method used (e.g. PCA, LDA, or PLS), but in

general these criteria lead to factor axes that are more information dense. The first factor axis (F1) contains the most relevant information about the data, F2 the second most, and so on up to FN, where N = number of original variables. With our data, we obtained better results using PLS rather than PCA. PLS relates a matrix \mathbf{X} (the emission data from our array) to a property \mathbf{y} (concentration of the additions).⁸¹ PLS uses maximum covariance as its criterion for determining the new factor axes, unlike PCA which uses maximum variance. The advantage of using covariance is that it includes both variance of \mathbf{X} as well as correlation of \mathbf{X} and \mathbf{y} .

By rebuilding the data set from fewer factor axes than the total number (for example, from F1-F3, discarding F4-F10), irrelevant information like noise, background, or interferences can be removed from \mathbf{X} . From the rebuilt data set (\mathbf{X}'), the NAS for the analyte of interest is calculated by using a type of single value decomposition method called rank annihilation (see reference 78 for a detailed account of this series of computations). The output is a vector that describes the NAS ($\|\text{NAS}\|$). By plotting the Euclidean norm of this vector (i.e. its magnitude) against the concentrations of the additions, one obtains a plot that is equivalent to the one described for the univariate case described above. The x-intercept of the linear fit has the same meaning as in the univariate case, that is, the concentration of the analyte of interest in the mixture.

2.11.2 Applying SANAS to mixtures of glycerides

The first step in performing the quantitative analysis with our array was to generate mixtures of triglycerides with a known composition to study the accuracy of the method. Mixtures consisting of **T1**, **T3**, **T5**, **T6**, and **T7** were carefully prepared (**Table 2.9**). Weight/volume concentrations were used since the molecular weights differ for different glycerides. The mixture solution was aliquoted into vials and each aliquot was spiked with

an increasing amount of pure **T5** at a constant total volume for seven additions. The eight resultant solutions were then subjected to the array, and the fluorescence data was obtained as described previously. The final concentration of **T5** from the mixture that was in the well plates of the array is shown in **Table 2.6**. This concentration is the one that we are measuring with the array. The data from the array was normalized and then subjected to PLS analysis as this method of dimensionality reduction was found to give a slightly more accurate result compared to PCA. The data was rebuilt from F1 only (mixtures A, B, C) or from F1 and F2 (mixture D) to obtain the closest estimates. In all cases, F1 described 90% or more of the covariance in the data sets.

From the rebuilt data matrix, the net analyte signal ($\|NAS\|$) was calculated and plotted against the additions to give SANAS plots (**Figures 2.34-2.37**). From these plots, the concentration of **T5** in the mixture was determined by finding the x-intercept of the linear fit (**Table 2.6**). Good estimates of the concentration of **T5** in the mixtures were obtained for mixtures B, C, and D; however, a larger overestimation of the concentration of **T5** in mixture A was obtained. We attribute this error to poor linearity in the response of the sensors to the additions of **T5** that is evidenced by the lower correlation coefficient for the linear fit in the SANAS plot for mixture A (**Table 2.6**). Some of the non-linearity is likely due to random experimental errors in solution and plate preparation. However, some deviations from linearity of the responses to the additions to all of the mixtures are inherent to the experimental method. The fluorescence signal collected from each well of the plate is mediated by the serum albumin, fluorophore, and glycerides in the solution and is an indirect measurement of the **T5** content. Hence, we expect that non-linearity in the response results from both the saturation behavior of ligand binding to serum albumin as well as binding competition from glycerides in the mixture that are not **T5**. Nevertheless, for mixtures B, C, and D the array response showed a good linear dependence on the

additions of **T5**. The SANAS method was able to provide quite accurate values for the **T5** content without the necessity for removing or otherwise separating out the other triglycerides in the mixture.

Table 2.6 Results for **T5** concentration in triglyceride mixtures from SANAS.

Mixture	Total concentration of the mixture (mg/mL)	Actual T5 concentration (mg/mL)	Predicted concentration (mg/mL)	R ² of SANAS plot
A	0.04	0.0113	0.0338	0.7403
B	0.04	0.0158	0.0160	0.9517
C	0.01	0.0053	0.0056	0.9519
D	0.01	0.0042	0.0047	0.9664

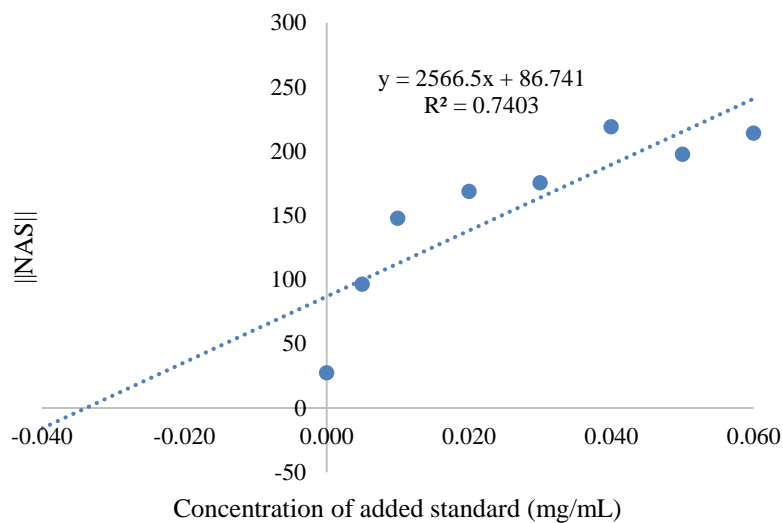


Figure 2.34 SANAS plot generated from the standard addition of mixture A.

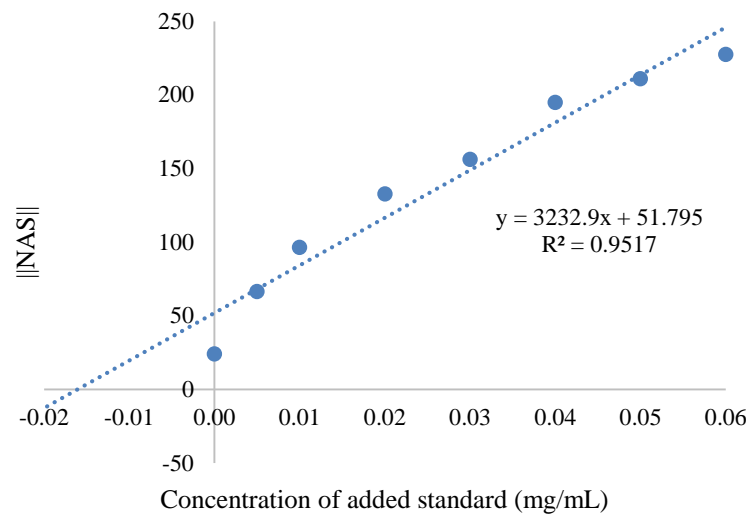


Figure 2.35 SANAS plot generated from the standard addition of mixture B.

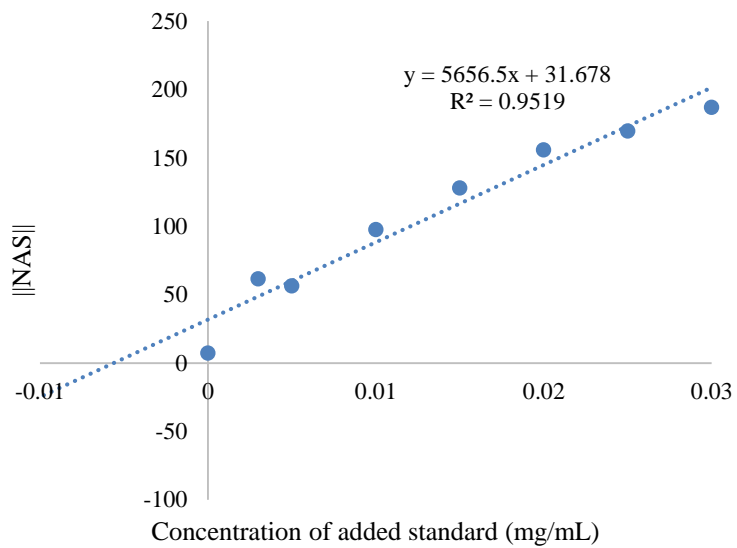


Figure 2.36 SANAS plot generated from the standard addition of mixture C.

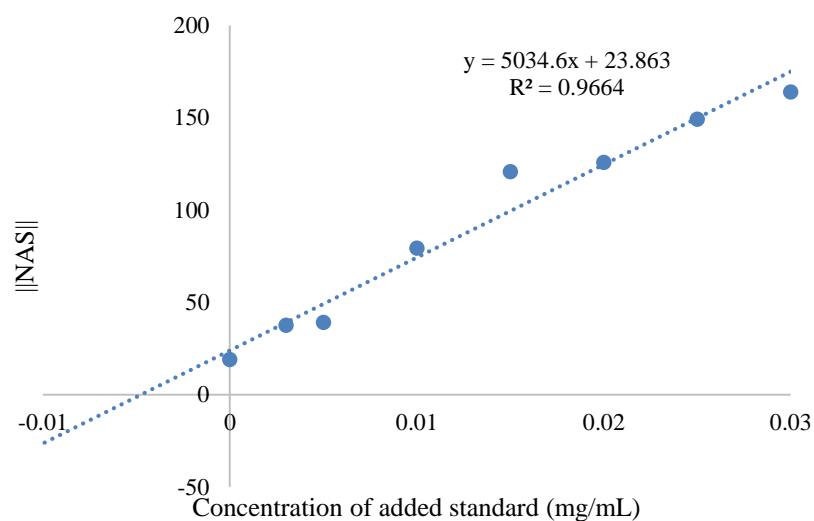


Figure 2.37 SANAS plot generated from the standard addition of mixture D.

2.12 CONCLUSIONS

In this work, we developed an array of cross-reactive serum albumins and fluorescent indicators to discriminate a panel of structurally similar glycerides. An olefin cross metathesis reaction with an allyl fluorescein species was used to pretreat the samples prior to array analysis in order to distinguish unsaturated glycerides. Using this simple assay, we successfully discriminated twenty mono-, di-, and triglycerides, including stereo- and regioisomeric pairs. These isomer types included *cis/trans* stereoisomers, double bond positional regioisomers, and positional regioisomers on the glycerol core. Additionally, we used the array in conjunction with LDA to predict structural features of glycerides. However, the application of the array to adipocyte fat extracts from humans and rats did not result in reproducible discrimination of these samples. It is unclear that the insulin and isoproterenol treatments were the appropriate concentration and length of time to stimulate differences in the cells. Thus, there is a question as to whether the poor discrimination is

indicative of a sensitivity problem with the array or of a true lack of differences in the samples.

We demonstrated the first use of multivariate standard addition with a cross-reactive array of chemosensors. We used our array to perform a standard addition of trilinolein to mixtures of triglycerides. By applying the established standard addition by net analyte signal (SANAS) method, we were able to quantitate the trilinolein in the mixtures with good accuracy.

2.13 FUTURE WORK

While we have clearly demonstrated that our array can be used successfully for a number of applications, the method could be further improved to be applicable to a range of glyceride-based analytical challenges. The array could be improved for discrimination of mono-, di-, and triglycerides by functionalizing the free hydroxyls of mono- and diglycerides to accentuate their differences. While the array was not successful at distinguishing the adipocyte extracts, the treatment protocols could be improved or changed to ensure that there would be differences in the fat samples. Other treatment conditions such as different diets or induced diabetic states in whole organisms could be used instead of just treatment of cells with insulin or isoproterenol. A series of artificially prepared mixtures of glycerides could be used to test the sensitivity of the array to differences in composition. Finally, the quantitation of glycerides of interest in a real world sample like an adipocyte fat extract could be tested.

2.14 CONTRIBUTIONS

The experiments that Michelle Adams Ivy performed for this work were indicated by citing her dissertation (Ref. 50) where her results can be found. The adipocyte samples

were prepared by Günter Müller of the Helmholtz Institute München, formerly of Sanofi-Aventis. The SANAS code was written by Scott Rabidoux (UT Austin).

2.15 EXPERIMENTAL METHODS

2.15.1 General Information

Human and bovine serum albumin were purchased fatty acid free from Sigma Aldrich. Glycerides were purchased from Nu-Chek Prep (**D1, D2, D4**), Sigma Aldrich (**T1, T2, T4, T8, D4, D7**), ABCR (**T3, T5, T6, M1, M2, M3, M4, M5**), Acros (**T7**), Bachem (**D6**), and Zerenex (**D5**). Grubbs Second Generation Catalyst and 1,8-anilinonaphthalene sulfonic acid was purchased from Sigma Aldrich. Dansyl amide and 2-anthracene carboxylate were purchased from TCI America. The cuvette experiments were performed with a PTI fluorimeter using an 814 photomultiplier detection system and a 75W xenon short arc lamp. The arrays were prepared in Costar 96-well plates (#3915). The plates were analyzed using a Biotek Synergy 2 Multimode Microplate Reader or a Biotek Cytation 3 Microplate Reader. LC-MS analysis was performed on an Agilent 1200 Series HPLC with an Agilent 6130 single quadrupole mass spectrometer (ESI and APCI ionization). NMR spectra were taken on a 500 MHz Varian NMR. Linear discriminant analysis, principal component analysis, and data normalization were done using XLSTAT 2011. The SANAS code was run on MATLAB 2014.

2.15.2 Synthesis

2.15.2.1 *Synthesis of NBD-FA*

To 100 mL methanol was added 12-aminododecanoic acid (5.61 mmol, 1.2 g) and Hünig's base (10.6 mmol, 1.85 mL). Then, 7-chloro-4-nitrobenzo-3-oxa-1,3-diazole (5.01 mmol, 1.0 g) was dissolved in 50 mL of methanol. This solution was added to the first one, and the reaction was stirred overnight at room temperature. The reaction slowly

darkened from yellow to a brownish-red color. The methanol was removed under vacuum leaving an oil. The oil was taken up in dichloromethane and washed several times with aqueous 1M HCl. Finally, the solvent was removed under vacuum to yield a reddish-brown powder (1.23 g, 65% yield). No further purification was necessary. ¹H-NMR (400 MHz, CDCl₃, ppm): δ 8.50 (d, 1H, H-Ar), 6.37 (m, 1H, H-N), 6.17 (d, 1H, H-Ar), 3.48 (dt, 2H, -CH₂-NH₂), 2.35 (t, 2H, -CH₂-COOH), 1.8-1.25 (m, 18H, 9 -CH₂). MS (ESI): 377.2 (M-1).

2.15.2.2 *Synthesis of allyl fluorescein*

Fluorescein (4 g, 12.0 mmol) and sodium hydroxide (1 g, 12.0 mmol) were dissolved in 50 mL of water to yield the disodium salt. The water was removed by lyophilization or centrifugal evaporation. The disodium salt was dissolved in 150 mL water/80 mL acetone. Allyl bromide (2.4 mL, 27.6 mmol) was added to the reaction, and the reaction was refluxed for 2 hr. The reaction mixture was poured into 150 mL of water, and this solution was basified to yield an orange precipitate. A red oil remained in the reaction flask. This oil was dissolved in acetone and then poured into 100 mL of water. This solution was also basified to form the orange precipitate. The two orange precipitates were collected by filtration and combined. The fluorescein diether was recrystallized in ethanol/dilute NaOH (aq). Finally, the diether was dissolved in 30 mL acetone and heated to reflux. Then 20 mL of 1 M NaOH (aq) was added, and the reflux was continued for 12 min. The reaction mixture was poured into 100 mL of water, and this solution was acidified, becoming yellow and cloudy. The solution was extracted 3x with dichloromethane. The dichloromethane was dried with magnesium sulfate and removed under vacuum. The resulting product was a yellow powder (357 mg, 8% yield). ¹H-NMR (400 MHz, CDCl₃, ppm): δ 8.00 (dt, 1H, H-Ar-COOH), 7.65 (td, 1H, H-Ar-COOH), 7.60

(td, 1H, H-Ar-COOH), 7.15 (dt, 1H, H-Ar-COOH), 6.76 (d, 1H, H-Ar), 6.71 (d, 1H, H-Ar), 6.67-6.50 (m, 6H, H-Ar), 6.02 (m, 1H, CH₂-CH=CH₂), 5.40 (dq, 1H, geminal *cis*-CH₂-CH=CH₂), 5.30 (dq, 1H, geminal *trans*-CH₂-CH=CH₂), 4.55 (dt, 2H, CH₂-CH=CH₂). MS (ESI): 373.2 (M+1).

2.15.2.3 *Synthesis of 1*

Michelle Adams Ivy synthesized **1**.⁵⁰

2.15.3 Array Procedures

2.15.3.1 *Glyceride Array (Part 1)*

Concentrated stock solutions of each glyceride were prepared in THF (concentrations on the order of millimolar). Stock solutions of human and bovine serum albumin were prepared in 10 mM phosphate buffer (pH 7.0, 0.02% NaN₃) at a concentration of 500-700 μM. Serum albumin-glyceride solutions were prepared from these stocks at concentrations of 150 μM serum albumin and 135 μM glyceride (3% THF, 97% 10 mM phosphate buffer, pH 7.0, 0.02% NaN₃). Concentrated stock solutions of dansyl amide and NBD-FA were prepared in DMSO. A concentrated stock solution of ANS was prepared in 10 mM NaOH (aq). Then solutions of each of the three indicators at a concentration of 180 μM in 10 mM phosphate buffer, pH 7.0, 0.02% NaN₃ were prepared.

The plates were made by placing 200 μL of each serum albumin-glyceride solution and 100 μL of the 180 μM indicator solution in each well (final concentrations: 100 μM serum albumin, 90 μM glyceride, 60 μM indicator, 2% THF). Each plate contained a column of indicator alone and a column of indicator and serum albumin as controls. Eight replicates were performed for each glyceride/serum albumin/indicator combination and for the controls.

The parameters for the reading of the well plates can be found in **Table 2.7**. The fluorescence data was normalized by first subtracting the emission of the control (indicator alone) from each data point, and then the data set was rescaled from 0 to 100.

Table 2.7 *Plate read parameters for part 1 of the glyceride array*

	NBD-FA	DNSA	ANS
λ_{ex}	485/20	360/40	380/20
λ_{em}	540/25	485/20	485/20
Mirror	Top 510	Top 400	Top 400

2.15.3.2 *Well Plate Metathesis Reaction*

Concentrated stock solutions of each glyceride were prepared in chloroform (concentrations were on the order of millimolar). Solutions of allyl fluorescein and Grubbs Second Generation Catalyst (G2) were also prepared in chloroform. These solutions were applied to a deep well polypropylene plate such that each well contained 2 mM glyceride, 2.2 mM allyl fluorescein, and 0.4 mM G2 in a total volume of 1 mL of chloroform. Two controls were also included that contained only allyl fluorescein and G2. The plate was placed in an oven and heated at 50°C for 2 hr. Then the plate was removed and left in the fume hood to finish evaporating the chloroform overnight at ambient temperature and pressure. The next day, the dry material that remained in each well was taken up in 1 mL of THF with thorough mixing via a glass pipette to completely dissolve all of the material in the THF.

2.15.3.3 *Glyceride Array with Metathesized Glycerides (Part 2)*

The solutions from the metathesis reaction in 1 mL of THF were used directly. Solutions of human and bovine serum albumin were prepared in 10 mM phosphate buffer (pH 7.0, 0.02% NaN₃) at a concentration of 550 μ M. Serum albumin-glyceride solutions

were prepared from these stocks at concentrations of 150 μM serum albumin, 135 μM glyceride, 149 μM AF, and 27 μM G2 (6.8% THF, 93.2% 10 μM phosphate buffer, pH 7.0, 0.02% NaN_3). A concentrated stock solution of dansyl amide was prepared in DMSO. Then a solution of dansyl amide at a concentration of 180 μM in 10 μM phosphate buffer, pH 7.0, 0.02% NaN_3 was prepared.

The plates were made by placing 200 μL of each serum albumin-glyceride solution and 100 μL of buffer or of the 180 μM dansyl solution in each well (final concentrations: 100 μM serum albumin, 90 μM glyceride, 99 μM AF, 60 μM dansyl amide, 4.5% THF). Each plate contained a column of indicator alone and a column of indicator and serum albumin as controls. Eight replicates were performed for each glyceride/serum albumin combination with either dansyl amide/AF or only AF and for the controls.

The parameters for the reading of the well plates can be found in **Table 2.8**. The fluorescence data was normalized by first subtracting the emission of the control (indicator alone) from each data point, and then the data set was rescaled from 0 to 100.

Table 2.8 *Plate read parameters for part 2 of the glyceride array*

	NBD-FA	DNSA1	ANS	AF1/2	FRET	DNSA2
λ_{ex}	485/20	360/40	380/20	485/20	360/40	360/40
λ_{em}	540/25	485/20	485/20	528/40	528/40	485/20
Mirror	Top 510	Top 400	Top 400	Top 510	Top 510	Top 400

2.15.4 Application of the Array to Adipocyte Fat Extracts

2.15.4.1 *Preparation of the fat samples*

To the fat was added 500 μL of ethyl acetate and 500 μL of brine. The solution was shaken for 5 min and then centrifuged at 5,000 \times g for 30 min. The ethyl acetate layer was removed by pipette, filtered, and placed in a weighed vial. Another 500 μL was added to

the brine, and the process was repeated two more times. The combined ethyl acetate layers were evaporated thoroughly by drying on the high vacuum line overnight. The weights of the fat samples were determined carefully. The presence of glyceride could be confirmed by TLC on a silica plate with 1:1 ethyl acetate/hexanes. The R_f of triglycerides was about 0.8.

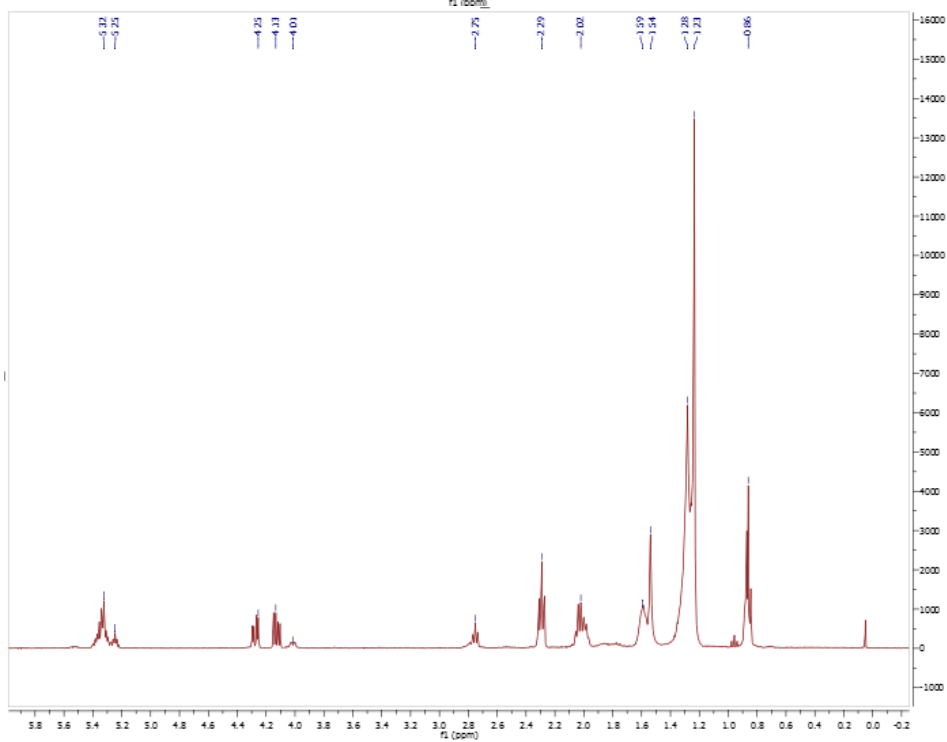
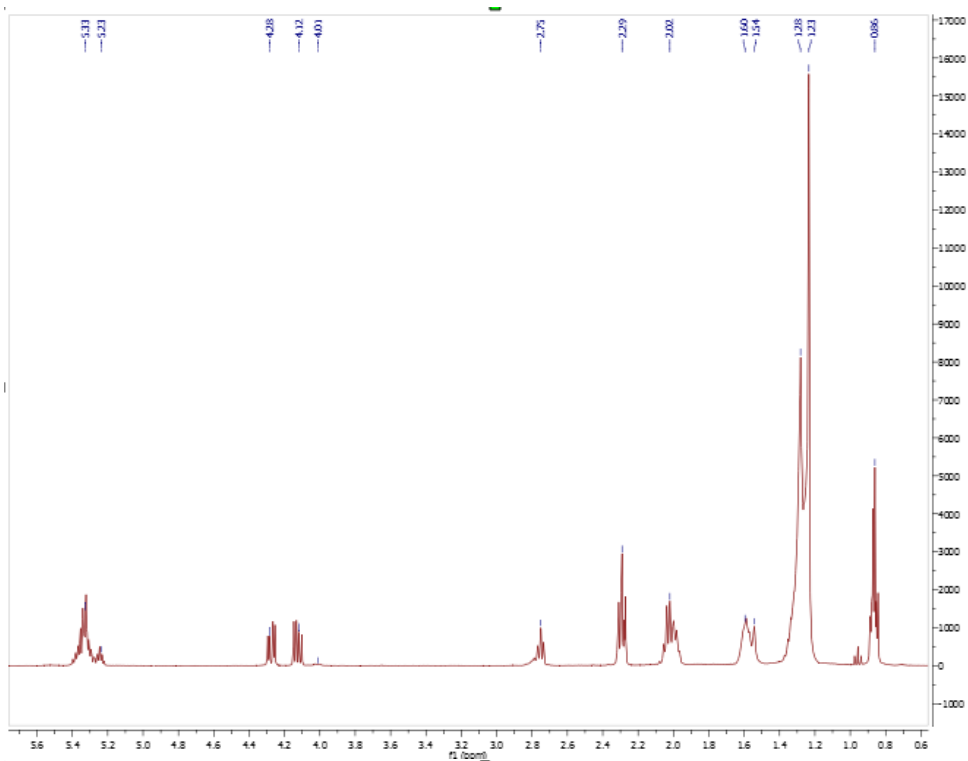
2.15.4.2 *Use of the fat in the array*

The human fat was dissolved in THF at a concentration of 4.8 mg/mL (\approx 6 mM) for the non-metathesis part of the array and in CHCl_3 at a concentration of 7.2 mg/mL (\approx 9 mM) for the metathesis reaction. In the array, the final concentration of human fat in the well plates was 0.072 mg/mL (\approx 90 μM). The other array components were used at the same molar concentrations as already indicated.

For the first analysis, the rat fat was used at the same concentrations as the human fat. For the subsequent analyses, the final concentration of the rat fat in the well plates was 0.6 mg/mL.

2.15.4.3 *$^1\text{H-NMR}$ Spectra of the Fat Samples*

The samples were dissolved in CDCl_3 .



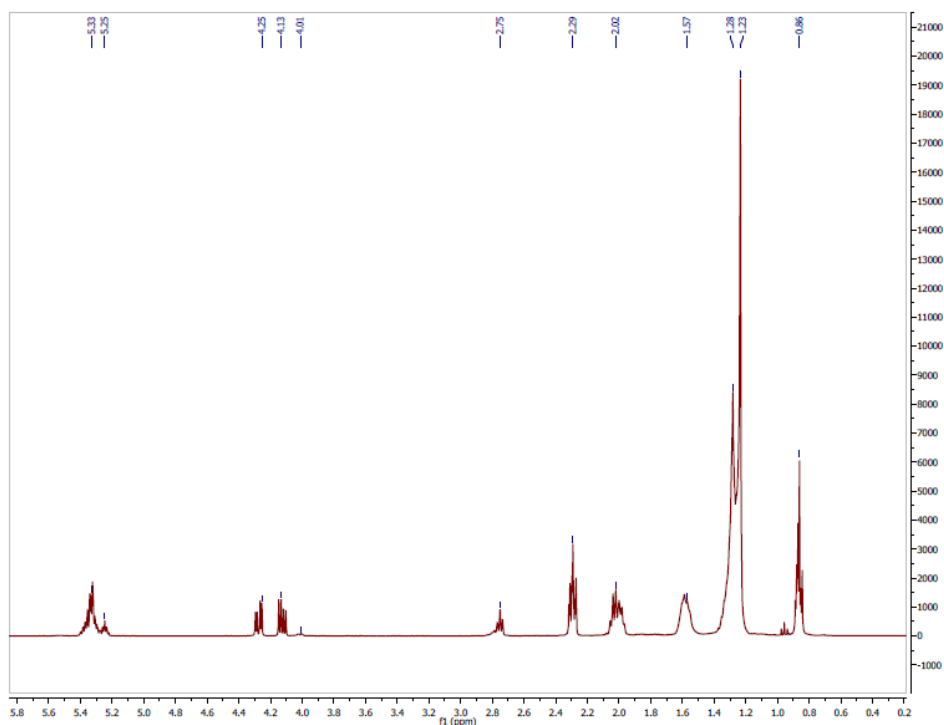


Figure 2.38 $^1\text{H-NMR}$ spectra of control (top), insulin-treated (middle), and isoproterenol (bottom) rat fat samples in CDCl_3 .

2.15.5 Quantitation of Trilinolein in Triglyceride Mixtures

2.15.5.1 Preparation of the Glyceride Mixtures

Mixtures of triglycerides were prepared by weighing out each component in a vial, dissolving each in chloroform, and consolidating the component solutions into a single weighed vial. The chloroform was removed under vacuum.

Table 2.9 Composition of glyceride mixtures by weight

Mixture	% T1 by weight	% T3 by weight	% T5 by weight	% T6 by weight	% T7 by weight	Total weight of glyceride material (mg)
A	11.61	26.47	22.13	16.01	23.79	17.49
B	7.74	17.95	31.63	24.00	18.67	34.87
C	0	16.34	52.81	30.86	0	23.69
D	0	32.37	41.88	25.75	0	25.24

2.15.5.2 *Standard Addition Procedure*

A fixed amount of each mixture was used for each analyte solution/metathesis reaction well. To the constant quantity of mixture in each vial/metathesis reaction well was added increasing amounts of pure **T5** as described in the text. These solutions were then subjected to the same conditions as described for parts 1 and 2 of the array. The final concentration of the glyceride mixture in the well plates of the array is given in **Table 2.6** as well as the final concentration within that mixture of **T5**. The content of the other triglyceride components is detailed in **Table 2.9**. For the additions, the final amount of **T5** added was 0, 0.005, 0.01, 0.02, 0.03, 0.04, 0.05, and 0.06 mg/mL for mixtures A and B and 0, 0.003, 0.005, 0.01, 0.015, 0.02, 0.025, and 0.03 mg/mL for mixtures C and D. The raw fluorescence data was normalized by rescaling within each variable from values 0-100 using XLSTAT. The eight replicates collected for each addition were averaged after normalization, and this averaged data matrix was subjected to the SANAS calculation in MATLAB.

2.16 **ADDITIONAL FIGURES**

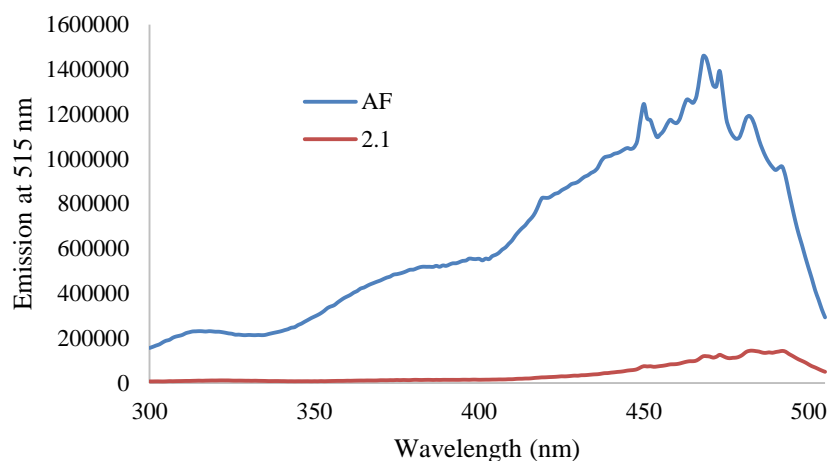


Figure 2.39 *Excitation spectra for AF and 2.1.* AF (30 μM) and 2.1 (30 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 ; $\lambda_{\text{em}} = 515 \text{ nm}$. (Reproduced from Ref. 50: © University of Texas at Austin, 2013)

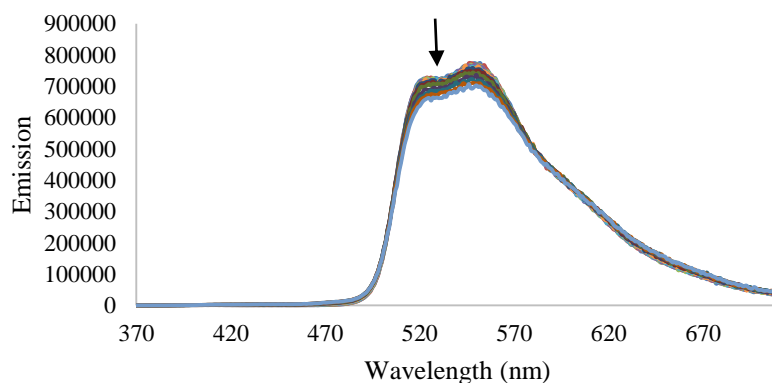


Figure 2.40 *Titration of DNSA into AF/BSA.* Addition of DNSA (0–253 μM) to AF (100 μM) and BSA (100 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 ; $\lambda_{\text{ex}} = 360 \text{ nm}$.

2.17 REFERENCES

- (1) Voet, D.; Voet, J. G.; Pratt, C. W. *Fundamentals of Biochemistry*; 3rd ed.; Wiley & Sons: New York, 2008; p. Chapter 9.
- (2) Andrikopoulos, N. K. *Crit. Rev. Food Sci. Nutr.* **2002**, *42*, 473–505.
- (3) Alemany, L. B. *Chem. Phys. Lipids* **2002**, *120*, 33–44.

- (4) *Oils and Fats Authentication*; Jee, M., Ed.; CRC Press: Boca Raton, FL, 2002.
- (5) Larsson, K. In *The Lipid Handbook*; Gunstone, F. D.; Hawood, J. L.; Padley, F. B., Eds.; Chapman and Hall: London, 1986; pp. 321–384.
- (6) Small, D. M. *Annu. Rev. Nutr.* **1991**, *11*, 413–434.
- (7) Kovac, A.; Scheib, H.; Pleiss, J.; Schmid, R. D.; Paltauf, F. *Eur. J. Lipid Sci. Technol.* **2000**, *102*, 61–77.
- (8) Warwel, S.; Borgdorf, R.; Brühl, L. *Biotechnol. Lett.* **1999**, *21*, 431–436.
- (9) Rodriguez, J. A.; Ben Ali, Y.; Abdelkafi, S.; Mendoza, L. D.; Leclaire, J.; Fotiadu, F.; Buono, G.; Carrière, F.; Abousalham, A. *Biochim. Biophys. Acta* **2010**, *1801*, 77–83.
- (10) Kritchevsky, D. *J. Nutr. Biochem.* **1995**, *6*, 172–178.
- (11) Raclot, T.; Holm, C.; Langin, D. *J. Lipid Res.* **2001**, *42*, 2049–2057.
- (12) Aluyor; O., E.; Ozigagu; E., C.; Oboh; I., O.; *P. Sci. Res. Essays* **2009**, *4*, 191–197.
- (13) Murphy, R. C.; Leiker, T. J.; Barkley, R. M. *Biochim. Biophys. Acta* **2011**, *1811*, 776–783.
- (14) Murphy, R. C.; Fiedler, J.; Hevko, J. *Chem. Rev.* **2001**, *101*, 479–526.
- (15) Li, X.; Evans, J. J. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 2528–2538.
- (16) Murphy, R. C.; Gaskell, S. J. *J. Biol. Chem.* **2011**, *286*, 25427–25433.
- (17) Thomas, M. C.; Mitchell, T. W.; Harman, D. G.; Deeley, J. M.; Murphy, R. C.; Blanksby, S. J. *Anal. Chem.* **2007**, *79*, 5013–5022.
- (18) Podlaha, O.; Töregård, B. *J. Chromatogr. A* **1989**, *482*, 215–226.
- (19) Adlof, R. *J. Chromatogr. A* **2007**, *1148*, 256–259.
- (20) Lévêque, N. L.; Héron, S.; Tchapla, A. *J. Mass Spectrom.* **2010**, *45*, 284–296.
- (21) Dobson, G.; Christie, W. W.; Nikolova-Damyanova, B. *J. Chromatogr. B Biomed. Sci. Appl.* **1995**, *671*, 197–222.

- (22) Brandão, L. F. P.; Braga, J. W. B.; Suarez, P. A. Z. *J. Chromatogr. A* **2012**, *1225*, 150–157.
- (23) Benitez-Sanchez, P. L.; Leon-Camacho, M.; Aparicio, R. *Eur. Food Res. Technol.* **2003**, *218*, 13–19.
- (24) Collins, B. E.; Wright, A. T.; V., A. E. *Top. Curr. Chem.* **2007**, *277*, 181–218.
- (25) Stitzel, S. E.; Aernecke, M. J.; Walt, D. R. *Annu. Rev. Biomed. Eng.* **2011**, *13*, 1–25.
- (26) De, M.; Rana, S.; Akpınar, H.; Miranda, O. R.; Arvizo, R. R.; Bunz, U. H. F.; Rotello, V. M. *Nat. Chem.* **2009**, *1*, 461–465.
- (27) Zhou, H.; Baldini, L.; Hong, J.; Wilson, A. J.; Hamilton, A. D. *J. Am. Chem. Soc.* **2006**, *128*, 2421–2425.
- (28) Diehl, K. L.; Anslyn, E. V. *Chem. Soc. Rev.* **2013**, *42*, 8596–8611.
- (29) Zhang, Y.; Askim, J. R.; Zhong, W.; Orlean, P.; Suslick, K. S. *Analyst* **2014**, *139*, 1922–1928.
- (30) Gunstone, F. D. *The Chemistry of Oils and Fats: Sources, Compositions, Properties, and Uses*; CRC Press: Boca Raton, FL, 2004.
- (31) Peters, T. *All About Albumin: Biochemistry, Genetics, and Medical Applications*; Academic Press: San Diego, CA, 1996.
- (32) Bujacz, A. *Acta Crystallogr. D. Biol. Crystallogr.* **2012**, *68*, 1278–1289.
- (33) Carter, D. C.; Ho, J. X. *Adv. Protein Chem.* **1994**, *45*, 153–203.
- (34) Squire, P. G.; Moser, P.; O’Konski, C. T. *Biochemistry* **1968**, *7*, 4261–4272.
- (35) Reed, R. G.; Feldhoff, R. C.; Peters, T. *Biochemistry* **1976**, *15*, 5394–5398.
- (36) Park, K. K.; Park, J. W.; Hamilton, A. D. *Org. Biomol. Chem.* **2009**, *7*, 4225–4232.
- (37) Richieri, G. V.; Anel, A.; Kleinfeld, A. M. *Biochemistry* **1993**, *32*, 7574–7580.
- (38) Roda, A.; Cappelleri, G.; Aldini, R.; Roda, E.; Barbara, L. *J. Lipid Res.* **1982**, *23*, 490–495.

- (39) Yates, F. E.; Urquhart, J. *Physiol. Rev.* **1962**, *42*, 359–433.
- (40) Ramsey, B. L.; Westphal, U. *Biochim. Biophys. Acta* **1978**, *529*, 115–122.
- (41) Pearlman, W. H.; Crepy, O. *J. Biol. Chem.* **1967**, *242*, 182–189.
- (42) Brodersen, R. *CRC Crit. Rev. Clin. Lab. Sci.* **1980**, *11*, 305–399.
- (43) Sudlow, G.; Birkett, D. J.; Wade, D. N. *Mol. Pharmacol.* **1975**, *11*, 824–832.
- (44) Sudlow, G.; Birkett, D. J.; Wade, D. N. *Mol. Pharmacol.* **1976**, *12*, 1052–1061.
- (45) Er, J. C.; Vendrell, M.; Tang, M. K.; Zhai, D.; Chang, Y.-T. *ACS Comb. Sci.* **2013**, *15*, 452–457.
- (46) Thumser, A. E. A.; Buckland, A. G.; Wilton, D. C. *J. Lipid Res.* **1998**, *39*, 1033–1038.
- (47) Kubarych, C. J.; Adams, M. M.; Anslyn, E. V. *Org. Lett.* **2010**, *12*, 4780–4783.
- (48) Adams, M. M.; Anslyn, E. V. *J. Am. Chem. Soc.* **2009**, *131*, 17068–17069.
- (49) Ivy, M. A.; Gallagher, L. T.; Ellington, A. D.; Anslyn, E. V. *Chem. Sci.* **2012**, *3*, 1773.
- (50) Ivy, M. A. Differential Sensing of Hydrophobic Analytes with Serum Albumin, 2012.
- (51) Petry, S.; Ben Ali, Y.; Chahinian, H.; Jordan, H.; Kleine, H.; Müller, G.; Carrière, F.; Abousalham, A. *J. Lipid Res.* **2005**, *46*, 603–614.
- (52) Mol, J. C. *Green Chem.* **2002**, *4*, 5–13.
- (53) Biermann, U.; Meier, M. A. R.; Butte, W.; Metzger, J. O. *Eur. J. Lipid Sci. Technol.* **2011**, *113*, 39–45.
- (54) Kirkland, T. A.; Lynn, D. M.; Grubbs, R. H. *J. Org. Chem.* **1998**, *63*, 9904–9909.
- (55) Hurd, C. D.; Schmerling, L. *J. Am. Chem. Soc.* **1937**, *59*, 112–117.
- (56) Corrie, J. E. T.; Trentham, D. R. *J. Chem. Soc. Perkin Trans. 1* **1995**, 1993.
- (57) Voigtritter, K.; Ghorai, S.; Lipshutz, B. H. *J. Org. Chem.* **2011**, *76*, 4697–4702.

- (58) Fairclough, R. H.; Cantor, C. R. *Methods Enzymol.* **1978**, *48*, 347–379.
- (59) Ouadahi, K.; Sbgoud, K.; Allard, E.; Larpent, C. *Nanoscale* **2012**, *4*, 727–732.
- (60) Schiano, V.; Laurenzano, E.; Brevetti, G.; De Maio, J. I.; Lanero, S.; Scopacasa, F.; Chiariello, M. *Clin. Nutr.* **2008**, *27*, 241–247.
- (61) Rhee, E. P.; Cheng, S.; Larson, M. G.; Walford, G. A.; Lewis, G. D.; McCabe, E.; Yang, E.; Farrell, L.; Fox, C. S.; O'Donnell, C. J.; Carr, S. A.; Vasan, R. S.; Florez, J. C.; Clish, C. B.; Wang, T. J.; Gerszten, R. E. *J. Clin. Invest.* **2011**, *121*, 1402–1411.
- (62) Cermenati, G.; Abbiati, F.; Cermenati, S.; Brioschi, E.; Volonterio, A.; Cavaletti, G.; Saez, E.; De Fabiani, E.; Crestani, M.; Garcia-Segura, L. M.; Melcangi, R. C.; Caruso, D.; Mitro, N. *J. Lipid Res.* **2012**, *53*, 300–310.
- (63) Guilherme, A.; Virbasius, J. V.; Puri, V.; Czech, M. P. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 367–377.
- (64) Grousse, A.; Langin, D. *Int. J. Obes. (Lond)*. **2012**, *36*, 581–594.
- (65) Large, V.; Arner, P. *Diabetes Metab.* **1998**, *24*, 409–418.
- (66) Brasaemle, D. L. *J. Lipid Res.* **2007**, *48*, 2547–2559.
- (67) Kraemer, F. B. *J. Lipid Res.* **2002**, *43*, 1585–1594.
- (68) Eichmann, T. O.; Kumari, M.; Haas, J. T.; Farese, R. V.; Zimmermann, R.; Lass, A.; Zechner, R. *J. Biol. Chem.* **2012**, *287*, 41446–41457.
- (69) Hazel, J. R.; Sidell, B. D. *J. Exp. Biol.* **2004**, *207*, 897–903.
- (70) Skoog, D. A.; Holler, J. F.; Crouch, S. R. *Principles of Instrumental Analysis*; 6th ed.; Thomson Brooks/Cole: Belmont, CA, 2007.
- (71) Saxberg, B. E. H.; Kowalski, B. R. *Anal. Chem.* **1979**, *51*, 1031–1038.
- (72) Herrero, A.; Cruz Ortiz, M.; Arcos, J.; López-Palacios, J.; Sarabia, L. *Anal. Chim. Acta* **1994**, *293*, 277–293.
- (73) Sena, M. M.; Trevisan, M. G.; Poppi, R. J. *Talanta* **2006**, *68*, 1707–1712.

- (74) Lozano, V. A.; Ibañez, G. A.; Olivieri, A. C. *Anal. Chim. Acta* **2009**, *651*, 165–172.
- (75) Yousefinejad, S.; Hemmateenejad, B. *Drug Test. Anal.* **2012**, *4*, 507–514.
- (76) Mohseni, N.; Bahram, M.; Olivieri, A. C. *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* **2014**, *122*, 721–730.
- (77) Karimi, M. A.; Mazloum Ardakani, M.; Behjatmanesh Ardakani, R.; Mashhadizadeh, M. H.; Zadeh, N. Z. *Anal. Bioanal. Electrochem.* **2009**, *1*, 142–158.
- (78) Hemmateenejad, B.; Yousefinejad, S. *Anal. Bioanal. Chem.* **2009**, *394*, 1965–1975.
- (79) Lorber, A.; Faber, K.; Kowalski, B. R. *Anal. Chem.* **1997**, *69*, 1620–1626.
- (80) Lorber, A. *Anal. Chem.* **1986**, *58*, 1167–1172.
- (81) Varmuza, K.; Filzmoser, P. *Introduction to Multivariate Statistical Analysis in Chemometrics*; CRC Press: Boca Raton, FL, 2009.

Chapter 3: Development of Dynamic Supramolecular Receptors from Serum Albumin and Fatty Acid-Appended Peptides

3.1 INTRODUCTION

Due to our success with using serum albumins as receptors for hydrophobic analytes, we next considered how we could use serum albumin in a different way to develop receptors for other types of analytes. The serum albumin could function as a scaffold onto which species dynamically assemble through hydrophobic interactions with the protein. These species would contain two parts: a hydrophobic portion to bind to serum albumin and a recognition unit to bind to the analyte. In this way, the association of the recognition units with the scaffold would be reversible. An advantage of dynamic receptors is that they can be templated to an analyte.¹⁻⁴ In other words, the scaffold and a variety of recognition units are simply combined in solution, and in the presence of an analyte, the assembly or assemblies that bind the analyte best are selected for upon thermodynamic equilibration. We envisioned targeting large, multivalent analytes such as proteins or cell surfaces (Chapter 1) with these receptors, so we wanted the potential for multiple recognition units to assemble on a single protein scaffold.

To explore these ideas, we hypothesized that we could build dynamic assemblies from serum albumin and fatty acid-appended peptides. The fatty acid portion binds in the hydrophobic pockets of the protein,⁵ leaving the peptidic portion protruding out into solution (**Figure 1**). Human serum albumin has been shown to have at least seven binding sites for long-chain fatty acids (LCFAs),⁶ so there is the potential for multiple fatty acid-peptides to bind. Since the association of the fatty acids with the protein is non-covalent and reversible, the assembly can rearrange freely until thermodynamic equilibrium is reached.

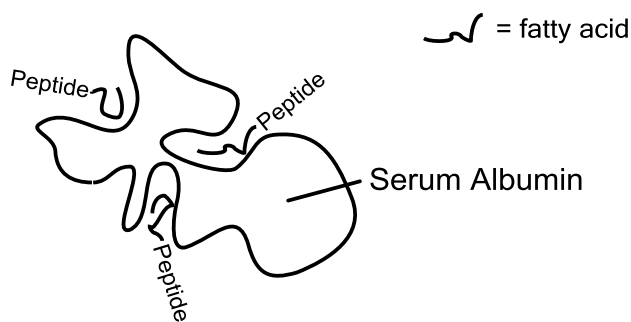


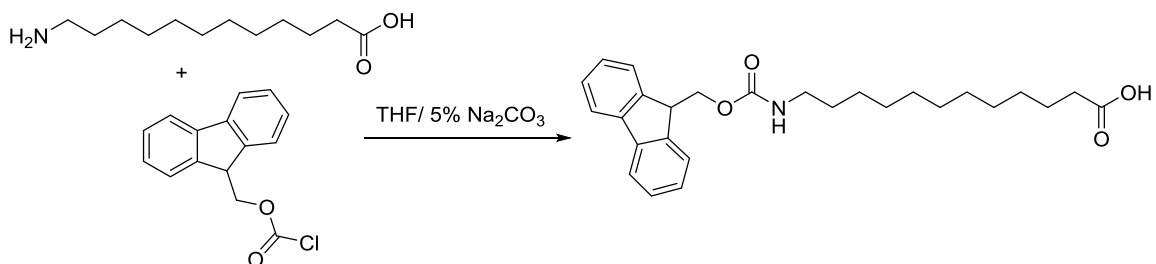
Figure 3.1 *Cartoon representation of the fatty acid-peptides assembled on the serum albumin scaffold.*

For the recognition unit part of the conjugate, peptides were chosen because they are easy to synthesize via solid-phase peptide synthesis (SPPS) and can be appended to compounds with an amine or carboxylic acid. Furthermore, previous work in our group toward the development of receptors built from a DNA scaffold used pentapeptides conjugated to a DNA intercalator, so a similar concept had already been tested with peptide-based species.⁷ The same peptide motifs from the DNA study were used for initial testing of the serum albumin-based assembly. These motifs were selected in the DNA study because they were known in the literature to bind to specific cancer cell types or cancer cell surface features, and the goal of that project was to make receptors to target cancer cell surfaces. However, in this work, we simply wanted to begin by determining if we could build assemblies from serum albumin and fatty acids that were attached to some kind of generic recognition unit. Hence, the choice of peptide was somewhat arbitrary at this point in the development of the receptors. We postulated that, if the components assembled as we hypothesized, then a variety of different kinds of recognition units could be attached to the fatty acids.

3.2 SYNTHESIS OF FATTY ACID-PEPTIDES

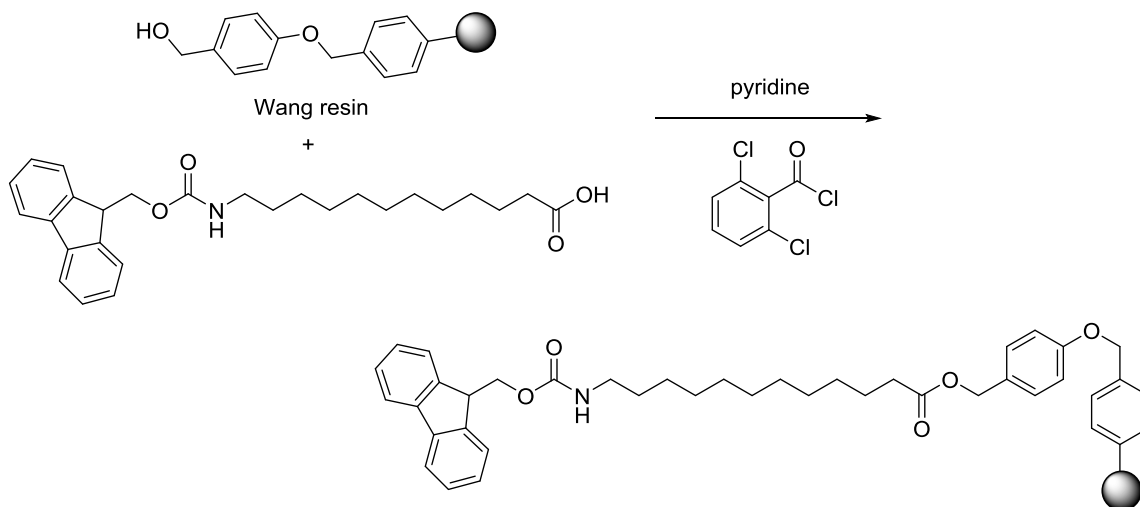
3.2.1 Attaching the Fatty Acid to the Resin

First, 12-aminododecanoic acid was Fmoc protected according to the literature procedure (Scheme 3.1).⁸ This medium-chain fatty acid (MCFA) was chosen because it is commercially available with the terminal amino group.



Scheme 3.1 *Fmoc protection of 12-aminododecanoic acid.*

Next, the Fmoc-protected fatty acid was coupled to a Wang resin by the Sieber method (Scheme 3.2).⁹

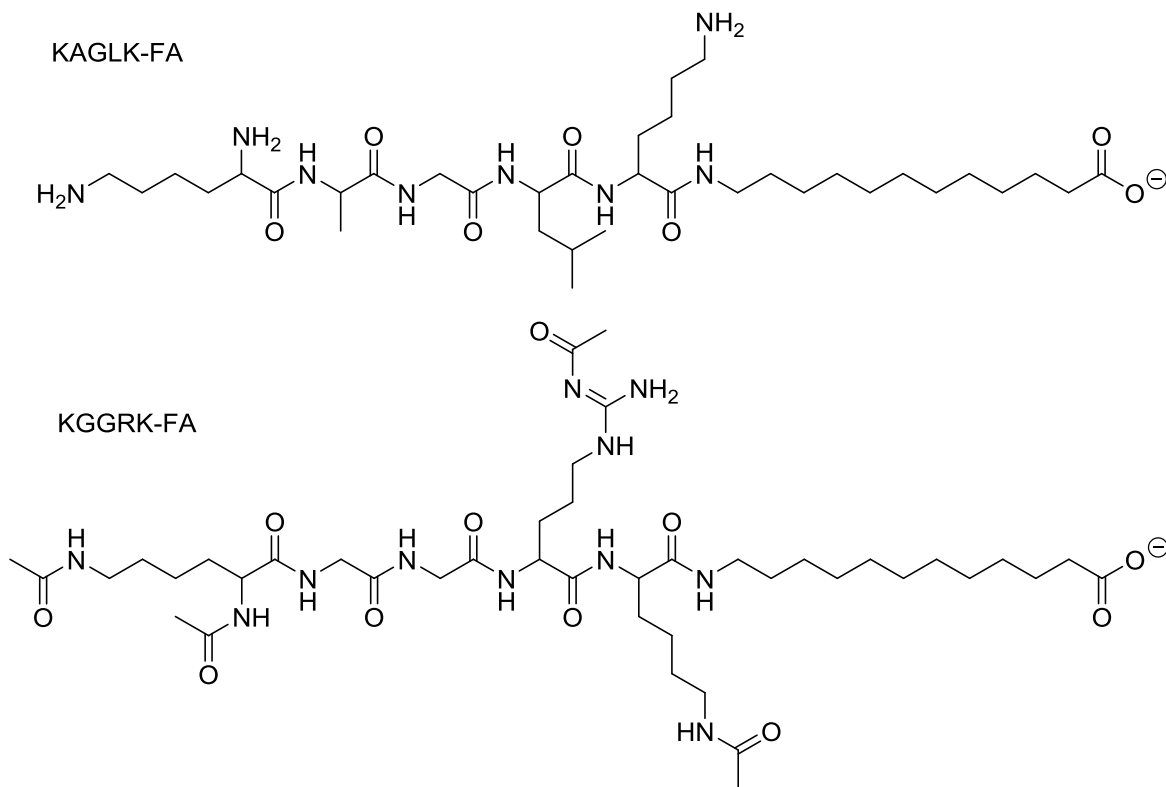


Scheme 3.2 *Loading of Fmoc-protected 12-aminododecanoic acid to Wang resin.*

The loading of the fatty acid on the resin was determined to be 0.6 mmol/g by quantitating the Fmoc content with absorbance spectroscopy.¹⁰

3.2.2 Synthesis of the Fatty Acid-Peptide Conjugates

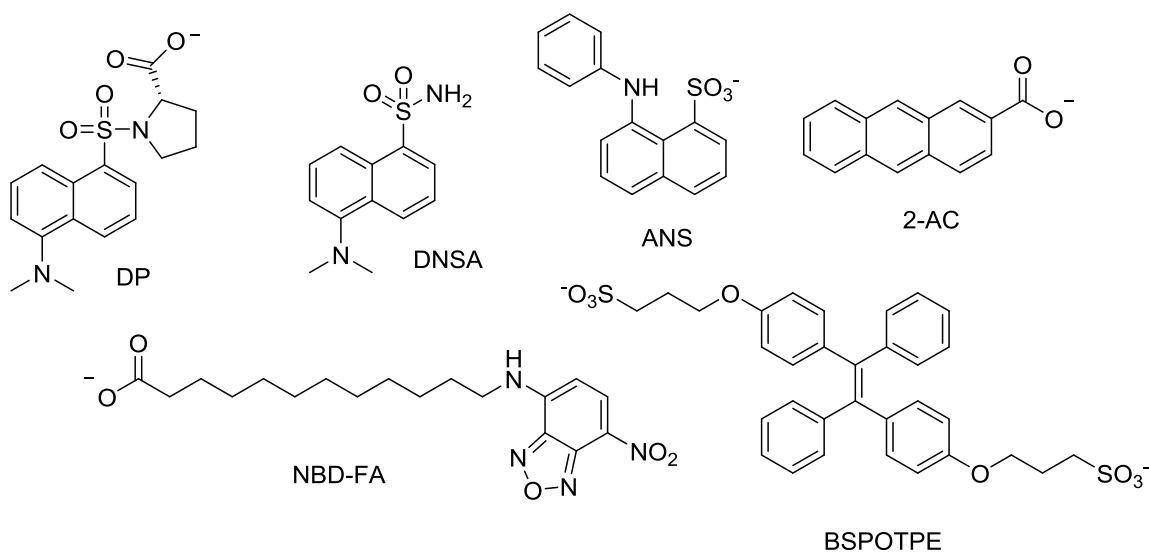
The peptides were built on the fatty acid-Wang resin using SPPS (**Scheme 3.3**). The peptides were cleaved from the Wang resin. After cleavage from the resin and deprotection of the side chains, ACKGGRK-FA was acetylated with acetic anhydride in order to try to control for the effect of hydrophobicity of the peptide on binding to serum albumin. The crude material was purified by reverse-phase HPLC (RP-HPLC).



Scheme 3.3 Structures of the fatty acid-peptides.

3.3 BINDING STUDIES OF FATTY ACID-PEPTIDES TO SERUM ALBUMIN

Once we had synthesized the fatty acid-conjugated peptides, we next performed a series of titrations to determine if the conjugates were binding to serum albumin. Since the conjugates do not contain a chromophore, we used indicator displacement and uptake experiments to measure binding indirectly with several fluorescent probes (**Scheme 3.4**).



Scheme 3.4 Structures of the fluorescent probes used in this work. DP = dansyl proline, DNSA = dansyl amide, ANS = 1,8-anilinonaphthalene sulfonate, AC = 2-anthracenecarboxylate, NBD-FA = nitrobenzofurazan-fatty acid, and BSPOTPE = 1,2-Bis[4-(3-sulfonatopropoxy)phenyl]-1,2-diphenylethene.

3.3.1 Indicator Displacement and Uptake Titrations with NBD-FA

The first fluorophore used to try to detect binding of KAGLK-FA to serum albumin was NBD-FA (see Chapter 2 for synthesis). NBD-FA is also derived from 12-aminododecanoic acid, so we postulated that it would bind similarly to the fatty acid-peptides. KAGLK-FA was titrated into a mixture of BSA and NBD-FA, and the fluorescence of the NBD-FA was monitored (**Figure 3.2**).

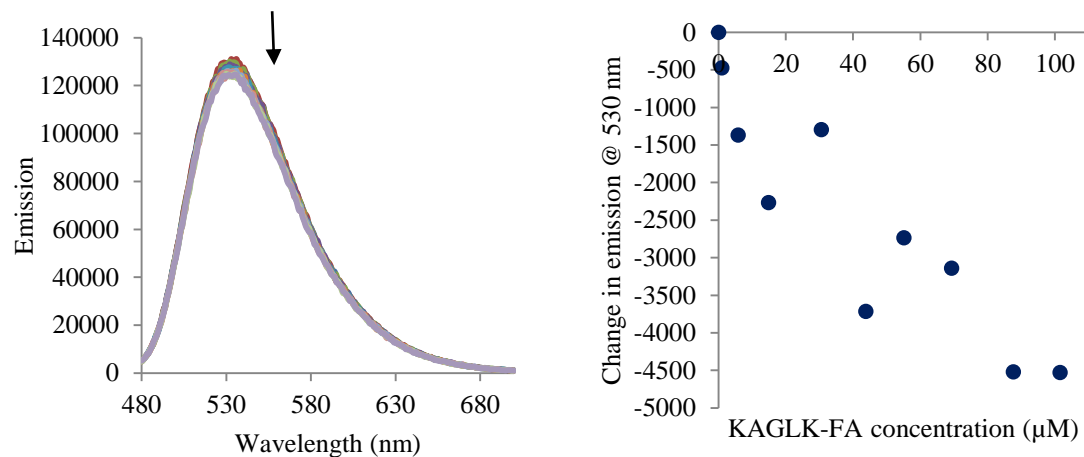


Figure 3.2 *Displacement of NBD-FA from BSA by KAGLK-FA.* Left: Addition of KAGLK-FA (0-101 μM) to BSA (33 μM) and NBD-FA (20 μM) in 10 mM phosphate buffer, 140 mM NaCl, H₂O, pH 7.40, 0.02% NaN₃, λ_{ex} = 470 nm. Right: Corresponding binding isotherm for λ_{em} = 530 nm.

The emission of NBD-FA decreased as the fatty acid-peptide was added, which is indicative of displacement of the fluorophore from the serum albumin. However, the modulation of the emission was very small. This small change could be indicative of very little displacement occurring or of some direct effect that KAGLK-FA was having on NBD-FA. Thus, we performed an indicator uptake experiment with NBD-FA to attempt to clarify these results. In this indicator uptake experiment, we compared the addition of NBD-FA to BSA to the addition of NBD-FA to BSA/KAGLK-FA to see if the presence of KAGLK-FA inhibits binding of NBD-FA to BSA (**Figure 3.3**).

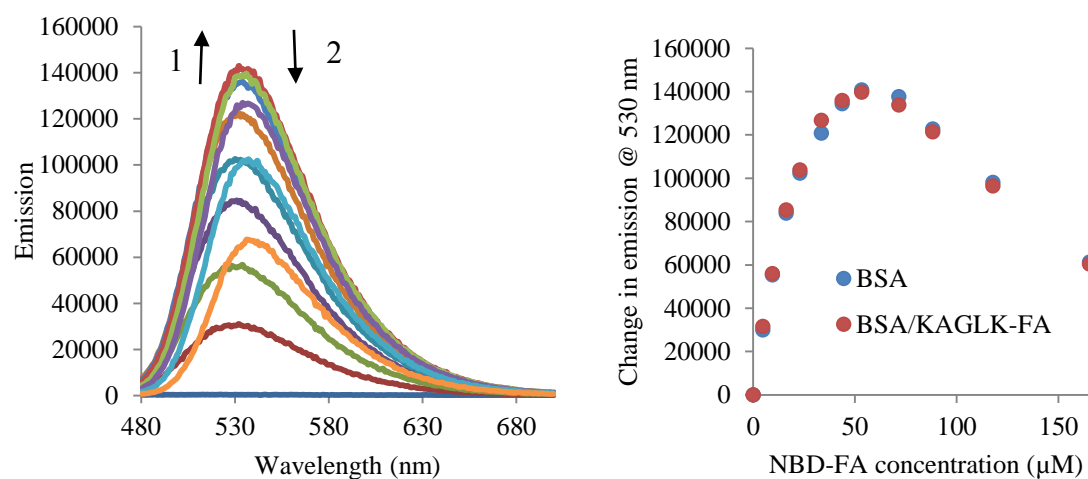


Figure 3.3 *Indicator uptake titrations with NBD-FA, BSA, and KAGLK-FA.* Left: Addition of NBD-FA (0-165 μM) to BSA (100 μM) in 10 mM phosphate buffer, 140 mM NaCl, H_2O , pH 7.40, 0.02% NaN_3 , $\lambda_{\text{ex}} = 470 \text{ nm}$. Right: Corresponding binding isotherm overlaid with isotherm from addition of NBD-FA (0-165 μM) to BSA (100 μM) and KAGLK-FA (90 μM), $\lambda_{\text{ex}} = 470 \text{ nm}$ and $\lambda_{\text{em}} = 530 \text{ nm}$.

The same self-quenching behavior of NBD-FA that was observed in other indicator uptake experiments (**Figure 2.6**) was also seen here. There was essentially no difference in the uptake of NBD-FA in the absence or presence of KAGLK-FA. This result indicates that the fatty acid-peptide is not binding to BSA or that it is binding to an extent and manner that does not affect the binding of NBD-FA. We hypothesized that the peptidic portion of the conjugate was inhibiting binding of the fatty acid portion due to its positive charge, and it was at this point that we acetylated the crude KGGRK-FA to block the amine and guanidine groups to test this postulate. The indicator uptake experiment was repeated with the acetylated KGGRK-FA (**Figure 3.4**).

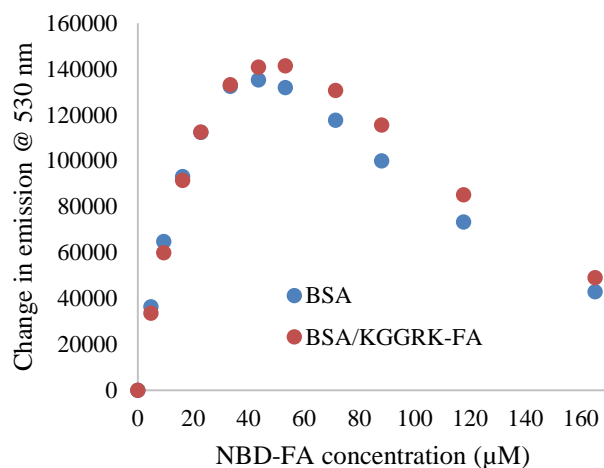


Figure 3.4 *Indicator uptake titrations with NBD-FA, BSA, and KGGRK-FA.* Addition of NBD-FA (0-165 μM) to BSA (100 μM) and addition of NBD-FA (0-165 μM) to BSA (100 μM) and KGGRK-FA (90 μM) in 10 mM phosphate buffer, 140 mM NaCl, H_2O , pH 7.40, 0.02% NaN_3 , $\lambda_{\text{ex}} = 470 \text{ nm}$ and $\lambda_{\text{em}} = 530 \text{ nm}$.

Again, the fatty acid-peptide did not significantly inhibit the uptake of NBD-FA by BSA. In fact, the uptake of NBD-FA appeared to be enhanced slightly by the presence of the acetylated KGGRK-FA. According to these experiments, the fatty acid-peptides were not binding to BSA. In order to ensure that these experiments with NBD-FA could provide evidence for ligand binding to BSA at all, we performed two positive controls. The uptake of NBD-FA by BSA was measured in the presence of palmitic acid and of monopalmitin, which are both known to bind to serum albumin (**Figure 3.5**).⁵

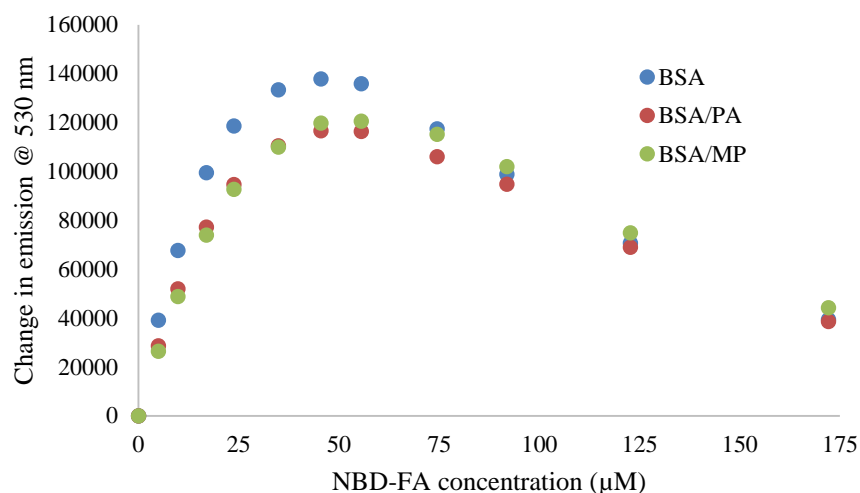


Figure 3.5 *Indicator uptake titrations with NBD-FA, BSA, palmitic acid (PA), and monopalmitin (MP). Addition of NBD-FA (0-172 μM) to BSA (100 μM), of NBD-FA (0-165 μM) to BSA (100 μM) and palmitic acid (90 μM), and of NBD-FA (0-165 μM) to BSA (100 μM) and monopalmitin (90 μM) in 10 mM phosphate buffer, 140 mM NaCl, H₂O, pH 7.40, 0.02% NaN₃, λ_{ex} = 470 nm and λ_{em} = 530 nm.*

The presence of palmitic acid and of monopalmitin inhibited the uptake of NBD-FA by BSA as evidenced by the lower emission increases and maxima of those isotherms compared to isotherm for the addition of NBD-FA to BSA alone. These positive controls confirmed our expectation of how inhibition of NBD-FA uptake would manifest in these experiments. Thus, the titrations with the fatty acid-peptides were not indicative of these species binding to BSA. We performed another control experiment with 12-aminododecanoic acid, the fatty acid from which NBD-FA and the peptides are derived (**Figure 3.6**). This fatty acid itself does not inhibit the uptake of NBD-FA by BSA. Hence, the NBD moiety is contributing significantly to the affinity of NBD-FA for BSA and simply attaching other groups like peptides to 12-aminododecanoic acid will not necessarily result in good ligands for BSA.

Since palmitic acid clearly inhibited the uptake of NBD-FA while the fatty acid-peptides and 12-aminododecanoic acid did not, we concluded that the conjugates were not binding in the same site(s) as long-chain fatty acids. Medium-chain fatty acids (C₆-C₁₄) can bind in long-chain fatty acid sites of serum albumin but not nearly as strongly as long-chain fatty acids bind (MCFA $K_a = 10^{3-5} \text{ M}^{-1}$, LCFA $K_a = 10^{7-8} \text{ M}^{-1}$).¹¹ Having an amine or even an entire peptide attached to the MCFA could further diminish its binding affinity by making it more hydrophilic and positively-charged. However, the conjugates could still bind in other parts of the protein.¹² There is evidence that medium-chain fatty acids bind in Sudlow Site II.¹³ Thus, we undertook more indicator uptake experiments with different fluorophores to probe other potential modes of binding to serum albumin.

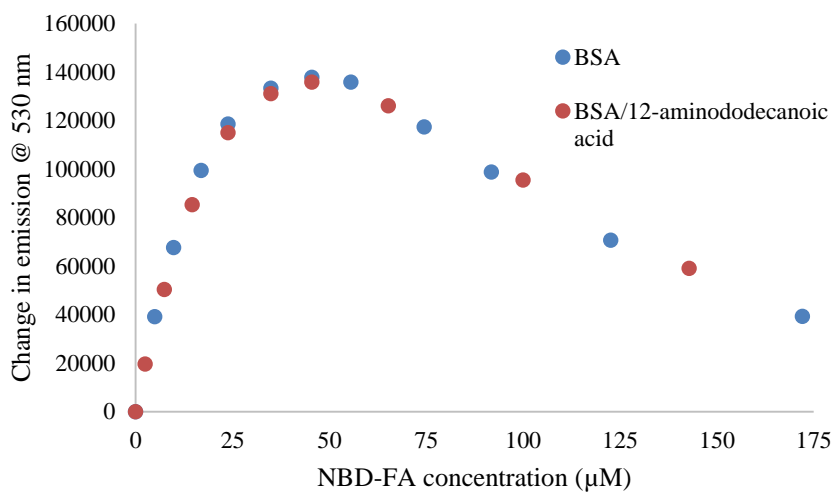


Figure 3.6 *Indicator uptake titrations with NBD-FA, BSA, and 12-aminododecanoic acid.* Addition of NBD-FA (0-172 μM) to BSA (100 μM) and of NBD-FA (0-165 μM) to BSA (100 μM) and 12-aminododecanoic acid (100 μM) in 10 mM phosphate buffer, 140 mM NaCl, H₂O, pH 7.40, 0.02% NaN₃, $\lambda_{\text{ex}} = 470 \text{ nm}$ and $\lambda_{\text{em}} = 530 \text{ nm}$.

3.3.2 Indicator Displacement and Uptake Titrations with Other Fluorophores

3.3.2.1 Dansyl Amide

Dansyl amide (DNSA) is a Sudlow Site I-specific ligand.¹⁴ Indicator uptake titrations were undertaken by adding DNSA to BSA, BSA/KGGRK-FA, and BSA/KAGLK-FA (**Figure 3.7**). An indicator displacement assay was performed by adding KAGLK-FA to DNSA/BSA (**Figure 3.8**).

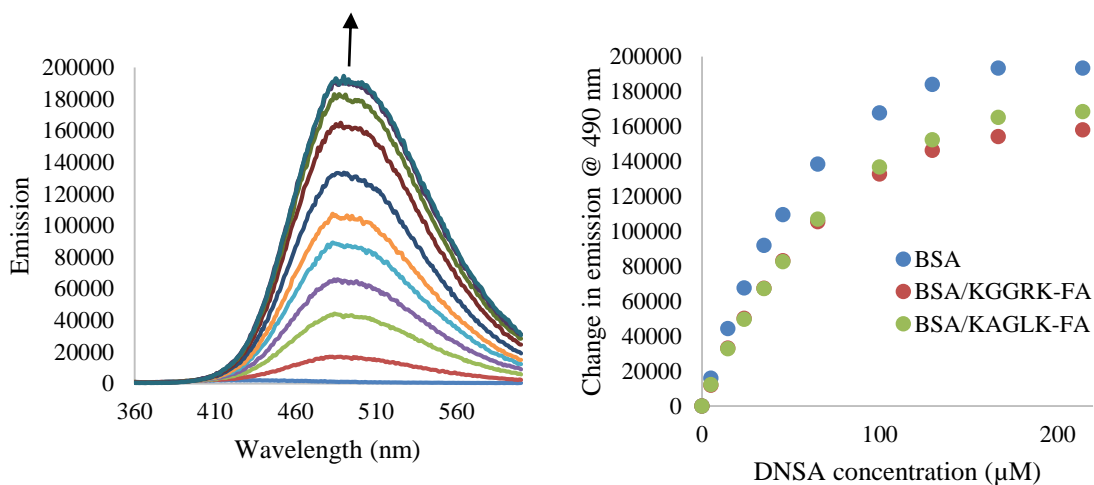


Figure 3.7 *Indicator uptake titrations with DNSA, BSA, ACKGGRK-FA, and KAGLK-FA.* Left: Addition of DNSA (0-214 μM) to BSA (100 μM), of DNSA (0-214 μM) to BSA (100 μM) and ACKGGRK-FA (90 μM), and of DNSA (0-214 μM) to BSA (100 μM) and KAGLK-FA (90 μM) in 10 mM phosphate buffer, 140 mM NaCl, H₂O, pH 7.40, 0.02% NaN₃, $\lambda_{\text{ex}} = 350 \text{ nm}$. Right: Corresponding binding isotherms for $\lambda_{\text{em}} = 490 \text{ nm}$.

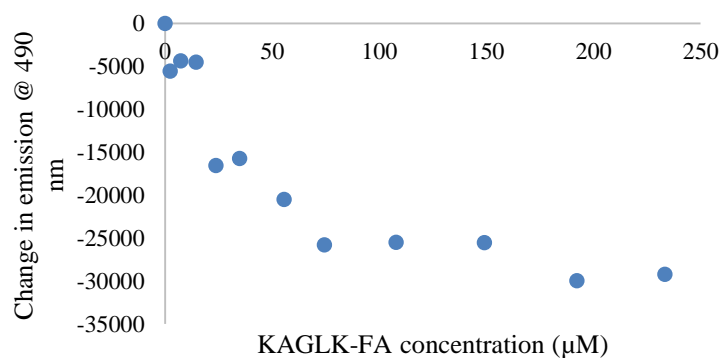


Figure 3.8 *Displacement of DNSA from BSA by KAGLK-FA.* Left: Addition of KAGLK-FA (0-233 μM) to BSA (20 μM) and DNSA (20 μM) in 10 mM phosphate buffer, 140 mM NaCl, H₂O, pH 7.40, 0.02% NaN₃, λ_{ex} = 350 nm. Right: Corresponding binding isotherm for λ_{em} = 490 nm.

The fatty acid-peptides are affecting the binding of DNSA to BSA, so perhaps the conjugates are binding in Sudlow Site I.

3.3.2.2 Dansyl Proline

Dansyl proline (DP) is a Sudlow Site II-specific ligand.¹⁴ An indicator displacement assay was performed by titrating ACKGGRK-FA into DP/BSA (**Figure 3.9**).

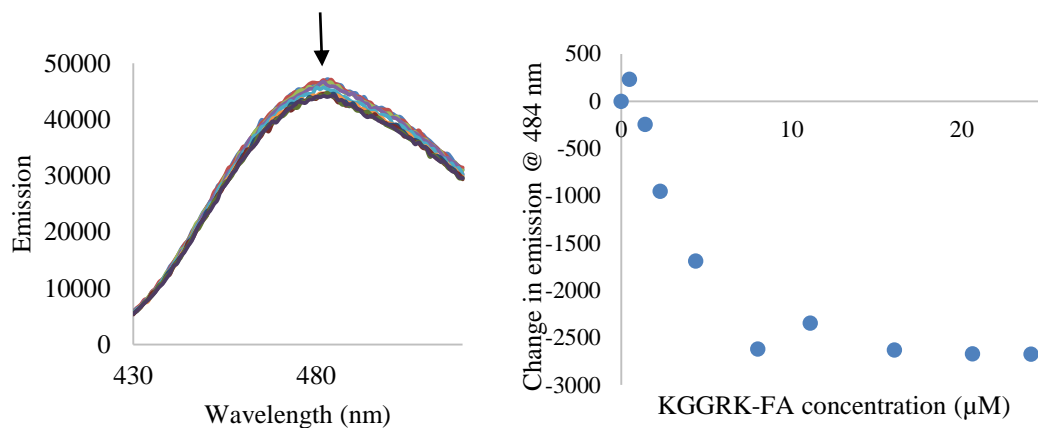


Figure 3.9 *Displacement of DP from BSA by ACKGGRK-FA.* Left: Addition of ACKGGRK-FA (0-24 μM) to BSA (10 μM) and DP (10 μM) in 10 mM phosphate buffer, 140 mM NaCl, H₂O, pH 7.40, 0.02% NaN₃, $\lambda_{\text{ex}} = 350$ nm. Right: Corresponding binding isotherm for $\lambda_{\text{em}} = 484$ nm.

The fatty acid-peptides are also displacing DP. The amount of modulation is lower than for DNSA by about an order of magnitude; however, fewer equivalents of conjugate are necessary to reach saturation, again by about an order of magnitude. Therefore, it is unclear whether DNSA or DP is being more directly affected by the fatty acid-peptides, and hence whether the conjugates are binding to Sudlow Site I, II, or both.

3.3.2.3 1,8-Anilinonaphthalene Sulfonate

ANS binds between the domains of serum albumin and is used as a probe for denaturation of the protein.¹⁵ LCFAs are known to depress ANS emission as do MCFAs, although to a lesser extent. Indicator uptake experiments were undertaken by adding ANS to BSA and to BSA/ACKGGRK-FA (**Figure 3.10**). No inhibition of ANS uptake was detected.

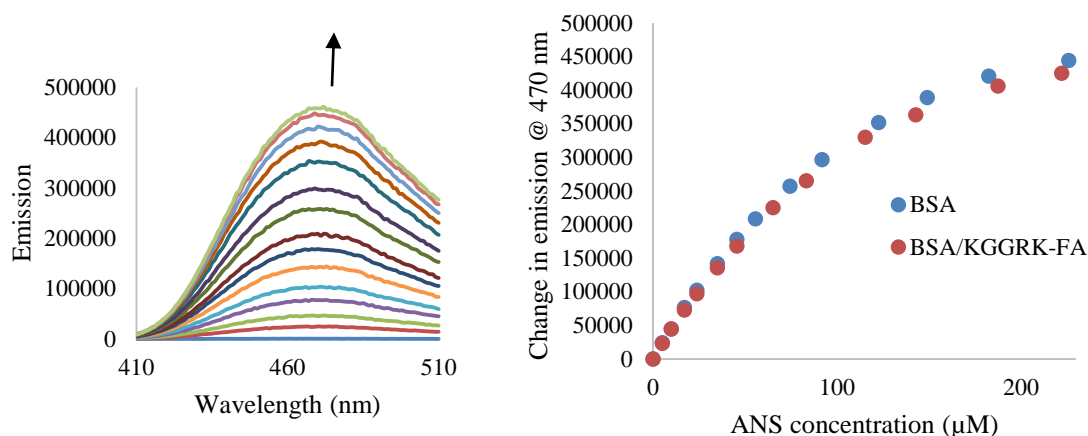


Figure 3.10 Indicator uptake titrations with ANS, BSA, and ACKGGRK-FA. Left: Addition of ANS (0-226 μM) to BSA (100 μM) and of ANS (0-226 μM) to BSA (100 μM) and ACKGGRK-FA (90 μM) in 10 mM phosphate buffer, 140 mM NaCl, H_2O , pH 7.40, 0.02% NaN_3 , $\lambda_{\text{ex}} = 350$ nm. Right: Corresponding binding isotherms for $\lambda_{\text{em}} = 490$ nm.

3.3.2.4 2-Anthracenecarboxylate

AC binds to serum albumin in four sites, including Sudlow Sites I and II and between subdomains IIA and IIB.^{16,17} Titration of KAGLK-FA into BSA/AC did not result in displacement of AC (**Figure 3.11**).

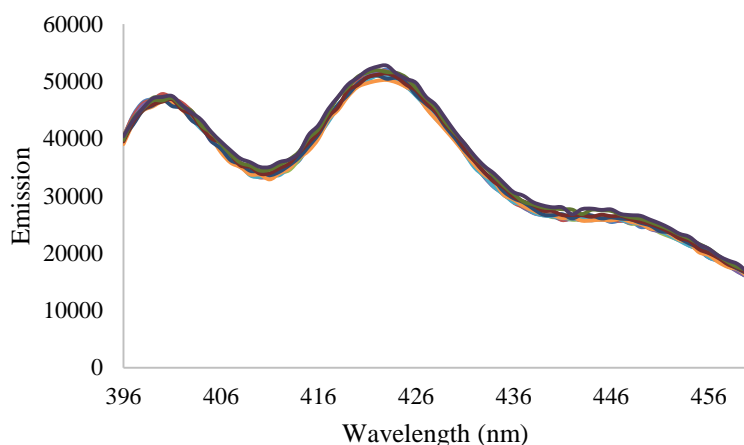


Figure 3.11 Displacement experiment of AC from BSA by KAGLK-FA. Addition of KAGLK-FA (0-200 μM) to BSA (10 μM) and AC (10 μM) in 10 mM phosphate buffer, 140 mM NaCl, H_2O , pH 7.40, 0.02% NaN_3 , $\lambda_{\text{ex}} = 386$ nm.

3.3.2.5 BSPOTPE

BSPOTPE has been shown by computational modeling to bind in the cleft between subdomains IIA and IIIA.¹⁸ Indicator uptake experiments show a small depression of binding of BSPOTPE to BSA in the presence of ACKGGRK-FA (**Figure 3.12**). Thus either the fatty acid-peptide can also bind in the cleft, or it binds elsewhere and causes allosteric changes in the protein that render the cleft a less attractive binding site for BSPOTPE.

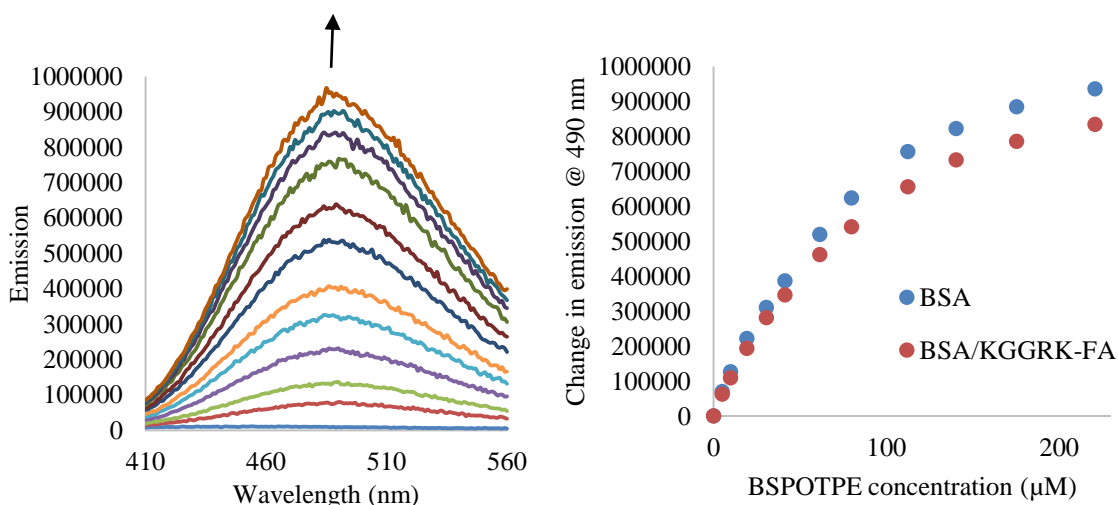


Figure 3.12 Indicator uptake titrations with BSPOTPE, BSA, and ACKGGRK-FA. Left: Addition of BSPOTPE (0-221 μM) to BSA (100 μM) and of BSPOTPE (0-221 μM) to BSA (100 μM) and ACKGGRK-FA (90 μM) in 10 mM phosphate buffer, 140 mM NaCl, H₂O, pH 7.40, 0.02% NaN₃, $\lambda_{\text{ex}} = 350$ nm. Right: Corresponding binding isotherms for $\lambda_{\text{em}} = 490$ nm.

3.3.3 Intrinsic Fluorescence

Human and bovine serum albumin both have a tryptophan residue (Trp-214) in Sudlow Site I.⁵ The intrinsic fluorescence of this tryptophan has been used to measure binding of ligands to this site and to the protein in general.¹⁹⁻²¹ Thus, we investigated the effect of the fatty-acid peptide on the tryptophan fluorescence. First, palmitic acid was titrated into BSA to determine the effect that this ligand would have on the intrinsic

fluorescence as a comparison (**Figure 3.13**). Then, ACKGGRK-FA was titrated into BSA (**Figure 3.14**).

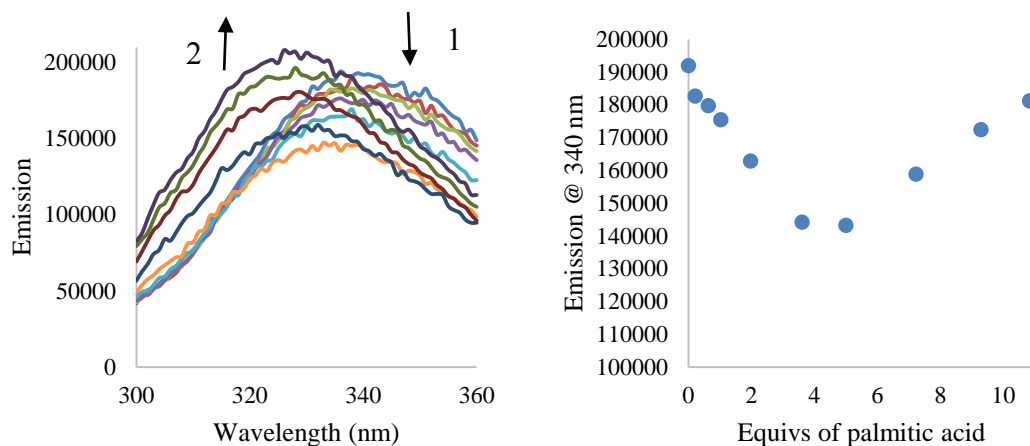


Figure 3.13 *Titration of palmitic acid into BSA monitoring intrinsic fluorescence of BSA.* Left: Addition of palmitic acid (0-33 μM) to BSA (3 μM) in 10 mM phosphate buffer, 140 mM NaCl, H_2O , pH 7.40, 0.02% NaN_3 , 0.3% THF, $\lambda_{\text{ex}} = 280$ nm. Right: Corresponding binding isotherm for $\lambda_{\text{em}} = 340$ nm.

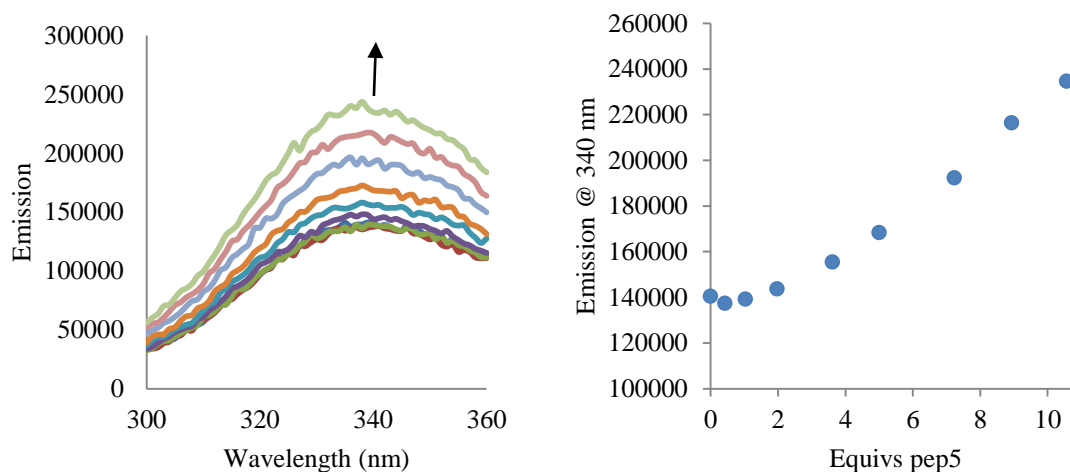


Figure 3.14 *Titration of ACKGGRK-FA into BSA monitoring intrinsic fluorescence of BSA.* Left: Addition of ACKGGRK-FA (0-33 μM) to BSA (3 μM) in 10 mM phosphate buffer, 140 mM NaCl, H_2O , pH 7.40, 0.02% NaN_3 , $\lambda_{\text{ex}} = 280$ nm. Right: Corresponding bind isotherm for $\lambda_{\text{ex}} = 340$ nm

For the addition of palmitic acid, the fluorescence is quenched at first, but then at about four equivalents, the fluorescence begins to increase. A gradual shift in the λ_{max} occurs from 340 nm to 330 nm as palmitic acid is added. For the addition of ACKGGRK-FA, the fluorescence decreases slightly upon the addition of about 0.5 equivalents of fatty acid-peptide and then increases continuously as more is added. There is no change in the λ_{max} . These results are consistent with the indicator displacement and uptake experiments in that the fatty acid-peptides do seem to be binding to some extent to serum albumin, but they are not binding to the same extent and in the same manner as LCFAs.

3.4 CONCLUSIONS AND FUTURE WORK

The goal of this project was to determine if multivalent, dynamic receptors could be made from serum albumin and fatty acid-appended recognition units. We hypothesized that multiple fatty acids would self-assemble by binding in the LCFA sites of the protein, leaving the recognition units out in solution to interact with analytes. To test this hypothesis, we synthesized fatty acid-peptides and performed a series of fluorescence experiments to measure the binding of these conjugates to bovine serum albumin. Through these experiments, we determined that the fatty acid-peptides seem to bind to the protein. However, the binding strength, stoichiometry, and mode are clearly different from the binding of LCFAs.

It is uncertain whether this difference in binding is due to the fatty acid portion of the conjugate, a MCFA, or to the hydrophilic peptide portion of the conjugate (or to both). A LCFA-appended peptide could be synthesized to test whether the fatty acid alone is the problem. Perhaps a longer-chain fatty acid would also provide a better spacer between the binding part of the conjugate and the peptide part, so that the peptide interferes less with binding. Finally, it is also possible that these conjugates, due to their hydrophobic and

hydrophilic ends, adopt stable structures in solution like micelles that diminish the driving force behind the binding of the fatty acid to serum albumin. In this case, less hydrophilic recognition units than peptides would need to be employed to try to avoid micelle formation. While there are conceivably other experiments that could be attempted to try to reach our goal, we instead decided to change direction away from fatty acids and toward squaraines (Chapter 4).

3.5 EXPERIMENTAL METHODS

3.5.1 General

Unless otherwise indicated, chemicals and reagents were obtained from Sigma Aldrich and used without further purification. The Wang resin, amino acids, and coupling reagents were obtained from NovaBioChem. Dansyl amide, dansyl proline piperidinium salt, and 2-anthracene carboxylic acid were obtained from TCI America. NBD-FA was synthesized according to the literature procedure described in Chapter 2. BSPOTPE was provided by Ben Zhong Tang (Hong Kong University of Science and Technology).

HPLC-grade solvents were prepared with 0.1% TFA (v/v) and filtered through 0.2 micron filter. HPLC was performed on a Shimadzu instrument with a preparative C-18 column using a water and acetonitrile mobile phase. LC-MS analysis was performed on an Agilent 1200 Series HPLC with an Agilent 6130 single quadrupole mass spectrometer (ESI and APCI ionization).

The titration experiments were performed with a PTI fluorimeter using an 814 photomultiplier detection system and a 75W xenon short arc lamp.

3.5.2 Synthesis

3.5.2.1 *Fmoc-Protected 12-Aminododecanoic Acid*

A 1:1 mixture of THF and 5% NaCO₃ (aq) was prepared and added to 12-aminododecanoic acid (1.5 g, 6.97 mmol, 1.0 eq). The solution was stirred for 30 min and then poured into a separatory funnel. The THF (top layer) was collected. To this solution was added fluorenylmethyl chloroformate (2.0 g, 7.73 mmol, 1.1 eq). The reaction was stirred at room temperature for 48 hr. Then the reaction was acidified with 1M HCl (aq) and extracted three times with diethyl ether. The ether was dried with MgSO₄, filtered, and removed under reduced pressure. The crude product was purified by silica column with chloroform and 0-10% methanol. The product was a white solid (0.95 g, 31% yield). LRMS (ESI): 438.2 (M+1).

3.5.2.2 *Loading Fmoc-Fatty Acid to Wang Resin*

The Wang resin (100-200 mesh, 1.1 mmol/g, 1.0 g, 1.09 mmol, 1.0 eq) was placed in a fritted peptide flask with the Fmoc-fatty acid (0.95 g, 2.17 mmol, 2.0 eq) in 10 mL of DMF. The flask was shaken for 15 min. Then pyridine (0.26 mL, 3.27 mmol, 3.3 eq) and 2,6-dichlorobenzoyl chloride (0.31 mL, 2.17 mmol, 2.0 eq) were added, and the reaction was shaken for 20 hr. The flask was drained, and the resin was washed with DMF, DCM, and methanol. The resin was dried slightly under nitrogen, and then 35 mL dichloroethane, 3 mL pyridine, and 3 mL benzoyl chloride were added. The reaction was shaken for 2 hr. Then flask was drained, and the resin was washed with DMF, DCM, and methanol. The resin was washed with dichloroethane, and then 10 mL dichloroethane, 2 mL pyridine, and 2 mL benzoyl chloride were added. The reaction was shaken for 30 min, and the draining and washes were repeated again. The resin was dried thoroughly on the high vacuum line.

The loading of the resin was measured by first carefully weighing out a small portion of dry resin (10.3 mg). To the resin was added 2 mL of a 2% DBU in DMF solution. The solution was shaken for 30 min and then diluted with 10 mL of acetonitrile. A 2-mL portion of this solution was taken and diluted to a total volume of 25 mL. A reference solution was prepared in the same way without the resin. The sample and reference solutions were placed in cuvettes, and the absorbance at 304 nm was read three times and averaged. From the following equation, the loading was calculated to be 0.60 mmol/g.

$$Fmoc \text{ loading } \left(\frac{mmol}{g} \right) = \frac{(Abs_{sample} - Abs_{ref}) \times 16.4}{mg \text{ resin}}$$

3.5.2.3 Peptide Synthesis

The peptides were synthesized on an automated peptide synthesizer (Protein Technologies) using Fmoc chemistry and the Wang resin pre-loaded with Fmoc-12-aminododecanoic acid. For a 200 μ mol-scale reaction, the Wang resin was swelled in DMF and then deprotected with 20% piperidine in DMF twice. In between the resin was washed with DMF and DCM. Each amino acid was dissolved in DMF (100 mM) and introduced to the resin after the deprotection step. PyBOP (300 mM) and DIPEA (1.2 M) in DMF were added along with the amino acid to perform each coupling (twice). The final residue was deprotected at the end. The resin was then removed from the synthesizer.

The resin was washed with DMF, DCM, MeOH, and glacial acetic acid and then dried overnight on the high vacuum line. The peptides were cleaved from the resin with 95% TFA, 2.5% TIPS, and 1% H₂O for 4 hr. The TFA solution was filtered from the resin, and the TFA was removed. To the residue was added diethyl ether to precipitate the peptides.

3.5.2.4 Acetylation of ACKGGRK-FA

The crude ACKGGRK-FA was dissolved in 5 mL of acetic anhydride and 200 μ L of DIPEA. The solution was stirred for 18 hr. Then another 2 mL of acetic anhydride and 100 μ L of DIPEA were added, and the reaction was stirred for another 5 hr. The acetic anhydride was quenched by adding a few milliliters of water, and the solution was concentrated.

3.5.2.5 Peptide Purification

The crude peptides were dissolved in water/DMSO (no more than 50% DMSO) and purified by HPLC. Fractions containing the desired peptide were verified by LC-MS, combined, and evaporated.

KAGLK-FA: 40.6 mg (39% yield); LRMS (ESI): 713.2 (M+1); HRMS (ESI): expected 713.53, found 713.53 (M+1)

ACKGGRK-FA: 16.1 mg (18% yield); LRMS (ESI): 907.8 (M-1), 910.0 (M+1); HRMS (ESI): expected 901.57, found 910.57 (M+1)

3.5.3 Titrations

Details about the titrations can be found in the captions of the figures in this chapter including the concentrations, buffer used, and excitation/emission wavelengths.

The KAGLK-FA stock was prepared in 10 mM phosphate buffer, 140 mM NaCl, H₂O, pH 7.40, 0.02% NaN₃ at a concentration of 770 μ M. The KGGRK-FA stock was prepared in DMSO (100%) at a concentration of 3.54 mM.

3.5.4 Additional Figures

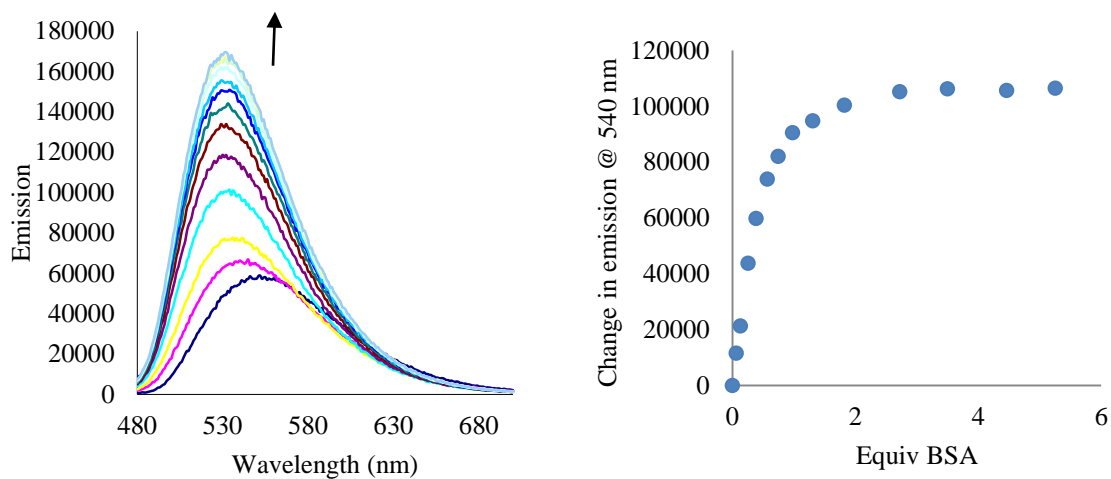


Figure 3.15 *Titration of BSA into NBD-FA.* Addition of BSA (0-79 μM) to NBD-FA (15 μM) in 10 mM phosphate buffer, H₂O, pH 7.00, 0.02% NaN₃, λ_{ex} = 470 nm.

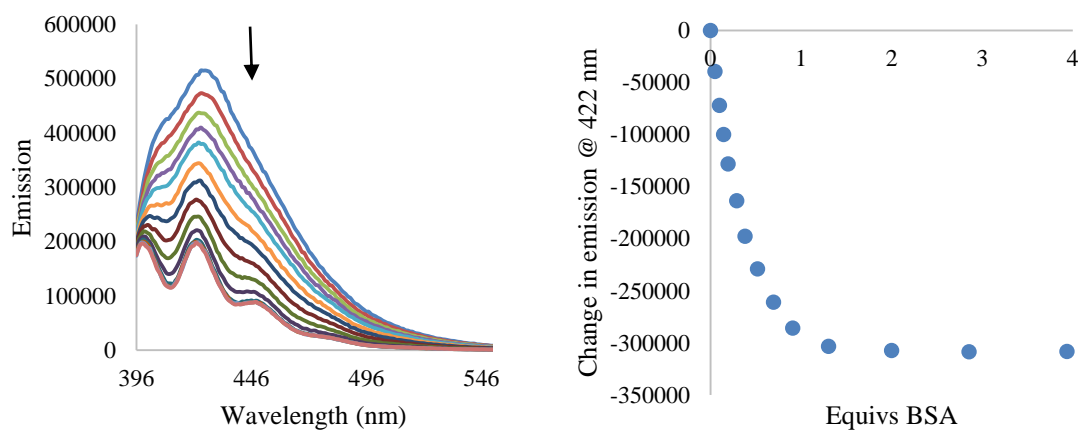


Figure 3.16 *Titration of BSA into AC.* Addition of BSA (0-39 μM) to AC (10 μM) in 10 mM phosphate buffer, H₂O, pH 7.00, 0.02% NaN₃, λ_{ex} = 386 nm.

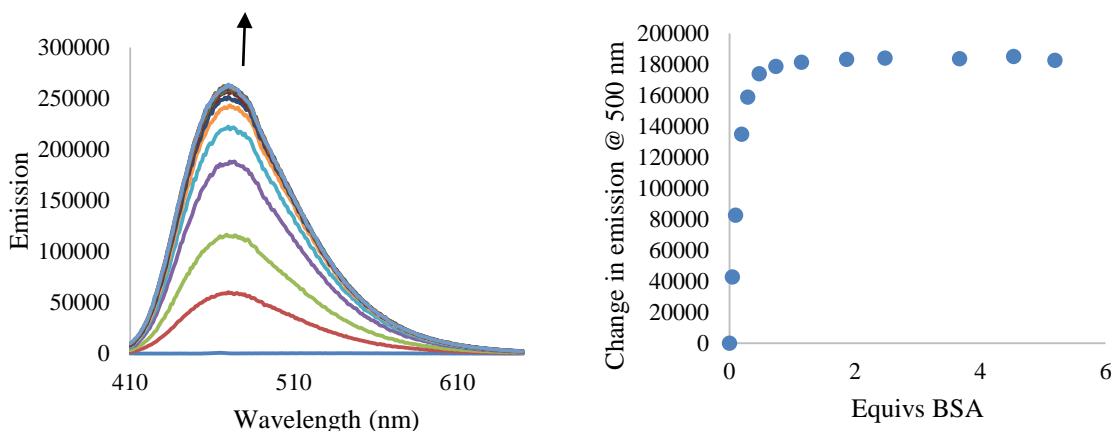


Figure 3.17 *Titration of BSA into ANS.* Addition of BSA (0-52 μM) to ANS (10 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 , $\lambda_{\text{ex}} = 400$ nm.

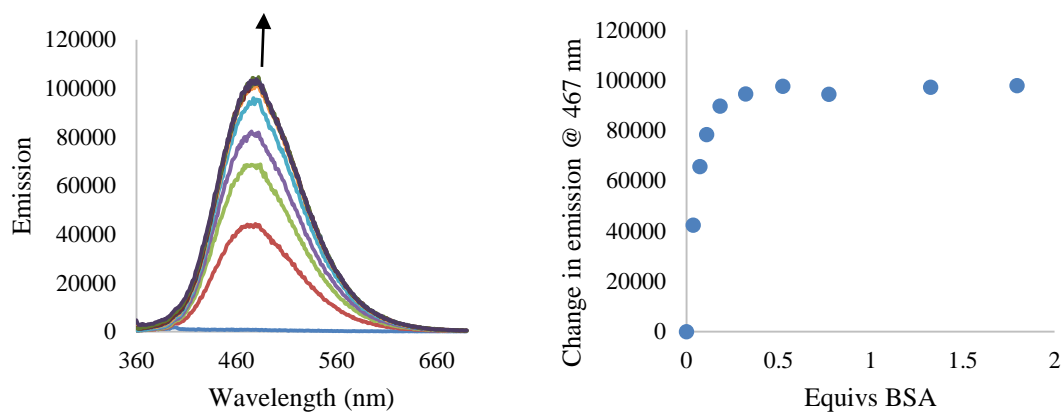


Figure 3.18 *Titration of BSA into BSPOTPE.* Addition of BSA (0- 36 μM) to BSPOTPE (20 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 , $\lambda_{\text{ex}} = 350$ nm.

3.6 REFERENCES

- (1) De Bruin, B.; Hauwert, P.; Reek, J. N. H. *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 2660–2663.
- (2) Lehn, J.-M. *Science.* **2001**, *291*, 2331–2332.
- (3) Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J.-L.; Sanders, J. K. M.; Otto, S. *Chem. Rev.* **2006**, *106*, 3652–3711.

- (4) *Organic Synthesis and Molecular Engineering*; Nielsen, M. B., Ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2013.
- (5) Peters, T. *All About Albumin: Biochemistry, Genetics, and Medical Applications*; Academic Press: San Diego, CA, 1996.
- (6) Simard, J. R.; Zunszain, P. A.; Hamilton, J. A.; Curry, S. *J. Mol. Biol.* **2006**, *361*, 336–351.
- (7) Gade, A. M. *Sensing Approaches for the Discrimination of Small Molecules and Multivalent Analytes*, 2014.
- (8) Marchesini, S.; Preti, A.; Aleo, M. F.; Casella, A.; Dagan, A.; Gatt, S. *Chem. Phys. Lipids* **1990**, *53*, 165–175.
- (9) Sieber, P. *Tetrahedron Lett.* **1987**, *28*, 6147–6150.
- (10) Gude, M.; Ryf, J.; White, P. D. *Lett. Pept. Sci.* **2002**, *9*, 203–206.
- (11) Ashbrook, J. D.; Spector, A. A.; Fletcher, J. E. *J. Biol. Chem.* **1972**, *247*, 7038–7042.
- (12) Sudlow, G.; Birkett, D. J.; Wade, D. N. *Mol. Pharmacol.* **1975**, *11*, 824–832.
- (13) Meisner, H.; Neet, K. *Mol. Pharmacol.* **1978**, *14*, 337–346.
- (14) Kasai, S.; Horie, T.; Mizuma, T.; Awazu, S. *J. Pharm. Sci.* **2006**, *76*, 387–392.
- (15) Weber, G.; Daniel, E. *Biochemistry* **1966**, *5*, 1900–1907.
- (16) Nishijima, M.; Kato, H.; Fukuhara, G.; Yang, C.; Mori, T.; Maruyama, T.; Otagiri, M.; Inoue, Y. *Chem. Commun.* **2013**, *49*, 7433.
- (17) Wada, T.; Nishijima, M.; Fujisawa, T.; Sugahara, N.; Mori, T.; Nakamura, A.; Inoue, Y. *J. Am. Chem. Soc.* **2003**, *125*, 7492–7493.
- (18) Hong, Y.; Feng, C.; Yu, Y.; Liu, J.; Lam, J. W. Y.; Luo, K. Q.; Tang, B. Z. *Anal. Chem.* **2010**, *82*, 7035–7043.
- (19) Takehara, K.; Yuki, K.; Shirasawa, M.; Yamasaki, S.; Yamada, S. *Anal. Sci.* **2009**, *25*, 115–120.

- (20) Zhang, Y.-Z.; Zhou, B.; Liu, Y.-X.; Zhou, C.-X.; Ding, X.-L.; Liu, Y. *J. Fluoresc.* **2007**, *18*, 109–118.
- (21) Sułkowska, A. *J. Mol. Struct.* **2002**, *614*, 227–232.

Chapter 4: Investigation of Serum Albumin Binding and Thiol Addition to Water-Soluble Squaraines

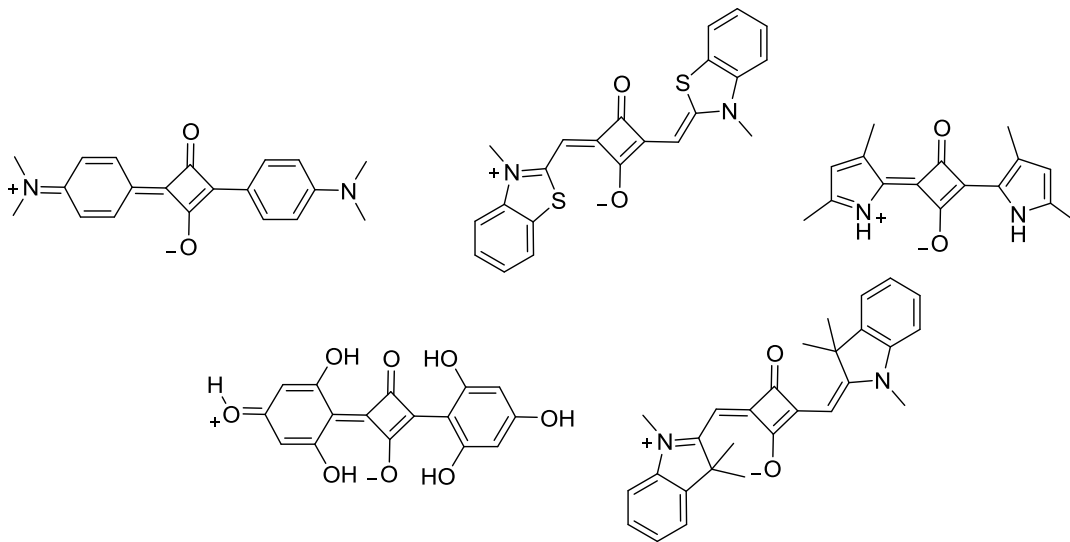
4.1 INTRODUCTION

Since the fatty acid-peptides did not bind to serum albumin as we had postulated, we became interested in using squaraines as part of an alternative strategy. The first squaraine was reported in 1965 by Treibs and Jacob, who synthesized the dye by reacting pyrrole and squaric acid (3,4-dihydroxy-3-cyclobutene-1,2-dione).¹ Despite being known in the literature for fifty years, squaraines have not been researched as extensively as other organic dye classes. A SciFinder search of the term “squaraine” yields only about 1,200 references, while searches for other dye classes like “rhodamine” or “cyanine” yield upwards of 25,000 references. Despite the comparative dearth of research on squaraines, they have found application in imaging,²⁻⁴ photovoltaics,⁵⁻⁷ ion sensing,⁸⁻¹² and other areas.^{13,14}

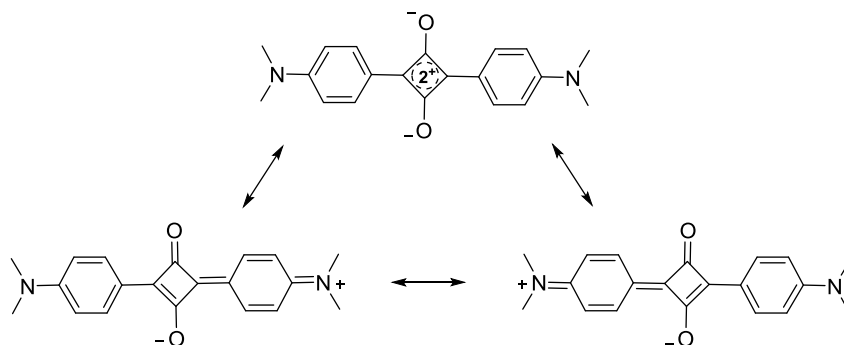
4.1.1 Synthesis, Structure, and Optical Properties of Squaraines

Squaraines are prepared by the reaction of squaric acid with electron rich aromatic or heterocyclic molecules such as *N,N*-dialkylanilines, benzothiazoles, phenols, and pyrroles with azeotropic removal of water (**Scheme 4.1**).² The structure of a squaraine is a resonance-stabilized zwitterion (**Scheme 4.2**). The central four-membered ring is electron deficient, while the oxygen atoms and the two aniline groups are electron donating, leading to a donor-acceptor-donor (D-A-D) structure. From theoretical calculations, Bigelow and Freund showed that during the $S_0 \rightarrow S_1$ transition there is a charge transfer that is largely from the oxygen atoms to the four-membered ring with some minor donation from the aniline groups.¹⁵ This intramolecular charge transfer and the extended conjugation of the molecule lead to absorption in the visible to near-IR region with narrow bands (half bandwidth $\sim 750 \text{ cm}^{-1}$)² and large extinction coefficients ($\epsilon \geq 10^5 \text{ cm}^{-1} \text{ M}^{-1}$).¹⁶

In the solid state, squaraines are microcrystalline powders that are often highly photoconductive.² They exhibit broad absorption between 700-850 nm and fluorescence quenching due to aggregation. For the aniline-derived squaraines, the absorption band is usually between 500-700 nm in solution.



Scheme 4.1 *Examples of squaraines derived from different electron-rich aromatic and heterocyclic molecules. From left to right, top to bottom: aniline-derived, benzothiazole-derived, pyrrole-derived, phloroglucinol-derived, and indolenine-derived squaraines.*



Scheme 4.2 *Resonance of the squaraine zwitterion.*

4.1.2 Aggregation Behavior of Squaraines in Solution

Squaraines can exist in solution as monomers, dimers, or higher order aggregates. The distribution of these species in solution is dependent on the structure of the squaraine as well as the solvent and pH.^{17,18} Squaraines can form H-type aggregates, which exhibit a hypsochromic shift relative to the monomer and are less or non-fluorescent than the corresponding monomer, or J-type aggregates, which exhibit a bathochromic shift and have a higher quantum yield of fluorescence than the monomer.^{19,20} H-type aggregates are stacked vertically like a deck of cards, while J-type aggregates are in a “slipped stack” arrangement (**Figure 4.1**). Squaraines in solution most commonly form H-type aggregates.¹⁶

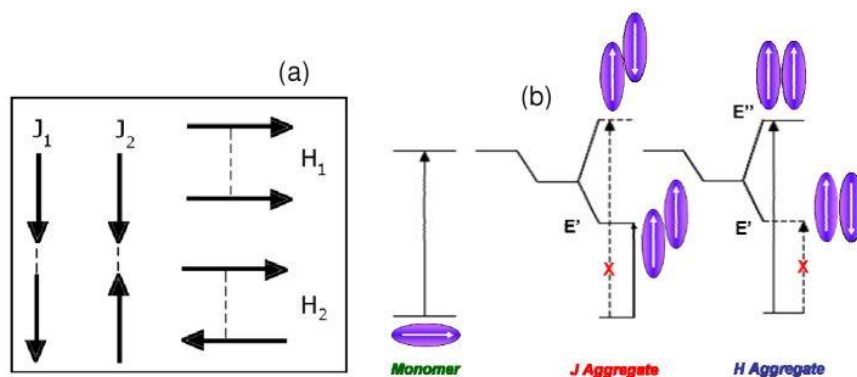
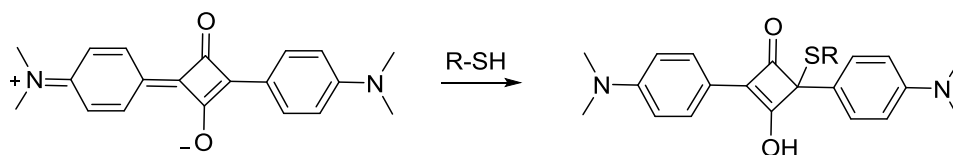


Figure 4.1 Simple schematic of H- and J-type aggregates. a) J and H dimer orientations b) Energy level splitting of J and H dimers.²¹ (Reproduced from Ref. 21: © World Scientific Publishing Co. Pte. Ltd., 2012)

4.1.3 Reaction of Squaraines with Nucleophiles

Martinez-Manez and coworkers first demonstrated in 2002 that the electron deficient four-membered ring in a squaraine is susceptible to nucleophilic attack.²² They used a squaraine as a cyanide probe in acetonitrile/water solutions buffered at pH 9.5. The addition of a nucleophile to a squaraine (**Scheme 4.3**) breaks up the extended conjugation

of the molecule, which switches off the distinctive absorbance and fluorescence of the molecule. This property allows for simple visual or spectroscopic detection. The sensor was selective for cyanide over other insufficiently nucleophilic anions, including halides, nitrate, acetate, thiocyanate, dihydrogen phosphate, and bisulfate. The researchers also used their squaraine as a probe for thiols.²³ The squaraine was exposed to ten equivalents of amino acid in acetonitrile/water solutions buffered at pH 6. Only the solution containing cysteine decolored (**Figure 4.2**), indicating that at pH 6 the probe is selective for thiols over amines and alcohols.



Scheme 4.3 *Thiol addition to a squaraine.*

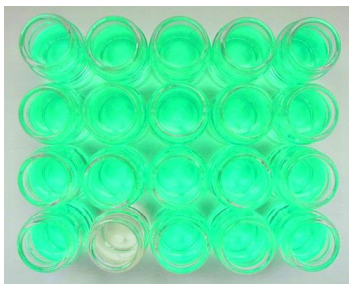


Figure 4.2 *Solutions of squaraine and ten equivalents of amino acids.* Addition performed in acetonitrile/water, 20:80 v/v, at pH 6. From left to right and top to bottom: no amino acid, phenylalanine, threonine, arginine, histidine, asparagine, leucine, alanine, proline, valine, glycine, lysine, glutamine, methionine, isoleucine, serine, cysteine, tryptophan, glutamic acid, and aspartic acid. (Reproduced from Ref. 23: © American Chemical Society, 2004).

Since this report, Martinez-Manez and other researchers have published further studies on squaraines as thiol probes including the detection of volatile thiols,²⁴ detection of thiols at physiological pH,²⁵ and a dual chromophore squaraine for thiol detection.²⁶

The addition of thiols to squaraines has also been used as a method of indirectly measuring other analytes of interest. For example, our group used an organic-soluble squaraine and thiols to discriminate metal ions based on their thiophilicity.²⁷ In the presence of thiol, the squaraine decolors due to nucleophilic addition of the thiol. Upon addition of mercury(II), palladium(II), or another thiophilic metal ion to the solution, the blue color returns as the metal binds to the thiol and pushes the equilibrium toward intact squaraine. This reaction with metal ions demonstrates the reversible nature of the thiol-squaraine reaction. Many researchers have used squaraines as probes for mercury(II) based on these general reactivity properties.²⁸⁻³⁰ Likewise, a squaraine has been used to detect the inhibition of acetylcholinesterase, which converts the thioester in acetylcholine into a free thiol.³¹ If the enzyme is turning over, thiol is generated in the solution, and the squaraine decolors. However, if the enzyme is inhibited, no thiol is generated, and the squaraine remains intact.

While the addition of thiols to squaraines is a well-known reaction, not all squaraines are susceptible to nucleophilic attack. Typically, the squaraines employed for the addition are of the *N,N*-dialkylaniline type. Squaraines derived from indoles, benzothiazoles, and other heterocyclic compounds have not been shown to react with nucleophiles and are stable in solution. Phenol-derived squaraines can be electrophilic if the phenols are substituted with electron withdrawing groups such as halogens.³²

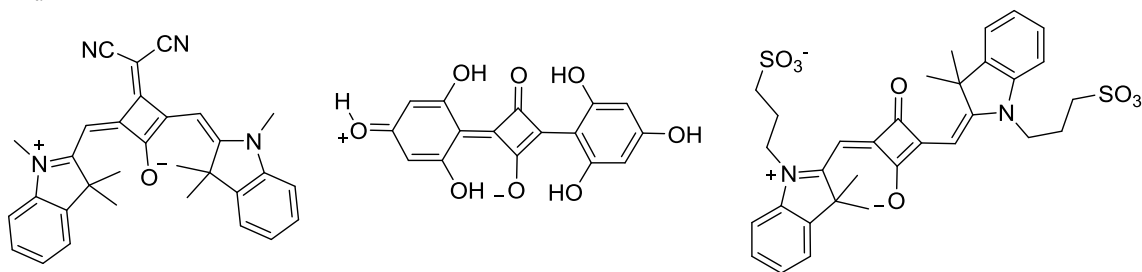
For the *N,N*-dialkylaniline-derived squaraines, electron-donating hydroxyl groups placed *ortho* to the four-membered ring of the squaraine render it much less electrophilic. Squaraines with these hydroxyl groups exhibit significantly greater stability in aqueous

solution compared to squaraines lacking hydroxyl groups.¹⁹ Furthermore, these hydroxyl-functionalized squaraines are essentially not reactive with thiols.³³

4.1.4 Binding of Squaraines to Serum Albumin

The binding of squaraines to serum albumin has been investigated primarily for protein detection and non-covalent labeling,^{34–38} but also for *in vitro* cell imaging.³⁹ In these examples, typically the indole- and benzothiazole-derived squaraines have been investigated, although one research group has used phenol-derived squaraines.^{37,38} Aniline-derived squaraines have not been studied in terms of their binding to serum albumin.

In the studies of indole-, benzothiazole-, and phenol-based squaraines (**Scheme 4.1**), researchers have invariably found that, when these squaraines bind to serum albumin, large increases in emission are observed. These increases are generally on the order of 10- to 200-fold enhancement in intensity and are accompanied by bathochromic shifts in the λ_{max} of emission.



Scheme 4.4 Structures of the squaraines studied in their binding to serum albumin.

The effect of serum albumin binding on the absorption spectra of squaraines is more variable. In the case of the 3-dicyanomethylene-substituted indolenine dye synthesized by Yarmoluk and coworkers (**Scheme 4.4**), the absorption of the squaraine increases about three-fold in the presence of BSA.³⁵ However, Ramaiah and coworkers observed more

complex absorbance behavior with their phloroglucinol-derived squaraine (**Figure 4.3, Scheme 4.4**).³⁷ As BSA is added, the absorbance decreases up to 0.5 equivalents of BSA and then increases up to about three equivalents of BSA. The authors attribute this phenomenon to site selectivity, which they further tested with lifetime and indicator displacement experiments with dansyl amide (DNSA) and dansyl proline (DP) as Sudlow Site I and II-specific ligands, respectively. They postulate that at low concentrations of serum albumin, the squaraine binds in Sudlow Site I, forming a tight complex through hydrophobic and π -stacking interactions. At more than 0.5 equivalents of BSA, the squaraine binds in Sudlow Site II, a more spacious binding site and one lacking in a tryptophan residue for π -stacking. Thus, the binding in this site is looser than in Site I.

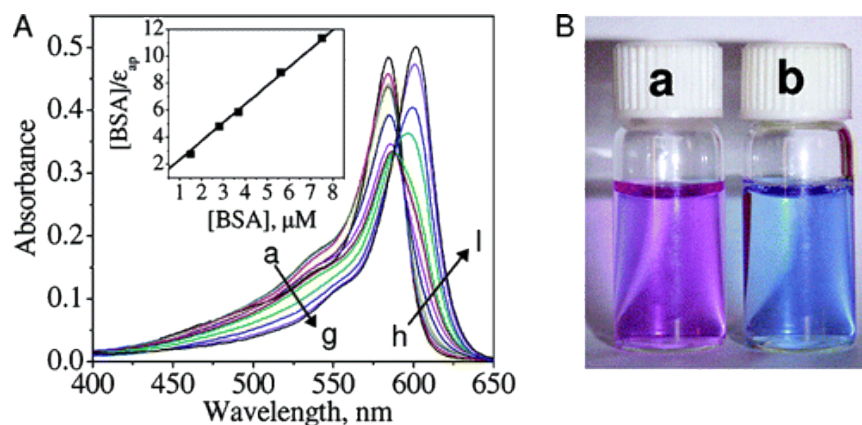


Figure 4.3 Addition of BSA to phloroglucinol-derived squaraine of Ramaiah *et al.* A) Change in absorbance of the squaraine (3 μ M) with increasing BSA in phosphate buffer a) 0 μ M BSA, g) 1.5 μ M BSA, h) 1.8 μ M BSA, and l) 7 μ M BSA. B) Visible detection of BSA a) squaraine alone b) squaraine + BSA. (Reproduced from Ref. 37: © American Chemical Society 2006)

Similarly, Belfield and coworkers observed an initial decrease in absorption of their benzindolenine-based squaraine upon addition of BSA (**Figure 4.4, Scheme 4.4**).³⁹ Upon addition of 0.4 equivalents of BSA, the absorption decreased and exhibited a bathochromic

shift. Another addition (0.8 equivalents total) of BSA caused the absorption to be further red-shifted and to increase. Further additions did not exhibit shifts but did increase the absorption up to about two equivalents of BSA. The authors did not provide a physical explanation for this behavior. They did observe that the peak at 590 nm, which corresponds to the dimer of the squaraine, diminished as serum albumin was added, implying that only the monomer binds and that the BSA breaks up the dimers.

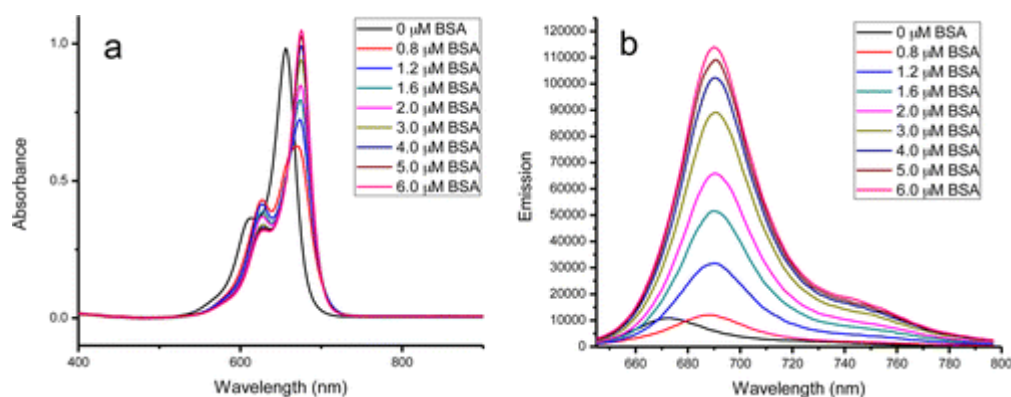


Figure 4.4 *Addition of BSA to benzoindolenine-derived squaraine of Belfield et al.* Titration of BSA (0-12 μM) to squaraine (5 μM) a) absorbance spectra and b) fluorescence spectra. (Reproduced from Ref. 39: © American Chemical Society, 2013)

4.1.5 Project Goals and Design

We hypothesized that squaraines could be used in conjunction with thiol-functionalized recognition units and serum albumin to build dynamic, multivalent receptors. The squaraines bind to the protein non-covalently, and thiol-containing units such as peptides reversibly add to the squaraines (**Figure 4.5**). The receptor would self-assemble and could be templated against large, multivalent analytes such as proteins or cell surfaces. Thus, we wanted to identify a squaraine that 1) binds to serum albumin with high stoichiometry ($n > 1$) and good affinity ($K_a \geq 10^5 \text{ M}^{-1}$) and 2) reversibly adds thiols.

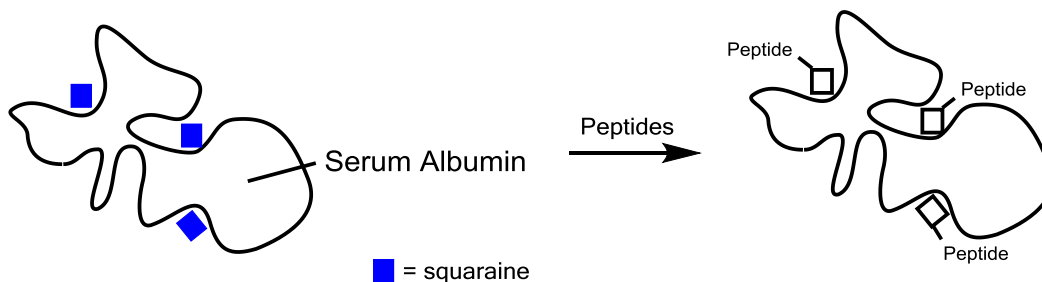


Figure 4.5 *Cartoon representation of the squaraines and peptides assembled on the serum albumin scaffold.*

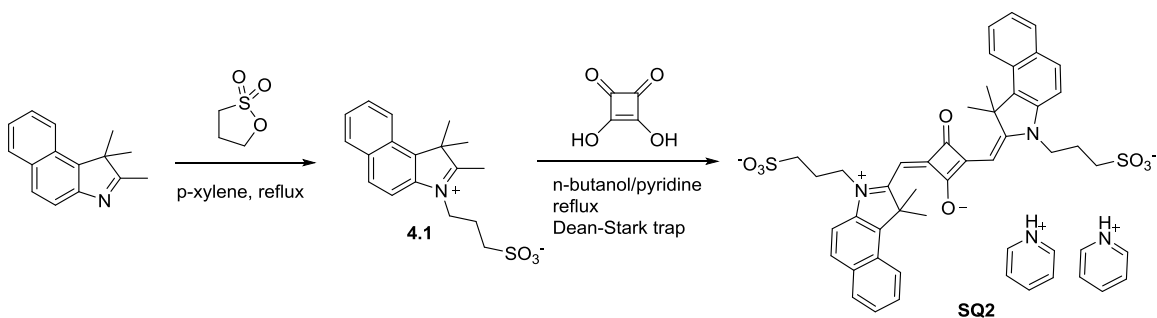
While our ultimate goal was to build receptors, we realized quickly upon beginning the project that we needed to first investigate squaraine structure and function in more depth than had previously been done in the context of thiol addition and serum albumin binding. Therefore, we synthesized a series of squaraines and used spectroscopic methods to study their properties.

4.2 SYNTHESIS OF SQUARAINES

The squaraines synthesized were all water soluble or able to be dissolved in water with a small amount (< 10%) of acetonitrile, methanol, or dimethyl sulfoxide (DMSO). Yields varied significantly based on the squaraine.

4.2.1 Indolenine-Based (SQ2)

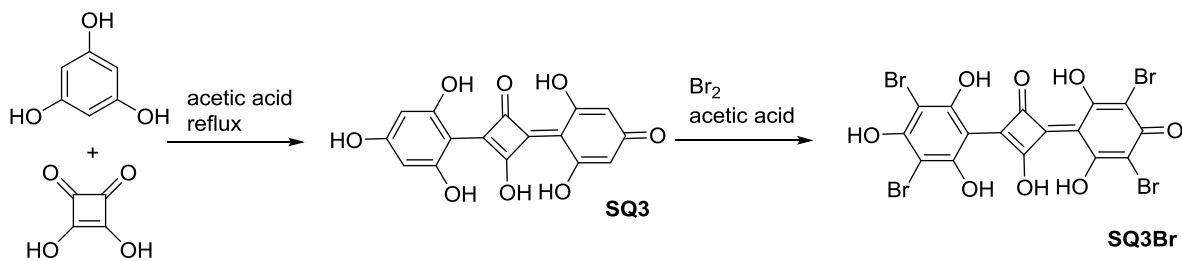
The indolenine-derived squaraine (SQ2) was synthesized by first attaching a sulfonate group to the benzoindolenine precursor with 1,3-propanesultone. The resultant compound was refluxed with squaric acid in *n*-butanol with azeotropic removal of water by a Dean-Stark trap according to a literature procedure (Scheme 4.5).³⁹ The product was purified by recrystallization from methanol (12% yield).



Scheme 4.5 *Synthesis of SQ2.*

4.2.2 Phloroglucinol-Based (SQ3 and SQ3Br)

These squaraines were also synthesized according to a literature procedure (**Scheme 4.6**).^{37,38} **SQ3** was synthesized by reacting phloroglucinol and squaric acid in acetic acid and purified by recrystallization from acetic acid. **SQ3** was brominated in acetic acid to yield **SQ3Br**. The product precipitated from the reaction and required no further purification. These squaraines were obtained in moderate yield (**SQ3**- 41%, **SQ3Br**- 35%).



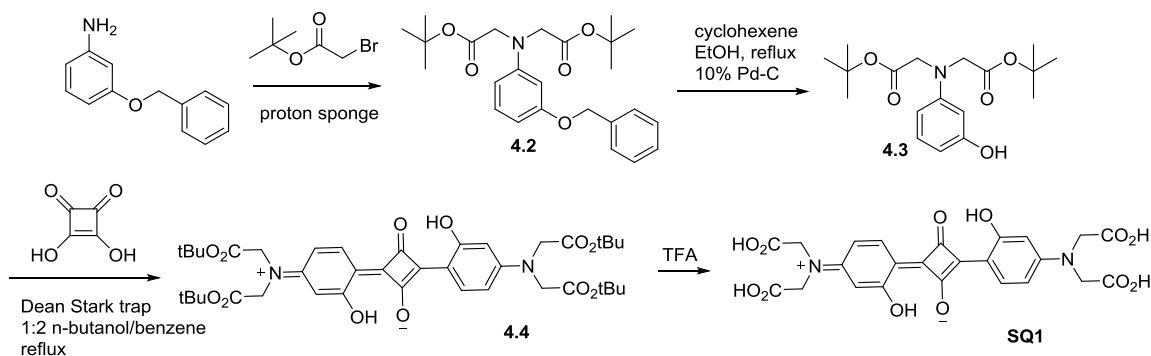
Scheme 4.6 *Synthesis of SQ3 and SQ3Br.*

4.2.3 Aniline-Based (SQ1, SQ4-9)

Most of the squaraines that were made for this work were aniline-derived. These types of squaraines are synthesized by heating squaric acid and the aniline in *n*-butanol and benzene to 100°C with azeotropic removal of water. In the literature, toluene is sometimes

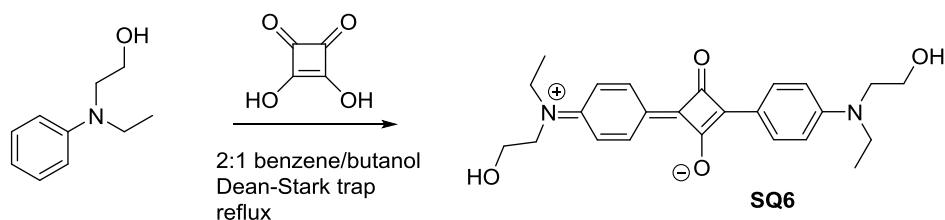
used in place of benzene; however, we had no success using toluene. The optimum ratio of butanol to benzene varies by aniline. Yields for aniline-based squaraines vary widely from less than 10% to more than 90% depending on the aniline used.² Since we wanted water soluble dyes, aniline precursors with carboxylates (**SQ1**), alcohols (**SQ6**), and sulfonates (**SQ4, 5, 7-9**) were used.

SQ1 was synthesized according to a literature procedure (**Scheme 4.7**).⁴⁰ First, 3-benzyloxyaniline was alkylated with *t*-butylbromoacetate. Next, the benzyl protecting group was removed by transhydrogenation. The squaraine was formed by refluxing in butanol/benzene (1:2) and was purified by silica chromatography (65% yield). The *t*-butyl protecting groups were removed with TFA, and this final product required no further purification (92% yield).



Scheme 4.7 *Synthesis of SQ1.*

SQ6 was synthesized according to a literature procedure (**Scheme 4.8**).²⁵ The pure product precipitated from the reaction, requiring no further purification (22% yield).

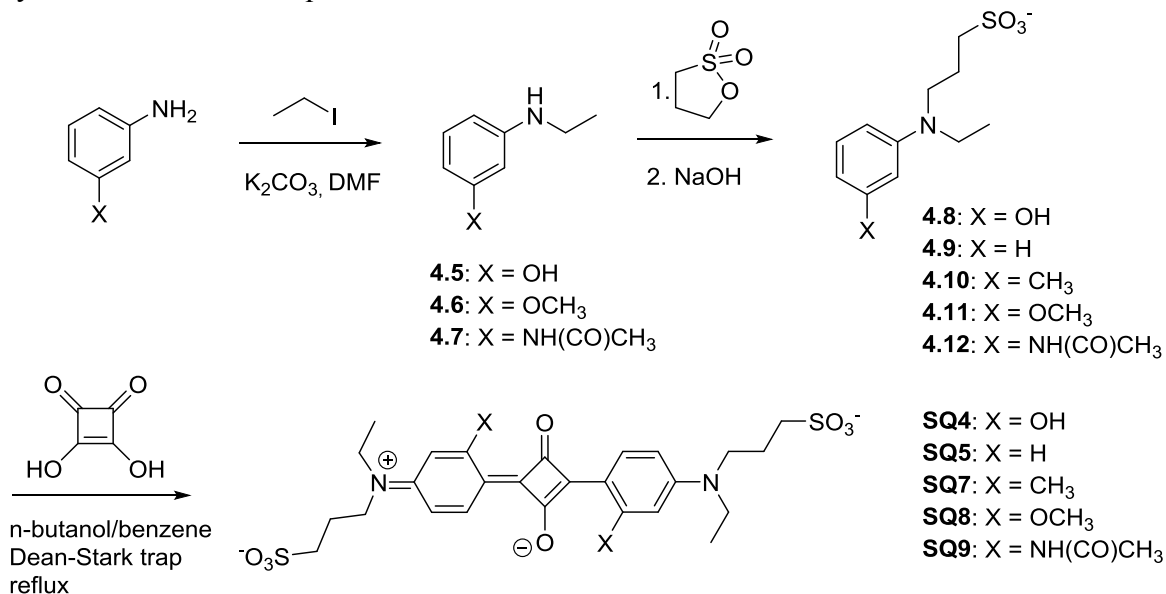


Scheme 4.8 *Synthesis of SQ6.*

The rest of the squaraines contained sulfonate groups for water solubility (**SQ4**, **5**, **7-9**). We wanted to investigate the effect of aniline substitution, so we synthesized a series of squaraines with different substituents in the benzene ring at the *ortho* position to the four-membered ring (**Scheme 4.9**). To do so, the appropriate aniline derivative was monoalkylated with iodoethane, and then sulfonated with 1,3-propanesultone. Then, the squaraine was formed with azeotropic removal of water. This procedure was based on the protocol described by Ghazarossian and coworkers to synthesize **SQ4**.⁴¹ We followed their procedure for the synthesis of **SQ4** exactly but modified it for the synthesis of the other squaraines. In their protocol, the sulfonated anilinium was used directly in the squaraine formation reaction, and they used sodium bicarbonate *in situ* to obtain the free base. The researchers also used DMSO in the squaraine formation reaction to improve solubility of the charged aniline species in the solvent.

While this protocol yielded a small amount of **SQ4**, we could not make **SQ5** in this way. The sodium bicarbonate interferes with the squaraine formation by deprotonating squaric acid (pK_a 1.5, 3.4),⁴² rendering it insufficiently electrophilic to be attacked by aniline. Thus, we obtained the free base of the sulfonated aniliniums by carefully adjusting an aqueous solution of the compound to pH 7 and then removing the water to isolate the sodium salt. The salt was used to form the squaraine. Additionally, we found that the

formation of **SQ5** would not occur in the presence of DMSO, so it was omitted from the synthesis of the other squaraines as well.



Scheme 4.9 Synthesis of **SQ4**, **5**, **7**, **8**, and **9**.

The yields were generally very low (< 8%), although **SQ9** was synthesized in a surprisingly good yield (62%). There were several reasons for the low yields. First, the squaraine formation reaction was performed at a high temperature (100°C) for extended periods (18-48 hours) which resulted in a large number of products, the distribution of which is determined by the stability of the various species.⁴³ Less stable squaraines (X = H, CH₃) in particular form in smaller amounts. These reaction conditions are necessary to overcome the poor nucleophilicity of aniline, but they result in many unwanted byproducts. Furthermore, the solubility of the sulfonated anilines is poor in butanol/benzene. While Ghazarossian *et al.* successfully used ethylene glycol and DMSO to ameliorate this solubility problem, we found that the presence of these solvents hindered the formation of

squaraine rather than improving it. We found that using a larger proportion of butanol to benzene (closer to 1:1 rather than 1:2) did seem to improve yield, although we did not investigate this factor systematically. Finally, the squaraines, with the exception of **SQ9**, had to be purified by reverse-phase chromatography, which resulted in poor recovery of pure product. **SQ9** precipitated from the reaction mixture and required no further purification. **SQ4** was purified by RP-HPLC in water/acetonitrile, 0.1% TFA. **SQ5**, **7**, and **8** were purified on a C-18 column in water/acetonitrile with a CombiFlash. They were not stable under acidic conditions, so TFA was not used. Even after method optimization, resolution on the column was poor, resulting in loss of product. Despite the poor yields, enough material was obtained for us to carry out the necessary experiments.

4.3 OPTICAL PROPERTIES OF THE ANILINE-BASED SQUARAINES

The absorbance spectra for **SQ4**, **5**, **7**, **8**, and **9** are shown in **Figure 4.6**, and their optical properties are summarized in **Table 4.1**. These squaraines exhibit typical squaraine absorbance properties: absorption bands around 500- 700 nm and extinction coefficients around $10^5 \text{ M}^{-1}\text{cm}^{-1}$. The squaraines are fluorescent in aqueous solution, but the intensity is relatively low.

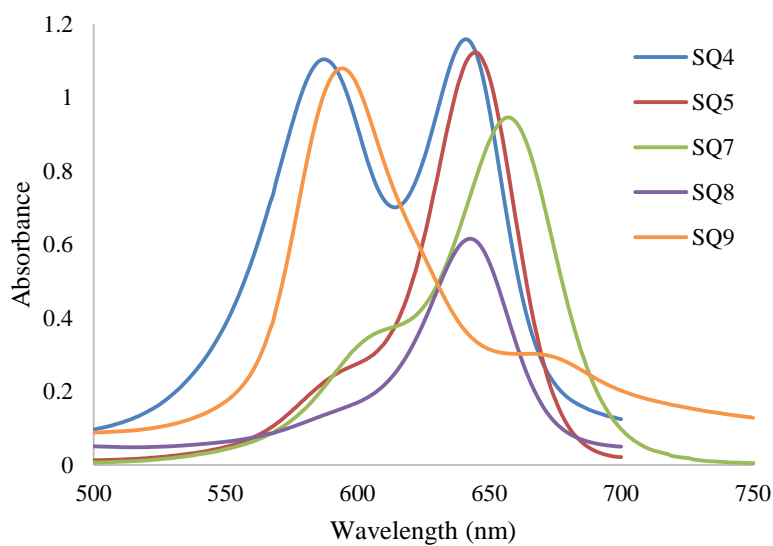


Figure 4.6 Absorbance spectra of **SQ4**, **5**, **7**, **8**, and **9**. The squaraines (10 μM) were dissolved in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 . 4 = OH, 5 = H, 7 = CH_3 , 8 = OCH_3 , and 9 = $\text{NH}(\text{CO})\text{CH}_3$.

Table 4.1 Summary of optical properties of aniline-based squaraines.

Squaraine	λ_{max} (nm)	$\lambda_{\text{emission}}$ (nm)	ϵ ($\text{M}^{-1}\text{cm}^{-1}$)
SQ4	585, 640	670	1.10×10^5 (585 nm) 1.16×10^5 (640 nm)
SQ5	595, 645	680	2.60×10^4 (595 nm) 1.12×10^5 (645 nm)
SQ7	605, 660	690	3.51×10^4 (605 nm) 9.35×10^4 (660 nm)
SQ8	595, 645	675	1.56×10^4 (595 nm) 6.10×10^4 (650 nm)
SQ9	595, 675	695	1.08×10^5 (595 nm) 2.93×10^4 (675 nm)

From the absorbance spectra, we can see that these squaraines exist as monomer and dimer in solution. The dimer peak is around 595 nm, and the monomer peak is around 650 nm. Only the monomer is emissive. These dimers are H-type aggregates, which are common for squaraines.¹⁶ **SQ5**, **7**, and **8** appear to exist primarily as the monomer in aqueous solution. **SQ4** seems to be a mixture of the two forms, while **SQ9** is mostly dimer.

4.4 STUDIES OF THIOL ADDITION TO SQUARAINES

In order to test the addition of thiols to the squaraines, *N*-acetylcysteine (NAC) was used for the thiol. Most of the studies undertaken entailed absorbance and fluorescence measurements, because these methods require considerably less material than ¹H-NMR. The decoloration of squaraines upon thiol addition can easily be followed visually or with UV-vis spectroscopy (**Figure 4.2**). The characteristic emission of squaraines also diminishes upon thiol addition. In the ¹H-NMR spectrum, desymmetrization of the dyes can be observed as well.

4.4.1 SQ2

While there was no specific evidence in the literature that indicated thiols will add to indolenine-derived squaraines, we wanted to confirm for ourselves whether or not this reaction would occur. These types of squaraines are much easier to make than the aniline-based dyes, so it would be desirable to use them if possible. To test this reaction, an excess of NAC was added to a solution of squaraine, and the absorbance was monitored over time (**Figure 4.7**). Even in the presence of 50 equivalents of thiol, the squaraine did not decolor.

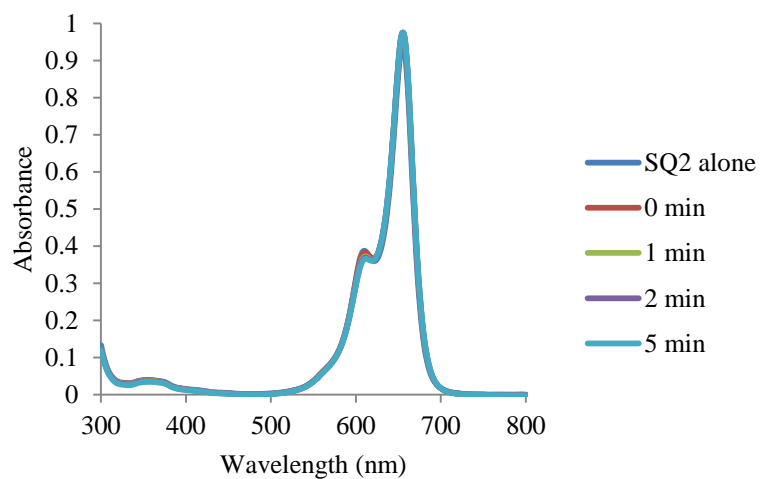


Figure 4.7 *Addition of thiol to SQ2.* Absorbance spectra of SQ2 (5 μ M) before and after NAC addition (250 μ M) in 10 mM phosphate buffer, H₂O, pH 7.00, 0.02% NaN₃.

Next, we tested whether a selenol could add to SQ2, since selenols are more nucleophilic than the corresponding thiols. The only selenol we had on hand was benzeneselenol, which is not very water soluble, so precipitation and cloudiness were problematic for our measurements (**Figure 4.8**). The absorbance did initially decrease (1 min) before the solution became too cloudy to obtain reliable measurements (2+ min). Due to the toxicity of selenols and the difficulties associated with working with them (e.g. oxidation), we decided not to pursue selenol addition further.

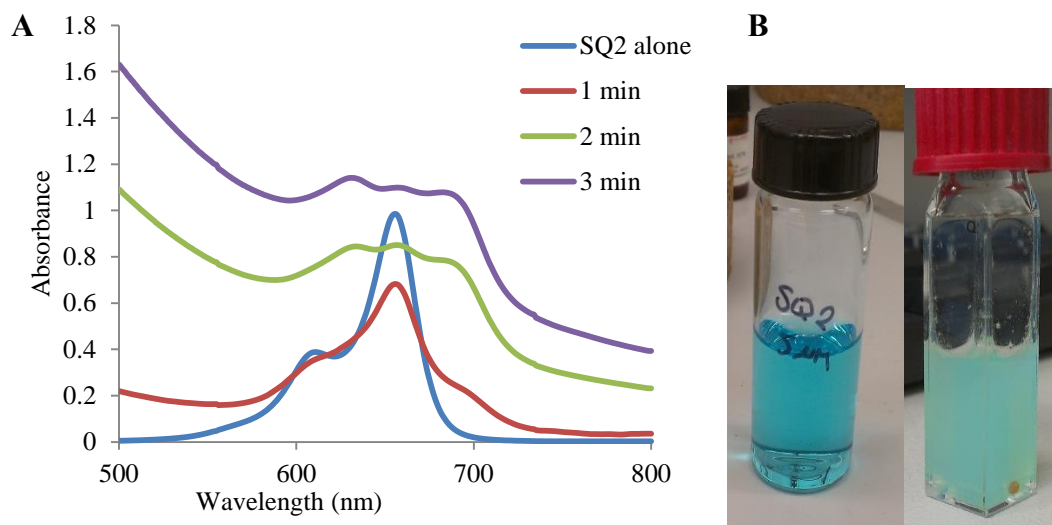


Figure 4.8 *Addition of selenol to SQ2.* A) Absorbance spectra of SQ2 before (5 μ M) and after addition of benzeneselenol (5 μ L, 24 mM) in 10 mM phosphate buffer, H₂O, pH 7.00, 0.02% NaN₃. B) Left: SQ2 alone; Right: SQ2 after addition of selenol.

4.4.2 SQ3Br

SQ3Br has been reported in the literature to react with thiols.³² We tested this reaction for ourselves by adding NAC to a solution of SQ3Br and monitoring the absorbance over time (Figure 4.9). The solution did decolor, but it was not buffered, so we were aware that these changes could be attributed to changes in pH. We dissolved the squaraine in a buffered solution (pH 7) and titrated NAC into it while monitoring the emission of the squaraine (Figure 4.10). In this experiment, the emission initially increased, and then shifted to a lower wavelength and decreased. Repeating the experiment gave a different result in which the emission only decreased. At this point, we became concerned about the stability of the squaraine in ambient conditions, since we were getting inconsistent results from repetitions of the same experiment. When working with SQ3Br, the solutions would change color from blue to purple or decolor while simply sitting on the bench. To test the photostability of SQ3Br, a solution of it was placed in the fluorimeter

and periodically excited (**Figure 4.11**). The emission decreased with each excitation, implying that the squaraine was either reacting (e.g. hydrolysis) over time or being photobleached by the excitation light. In either case, while **SQ3Br** may react with thiols, this squaraine is too unstable to be sure that this is the only reaction occurring in solution.

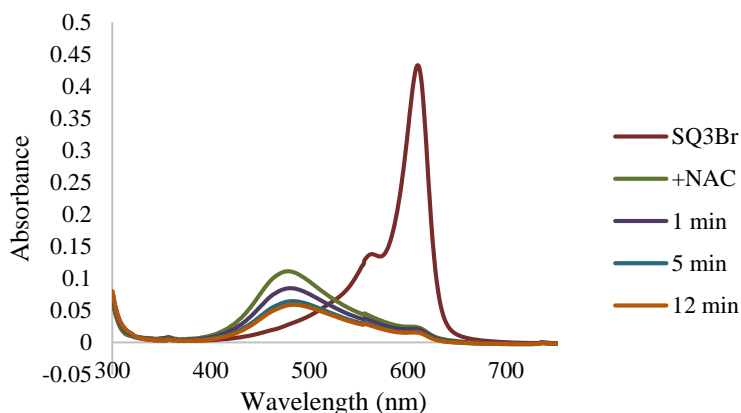


Figure 4.9 Absorbance monitoring of thiol addition to **SQ3Br**. Addition of NAC (3 mM) to **SQ3Br** (4 μ M) in water, <1% ethanol.

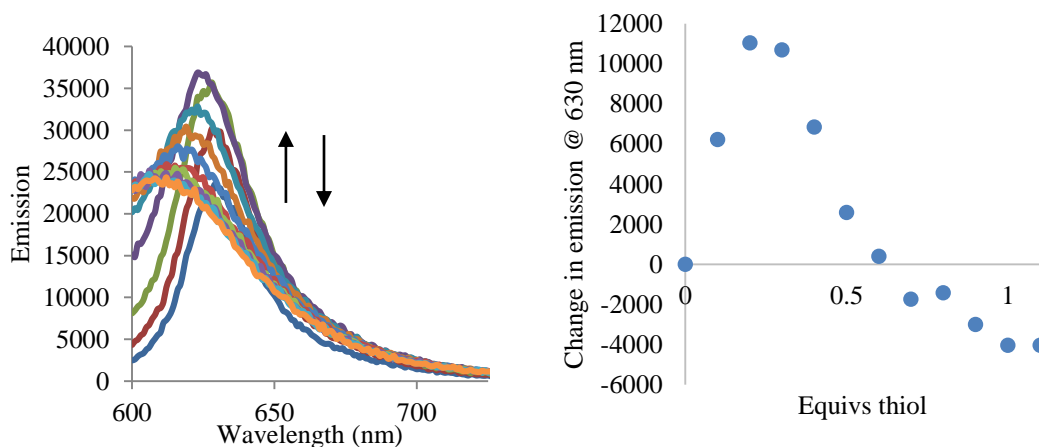


Figure 4.10 Addition of thiol to **SQ3Br** with fluorescence monitoring. Thiol (0-22 μ M) was added to **SQ3Br** (20 μ M) in 10 mM phosphate buffer, H₂O, pH 7.0, 0.02% NaN₃, 2% ethanol, λ_{ex} = 575 nm, λ_{em} = 630 nm.

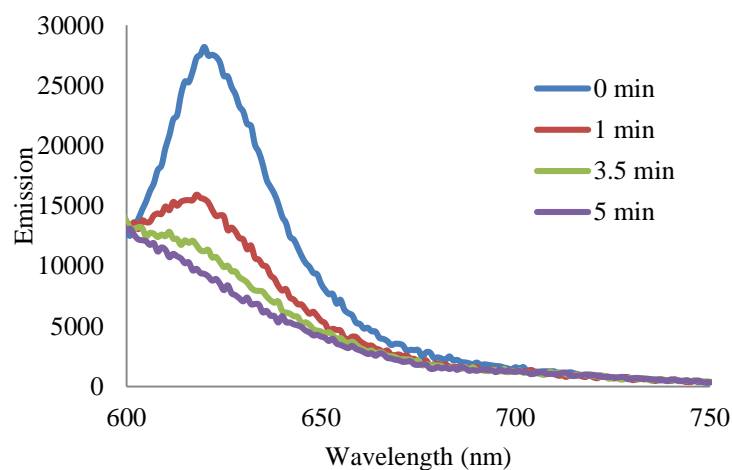


Figure 4.11 Fluorescence monitoring of **SQ3Br** alone over time. **SQ3Br** (2 μM) in 10 mM phosphate buffer, H_2O , pH 7.0, 0.02% NaN_3 , < 1% ethanol, $\lambda_{\text{ex}} = 575$ nm.

4.4.3 SQ6

SQ6 was reported by Liu and coworkers as a thiol probe,²⁵ so we expected to find that it would react with thiols. NAC (5 equiv.) was added to SQ6, and the solution decolorized almost immediately (**Figure 4.12**). Indeed, SQ6 reacts readily with thiols.

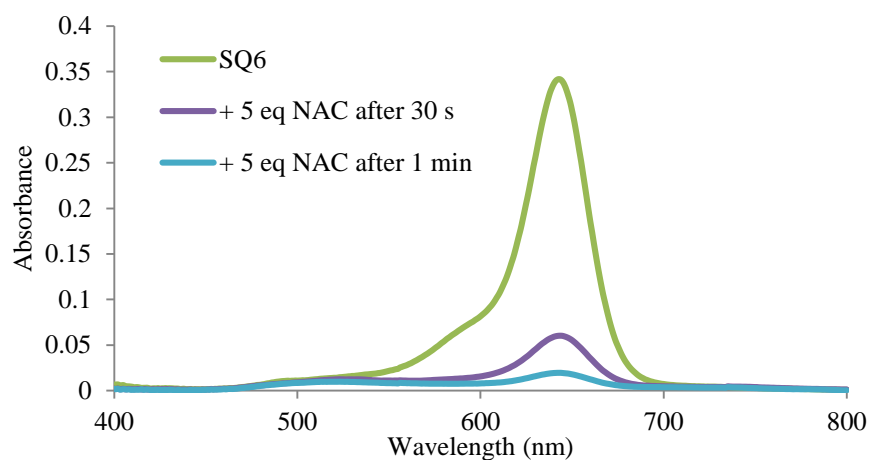


Figure 4.12 Absorbance monitoring of thiol addition to **SQ6**. Addition of NAC (12.5 μM) to **SQ6** (2.5 μM) in 50 mM phosphate buffer, 10 mM EDTA, H_2O , pH 7.5, < 1% DMSO.

4.4.4 Thiol Addition to the Aniline-Derived Squaraine Series (SQ4, 5, 7, 8, and 9)

One of the main things that we wanted to investigate with these aniline-based squaraines was if we could tune their reactivity toward thiols through the substitution at the *ortho* position on the benzene ring (X in **Scheme 4.8**). We knew from the literature that squaraines with no *ortho* substitution (X = H, **SQ5**) will react readily with thiols. However, these dyes also hydrolyze readily, which is a problem for working with squaraines in aqueous solutions. On the other hand, squaraines with a hydroxyl group in the *ortho* position (X = OH, **SQ4**) are essentially not reactive to thiols, presumably because the hydroxyl group donates electron density to the squaraine core, making it much less electrophilic. We hypothesized that placing other donating groups at the *ortho* position would result in intermediate reactivity between a hydrogen and a hydroxyl group. We synthesized squaraines with X = CH₃, OCH₃, and NH(CO)CH₃ (**SQ7**, **8**, and **9**) in order to test this hypothesis. NAC was titrated into each squaraine while monitoring the absorbance (**Figure 4.13-4.18**).

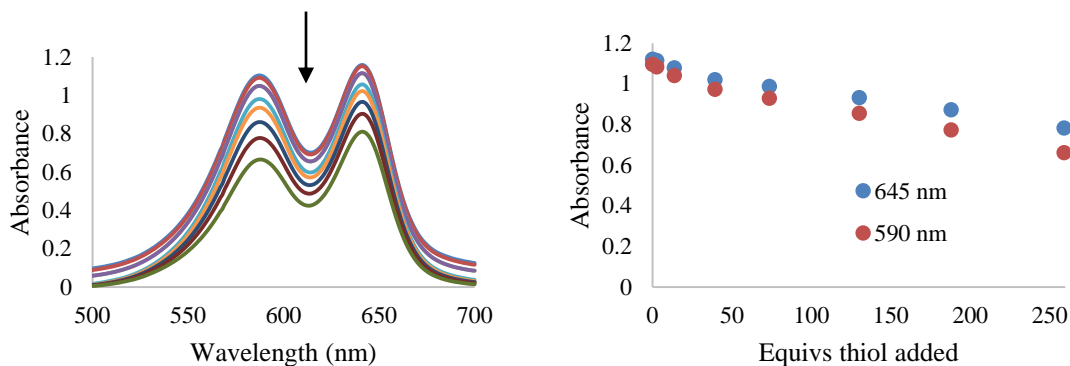


Figure 4.13 Absorbance titration of thiol into **SQ4**. NAC (0-2.59 mM) was added to **SQ4** (10 μ M) in 10 mM phosphate buffer, H₂O, pH 7.0, 0.02% NaN₃, 25°C, 5 min. equilibration between additions.

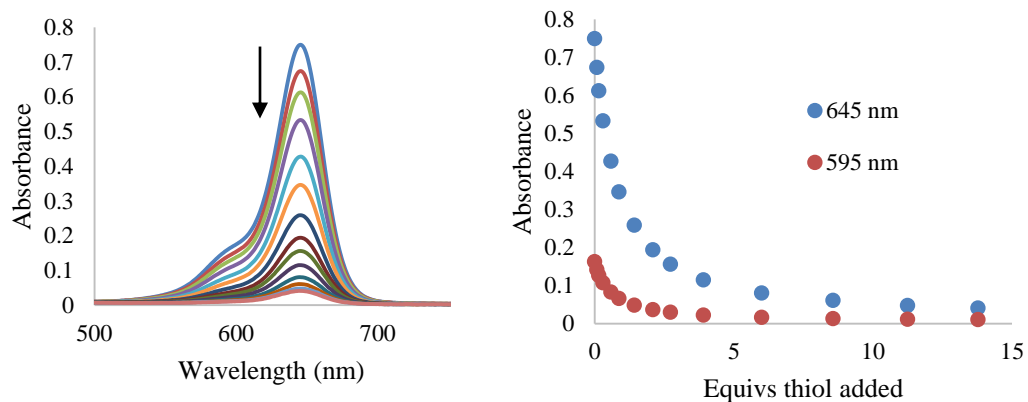


Figure 4.14 Absorbance titration of thiol into **SQ5**. NAC (0-138 μM) was added to **SQ5** (10 μM) in 10 mM phosphate buffer, H_2O , pH 7.0, 0.02% NaN_3 , 25°C, 3 min. equilibration between additions.

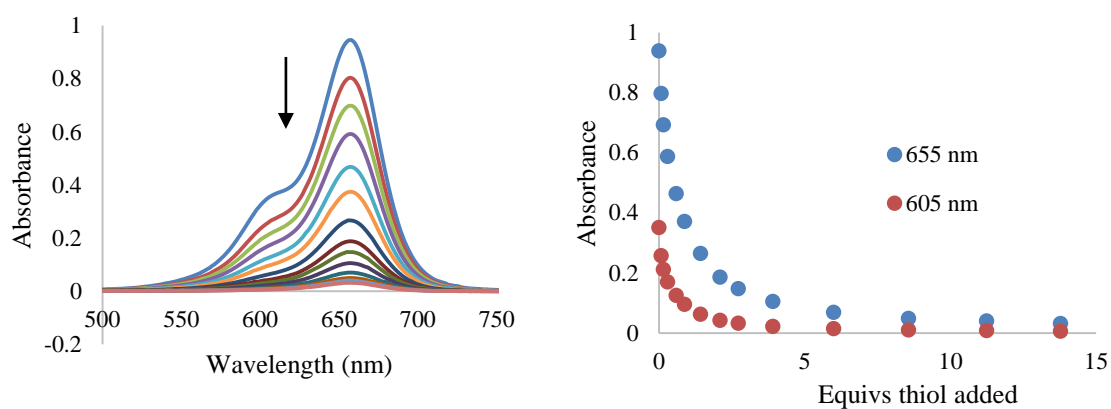


Figure 4.15 Absorbance titration of thiol into **SQ7**. NAC (0-138 μM) was added to **SQ7** (10 μM) in 10 mM phosphate buffer, H_2O , pH 7.0, 0.02% NaN_3 , 25°C, 3 min. equilibration between additions.

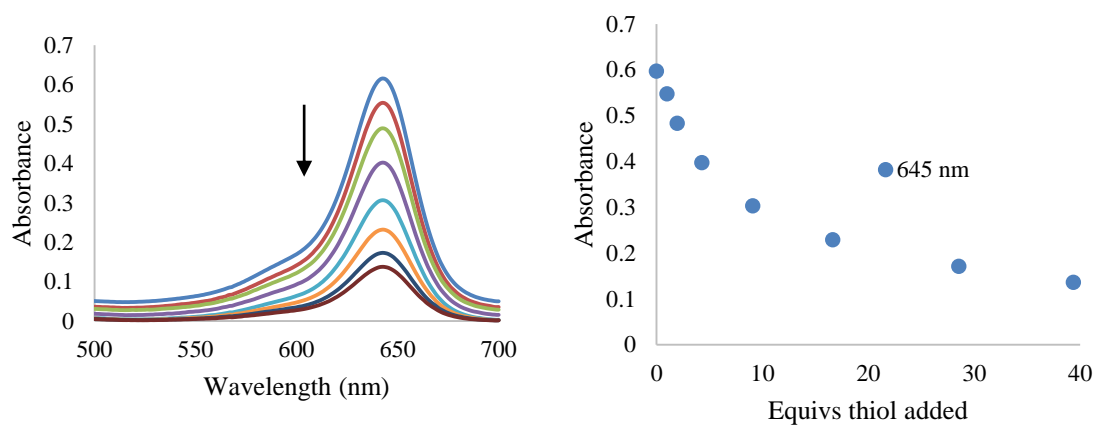


Figure 4.16 Absorbance titration of thiol into **SQ8**. NAC (0-394 μM) was added to **SQ8** (10 μM) in 10 mM phosphate buffer, H_2O , pH 7.0, 0.02% NaN_3 , 25°C, 30 min. equilibration between additions.

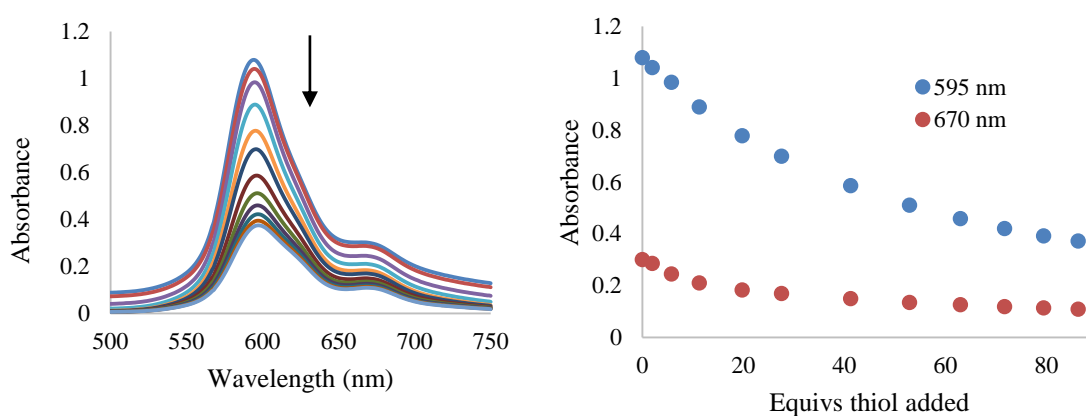


Figure 4.17 Absorbance titration of thiol into **SQ9**. NAC (0-864 μM) was added to **SQ9** (10 μM) in 10 mM phosphate buffer, H_2O , pH 7.0, 0.02% NaN_3 , 25°C, 15 min. equilibration between additions.

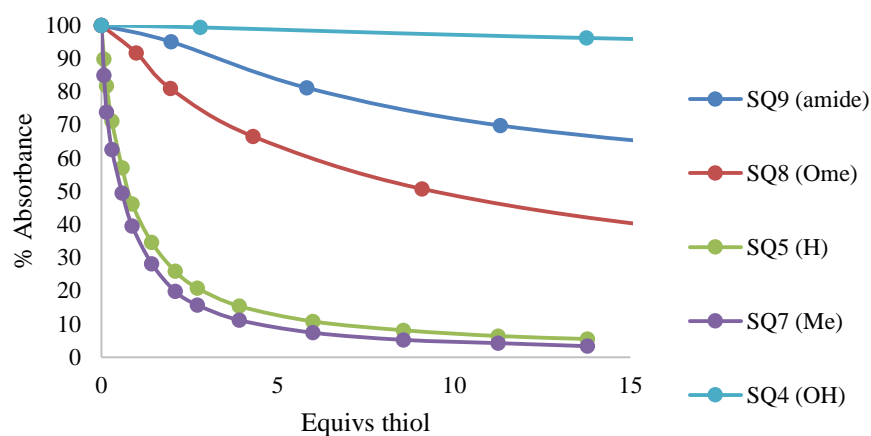


Figure 4.18 Isotherms for thiol addition to **SQ4**, **5**, **7**, **8**, and **9** overlaid. The decrease in absorbance for the monomer peak is described as a percentage of the initial absorbance of each squaraine. See **Figures 4.13-4.17** for experimental details.

From **Figure 4.18** and **Table 4.2**, it is clear that **SQ5** and **SQ7** react with thiols readily, while **SQ8** and **SQ9** do not react as readily. **SQ4** is essentially not reactive, as expected. The reactivity of **SQ7** appears to be slightly greater than for **SQ5**, which is unexpected because a methyl group is electron-donating. Furthermore, **SQ8** is expected to be less reactive than **SQ9**, but we observe the opposite order of reactivity. The observed trends do not seem to agree particularly well with σ values (**Table 4.3**), although the σ_+ values appear to be more appropriate than σ_{para} , which implies that resonance effects are more significant than inductive effects from the substituent for this reaction. A methyl group would not provide resonance stabilization, which would be consistent with the similar reactivity of **SQ5** and **SQ7**. Furthermore, perhaps **SQ7** is more reactive than **SQ5** due to relief of steric strain between the methyl group and the oxygen when planarity is broken. While a methoxy group should be more donating than an amide group, perhaps the monomer-dimer equilibrium for **SQ9** can explain its extra stability compared to **SQ8**. **SQ9** shows a strong tendency to dimerize compared to the other squaraines. A dimer

would be less susceptible to nucleophilic attack statistically (i.e. two sides per two squaraines rather than two sides per one squaraine). Additionally, if the **SQ9** dimer is more stable than or similar in stability to the **SQ9**-thiol complex, then the equilibrium would not clearly favor the SQ-thiol complex. These factors could make the σ value for amide inapplicable in this case.

Table 4.2 *Fitting of binding curves of NAC added to SQ to 1:1 regime.*

Squaraine	Binding constant (M ⁻¹)	Asymptotic Error (%)	Covariance of Fit	RMS of Residuals
SQ7	8.57 x 10 ⁵	5.2	0.017	0.032
SQ5	4.00 x 10 ⁵	2.0	0.005	0.012
SQ8	1.92 x 10 ⁴	0.7	0.003	0.006
SQ9	2.62 x 10 ³	0.6	0.009	0.015
SQ4	6.44 x 10 ²	1.3	0.038	0.019

Table 4.3 *Relevant σ values for the squaraine series.*⁴⁴

Substituent	σ_{para}	σ_+
CH ₃	-0.14	-0.31
NH(CO)CH ₃	0.00	-0.6
OCH ₃	-0.12	-0.78
OH	-0.38	-0.92

4.4.5 ¹H NMR of Thiol Addition to SQ5

The desymmetrization of the squaraine upon thiol addition can also be observed by ¹H NMR. Due to the small amounts of squaraine material we synthesized, we did not perform extensive ¹H NMR studies. However, we did examine the addition of NAC to **SQ5** using this method. The original spectra described here can be found in section 4.11.1. First, a spectrum of **SQ5** in D₂O alone was taken for a baseline. In the aromatic region, there were two doublets from the aromatic protons of **SQ5** (7.74, 6.66 ppm), as well as

additional small doublets from the hydrolyzed **SQ5** (7.62, 6.70, 6.67 ppm), indicating that hydrolysis had already begun even after only a few minutes in solution.

Then, one equivalent of NAC was added, and another spectrum was taken. The solution had decolorized completely immediately upon mixing the sample. The doublets at 7.74 and 6.66 ppm for intact **SQ5** had disappeared completely, and several new doublets had appeared from the asymmetric **SQ5**-thiol product (7.42, 7.30, 6.77, and 6.73 ppm). Finally, this same sample was analyzed again 18 hr later. The spectrum showed that the peaks for **SQ5**-thiol (7.42, 7.30, 6.77, and 6.73 ppm) had diminished relative to the peaks for **SQ5**-water (7.62, 6.70, 6.67 ppm). This evidence is indicative of the reversibility of thiol addition to squaraines. The **SQ5**-thiol peaks diminish over time because the thiols are constantly leaving and re-adding to **SQ5**, and the free thiols in solution are also being oxidized to disulfides in a separate reaction. Over time, the thermodynamic sink for the system is the state in which the thiols are oxidized to disulfides and the squaraines are hydrolyzed.

4.4.6 Testing Reactivity of Squaraines to Amines

The selectivity of squaraines for thiols over amines at pH 6 has been demonstrated in the literature.²³ However, we wanted to be sure that this selectivity was still observed at pH 7 with our squaraines. A large excess of Boc- α -lysine (175 equivalents of amine) was added to a solution of **SQ7**, which was the most reactive squaraine with thiols in our series (**Figure 4.19**). The solution did not decolor, even after one hour. In fact, the absorbance of the squaraine appeared to increase slightly, perhaps due to the change in ionic strength of the solution due to the lysine.

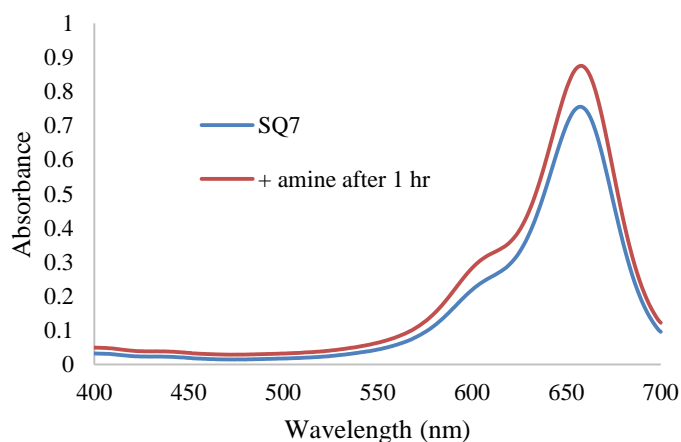


Figure 4.19 *Addition of an amine to SQ7.* Boc- α -lysine (2 mM) was added to SQ7 (11.4 μ M) in 10 mM phosphate buffer, H₂O, pH 7.0, 0.02% NaN₃.

4.5 STUDIES OF SQUARINE BINDING TO SERUM ALBUMIN

The other major component of this work was to investigate the binding behavior of the squaraines to serum albumin. In order to do this, BSA was titrated into each squaraine while monitoring the absorbance and fluorescence of the squaraine. Furthermore, we used indicator displacement studies with dansyl amide (DNSA) and dansyl proline (DP) to probe the site-specificity of the binding of some of the squaraines.

4.5.1 SQ1

BSA was titrated into SQ1 (**Figure 4.20**). The emission at 675 nm decreased as BSA was added, while the emission at 685 nm increased. Saturation was reached at about three equivalents of protein.

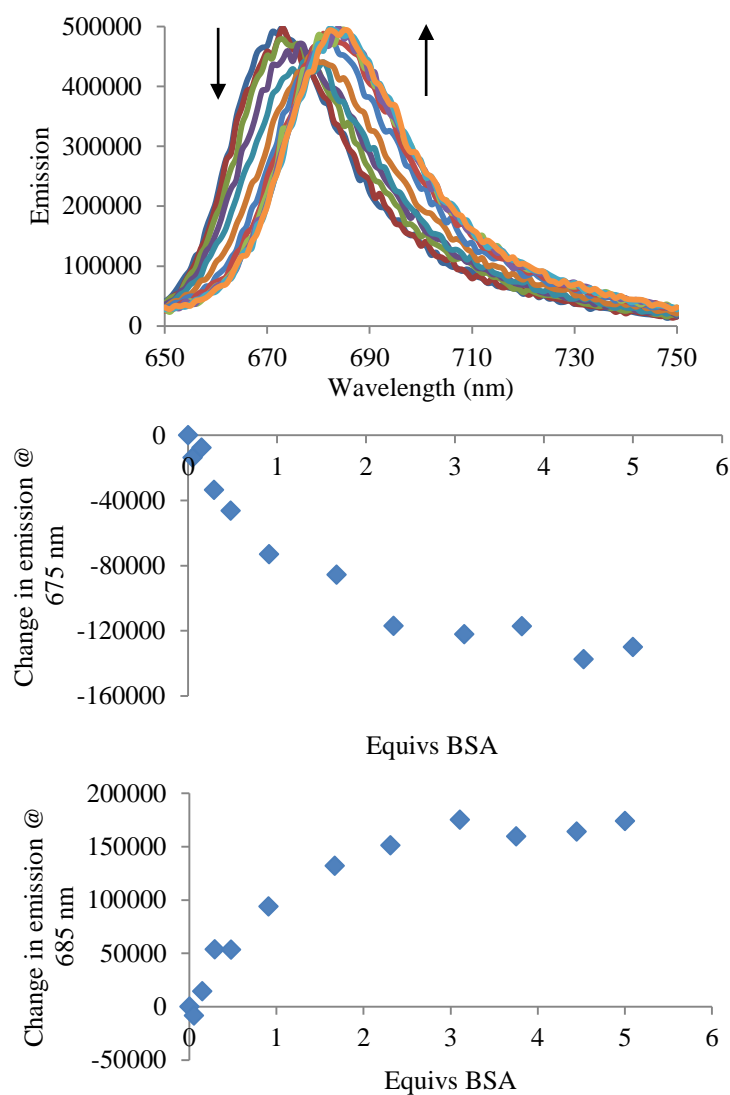


Figure 4.20 *Fluorescence titration of BSA into SQ1.* Addition of BSA (0-100 μM) to SQ1 (20 μM) in 10 mM phosphate buffer, H₂O, pH 7.00, 0.02% NaN₃, λ_{ex} = 640 nm. Top: Spectra, Middle: λ_{em} = 675 nm, Bottom: λ_{em} = 685 nm.

The new peak at 685 nm is likely due to binding of the squaraine into a hydrophobic pocket in serum albumin. Binding stabilizes the excited state of SQ1 resulting in a smaller HOMO-LUMO gap, which leads to a lower energy, higher wavelength emission.

4.5.2 SQ3Br

BSA was titrated into **SQ3Br** (**Figure 4.21**). The emission of the squaraine increased with the addition of BSA and saturated at about one equivalent of BSA. The results from this experiment are consistent with the work by Ramaiah *et al.*³⁸ Our binding isotherm is consistent with their 2:1 **SQ3Br**/BSA binding stoichiometry.

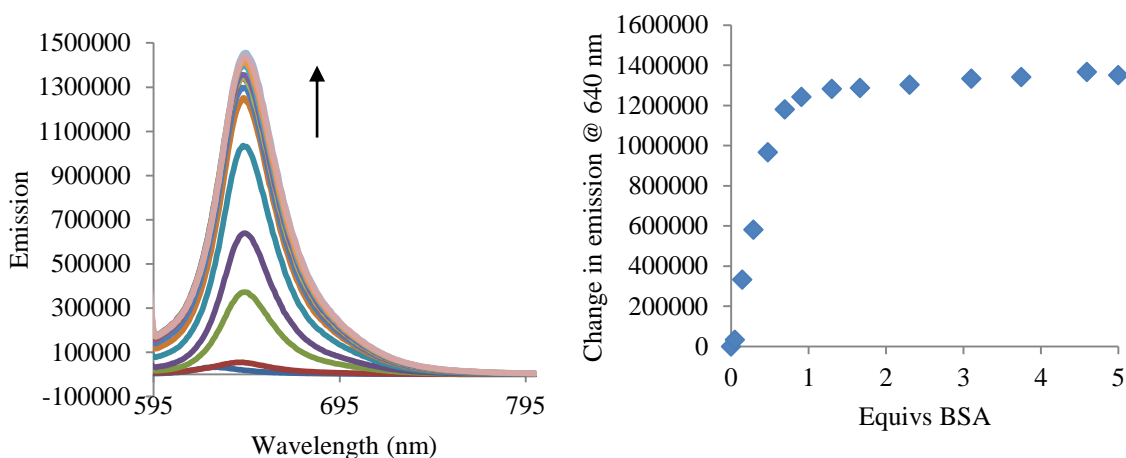


Figure 4.21 Fluorescence titration of BSA into **SQ3Br**. Addition of BSA (0-100 μM) to **SQ3Br** (20 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02 NaN_3 , $\lambda_{\text{ex}} = 575 \text{ nm}$.

4.5.3 SQ4

BSA was titrated into **SQ4**, and the absorbance and fluorescence spectra were recorded (**Figure 4.22** and **4.23**). **SQ4** dimerizes to a significant extent in aqueous solution (**Figure 4.6**), and this monomer-dimer equilibrium leads to complex absorbance and fluorescence behavior of the squaraine upon addition of BSA. In the absorbance titration, the dimer peak at 585 nm consistently decreases, while a new peak (605 nm) grows in. This new peak then diminishes after about 0.7 equivalents of BSA are added. This new peak could represent the bound dimer. The monomer peak increases up to about 0.2 equivalents of BSA, and the decreases up to 0.7 equivalents of BSA. After that, the

absorbance at that wavelength increases up to at least 6 equivalents of protein. The fluorescence titration also showed complex behavior: the emission decreases up to 0.5 equivalents of BSA and then increases. These results are indicative of a binding stoichiometry greater than 1:1. Probably both the monomer and dimer can bind to serum albumin.

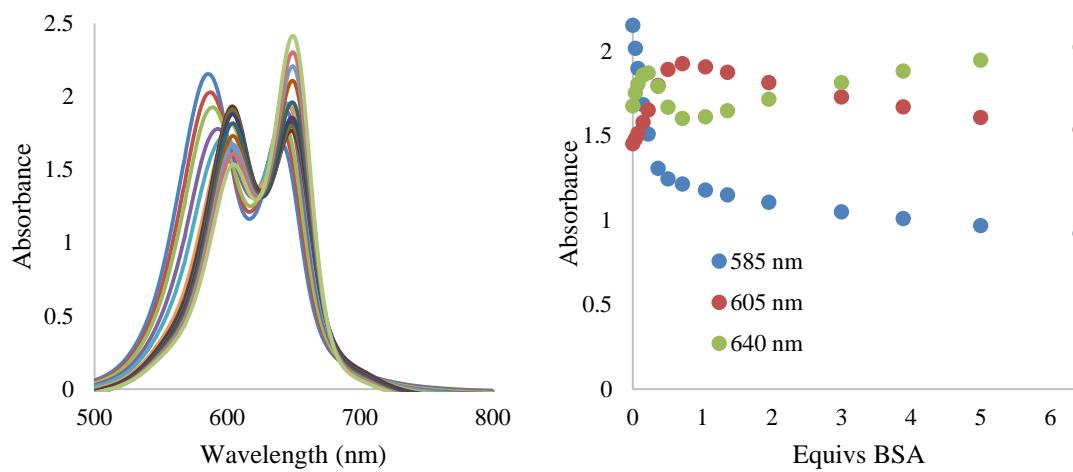


Figure 4.22 Absorbance titration of BSA into **SQ4**. Addition of BSA (0-129 μM) to **SQ4** (20 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 .

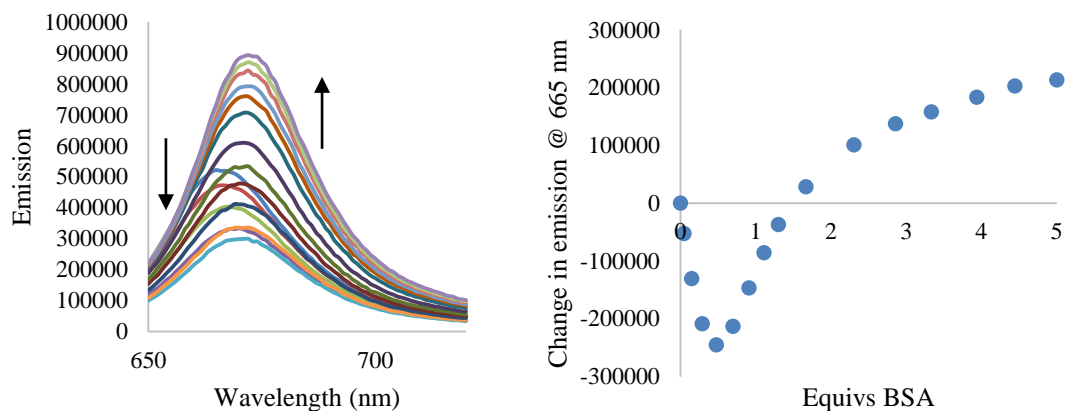


Figure 4.23 Fluorescence titration of BSA into **SQ4**. Addition of BSA (0-50 μM) to **SQ4** (10 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 , $\lambda_{\text{ex}} = 640 \text{ nm}$.

As a comparison, we performed the same titration with HSA (**Figure 4.24**). Similar absorbance behavior to the addition of BSA was observed; however, small differences indicate that the binding of **SQ4** is different to HSA than it is to BSA.

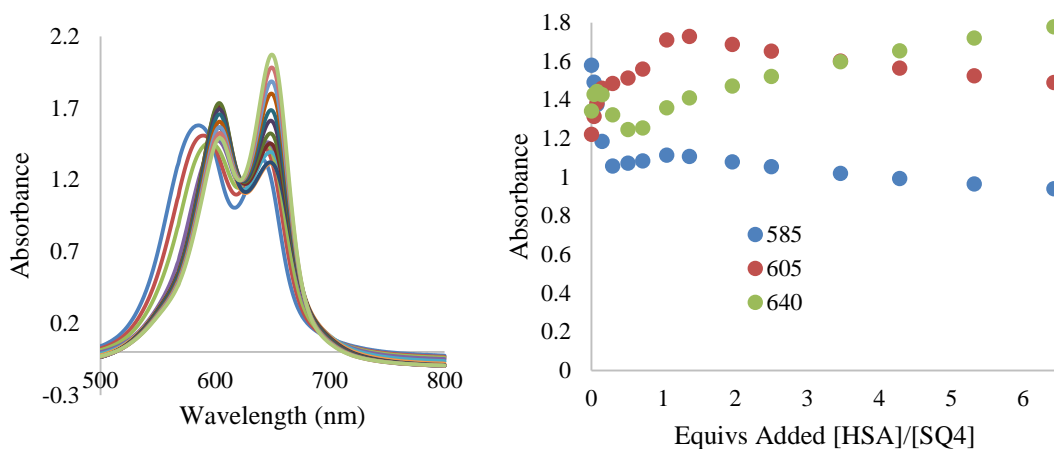


Figure 4.24 Absorbance titration of HSA into **SQ4**. Addition of HSA (0-129 μM) to **SQ4** (20 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 .

4.5.4 SQ5

BSA was titrated into **SQ5**, and the absorbance and fluorescence spectra were recorded (**Figure 4.25** and **4.26**). As BSA is added, the absorbance decreases but the fluorescence increases. We were puzzled by the observation of these two apparently opposing phenomena occurring. The absorbance decrease implies that the squaraine is becoming desymmetrized by reaction with a nucleophile, or the charge transfer band is otherwise being disrupted by steric or electronic factors upon binding. At the same time, the emission of the squaraine is increasing, which implies that binding improves the quantum yield of fluorescence. However, the squaraine is absorbing less light at the excitation wavelength, so the quantum yield increase must be quite large. For example, at three equivalents of BSA, the emission had increased by about fivefold, while the

absorbance at the excitation wavelength had decreased by about fivefold. It is likely that we are observing the net effect of multiple factors. Perhaps **SQ5** is reacting with nucleophilic amino acids in the protein (e.g. lysine, serine), resulting in the decoloration. At the same time, the intact squaraine that remains binds to the protein and experiences a huge increase in its emission due to binding in the hydrophobic pocket/s.

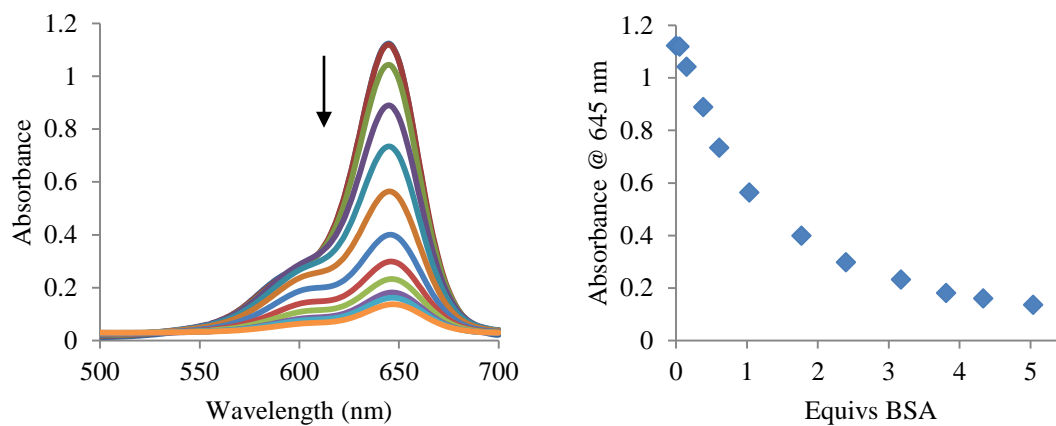


Figure 4.25 Absorbance titration of BSA into **SQ5**. Addition of BSA (0-50 μM) to **SQ5** (10 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 .

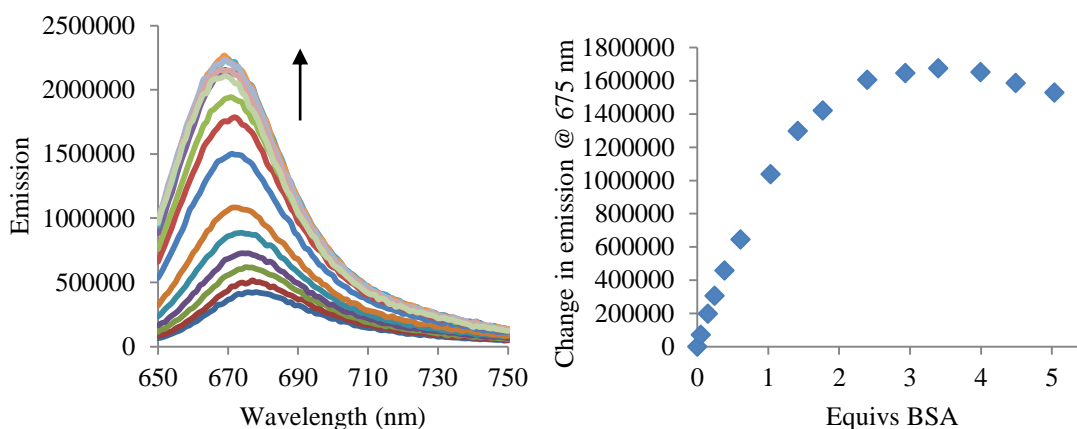


Figure 4.26 *Fluorescence titration of BSA into SQ5.* Addition of BSA (0-50 μM) to SQ5 (10 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 , $\lambda_{\text{ex}} = 640 \text{ nm}$.

In order to test if the squaraine was reacting with the one free cysteine residue of BSA, we prepared a sample of BSA in which that cysteine had been blocked with iodoacetamide (c-BSA). This c-BSA was titrated into SQ5, and the binding isotherm was compared with that of BSA (**Figure 4.27**). The two isotherms are similar, which indicates that blocking the thiol does not significantly change the emission behavior of SQ5 in the presence of serum albumin. Of course, SQ5 could be reacting with other amino acids like lysine. We found with SQ7 that nearly 200 equivalents of amine did not decolor the solution, which indicates that amines will not react with the squaraines. However, the lysines in BSA have an average pK_a of 9.2,⁴⁵ and some individual lysines in the protein could have even lower pK_a s than that. These lysines would be more nucleophilic than a free lysine in solution like the one we used to test amine reactivity.

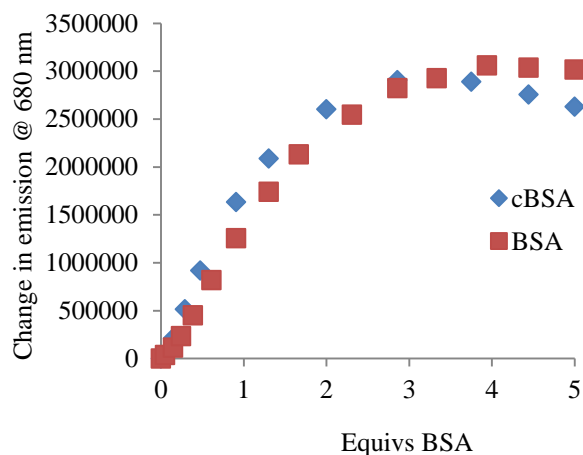


Figure 4.27 Fluorescence titration of BSA and c-BSA into **SQ5**. Addition of BSA (0-100 μM) to **SQ5** (20 μM) and of c-BSA (0-100 μM) to **SQ5** (20 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 , $\lambda_{\text{ex}} = 640 \text{ nm}$, $\lambda_{\text{em}} = 680 \text{ nm}$.

To determine the binding stoichiometry of **SQ5** to BSA, we used the method of continuous variation to generate a Job plot (**Figure 4.28**). The Job plot indicates a 1:1 stoichiometry, which is consistent with the fluorescence binding isotherm as well as the indicator displacement experiments with DNSA and DP that follow.

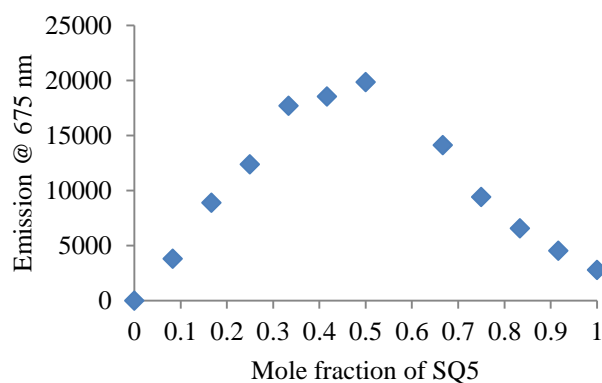


Figure 4.28 Job plot of BSA and **SQ5**. BSA (100 μM) and **SQ5** (100 μM) were mixed in different mole fractions in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 , $\lambda_{\text{ex}} = 640 \text{ nm}$, $\lambda_{\text{em}} = 675 \text{ nm}$.

To probe which site **SQ5** might be binding in, solutions of BSA with dansyl amide (DNSA, Site I-specific) and with dansyl proline (DP, Site II-specific) were titrated with **SQ5** while measuring the emission of DNSA and DP (**Figure 4.29**). Both probes experienced a decrease in their emission, which is indicative of their displacement from BSA. DP experienced a larger decrease than DNSA, so **SQ5** is probably binding in Sudlow Site II, which is consistent with the results for **SQ2** reported by Belfield *et al.*³⁹ Site II generally prefers aromatic molecules with peripherally-located negative charges,⁴⁵ so it makes sense that these sulfonate-functionalized squaraines would bind in Site II.

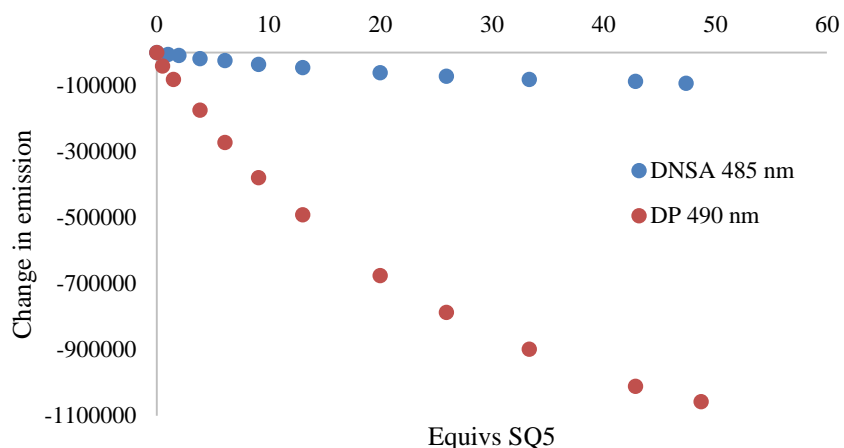


Figure 4.29 *Fluorescence titrations of SQ5 into BSA/DNSA and BSA/DP. SQ5 (0-47 μM) was added to a solution of BSA (10 μM) and DNSA (9 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 , $\lambda_{\text{ex}}=360$ nm, $\lambda_{\text{em}}=485$ nm. SQ5 (0-49 μM) was added to a solution of BSA (10 μM) and DP (9 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 , $\lambda_{\text{ex}}=350$ nm, $\lambda_{\text{em}}=490$ nm.*

As a way of further investigating the bizarre absorbance/emission behavior of **SQ5**, we titrated a solution of BSA/**SQ5** with DNSA and DP (**Figures 4.30-4.33**). If these probes displaced **SQ5** from BSA, we would expect to see a decrease in **SQ5**'s emission. If the absorbance behavior we observe is due to **SQ5** reacting with lysines in the protein, we would expect the probes to have no effect on this reaction, so the absorbance would not change (Scenario 1). However, if the absorbance behavior of **SQ5** can be attributed to some non-covalent effect of being bound in the hydrophobic pocket, then displacing **SQ5** would restore its absorbance (Scenario 2). While these were the scenarios we hypothesized, we did not exactly observe either of these outcomes.

The addition of DNSA slightly increased the emission of **SQ5**, which implies that DNSA has a small cooperative effect on **SQ5** binding to BSA. From the other indicator displacement experiments, we determined that **SQ5** probably binds in Site II, so perhaps the binding of DNSA in Site I induces allosteric changes in BSA that render the Site II

more hospitable to **SQ5**. The absorbance data echoes this apparent cooperative effect of DNSA. The decrease in absorbance of **SQ5** upon addition of DNSA tends to suggest that the decrease observed upon **SQ5** binding to BSA is not due to reaction with lysines. DNSA would not be expected to affect that reaction.

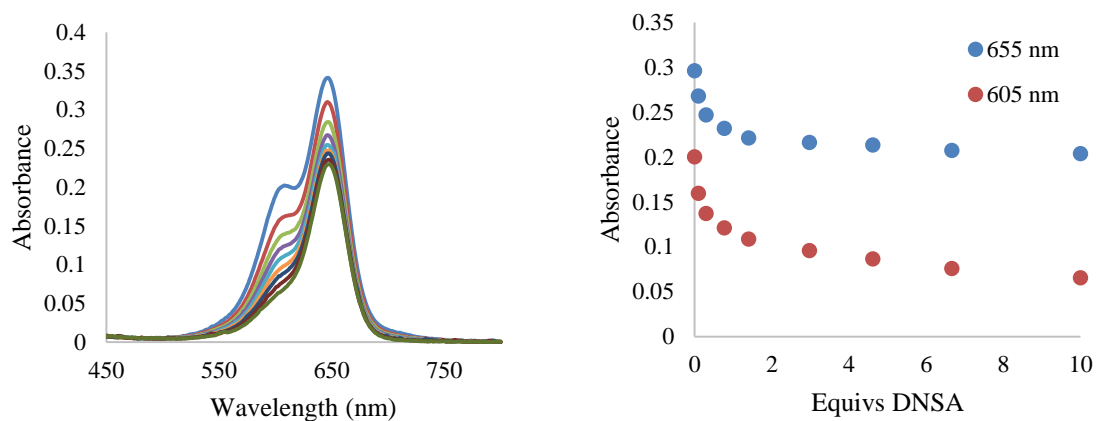


Figure 4.30 *Absorbance titration of DNSA into BSA/SQ5.* Addition of DNSA (0-100 μM) to BSA (10 μM) and **SQ5** (10 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 .

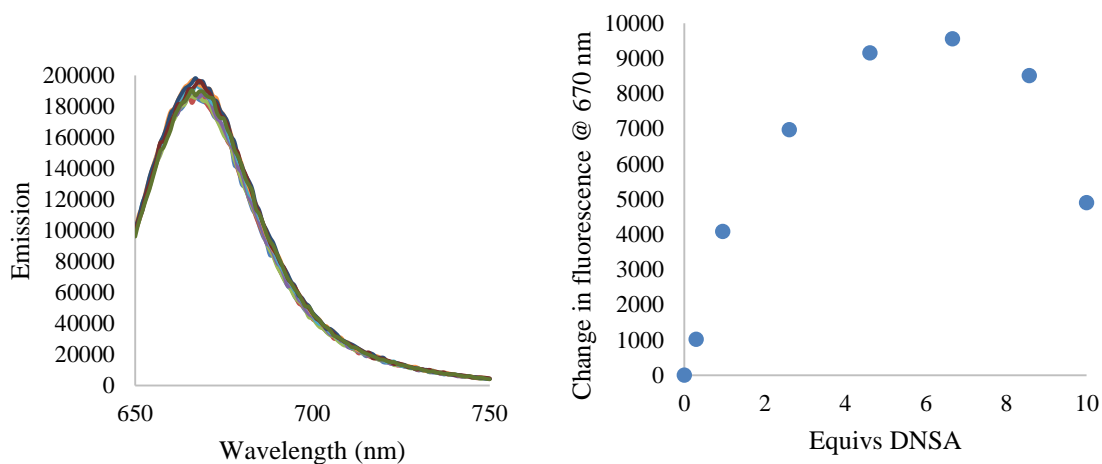


Figure 4.31 *Fluorescence titration of DNSA into BSA/SQ5.* Addition of DNSA (0-100 μM) to BSA (10 μM) and **SQ5** (10 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 , $\lambda_{\text{ex}} = 640 \text{ nm}$, $\lambda_{\text{em}} = 670 \text{ nm}$.

From both the absorbance and fluorescence data, DP displaces **SQ5** starting at about two equivalents, which is consistent with our Site II binding model and our Scenario 2 for indicator displacement. The increase in emission from 0-2 equivalents of DP matches with the decrease in the dimer peak in the absorbance spectrum. The presence of the DP seems to break up the dimer into monomer, which leads to the initial increase in emission upon addition of DP.

These experiments provided some insight into how **SQ5** interacts with BSA. However, it is still unclear exactly what phenomenon the decoloration of **SQ5** in the presence of BSA can be attributed to.

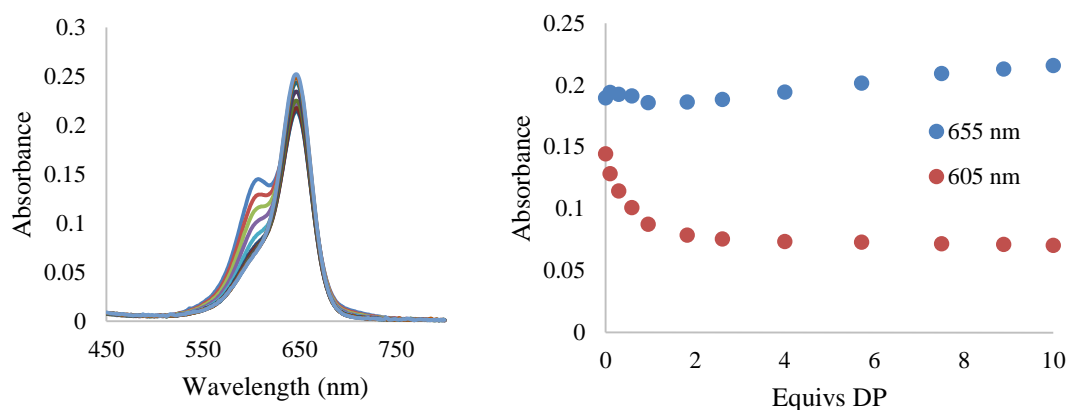


Figure 4.32 Absorbance titration of DP into BSA/SQ5. Addition of DP (0-100 μM) to BSA (10 μM) and SQ5 (10 μM) in 10 mM phosphate buffer, H₂O, pH 7.00, 0.02% NaN₃.

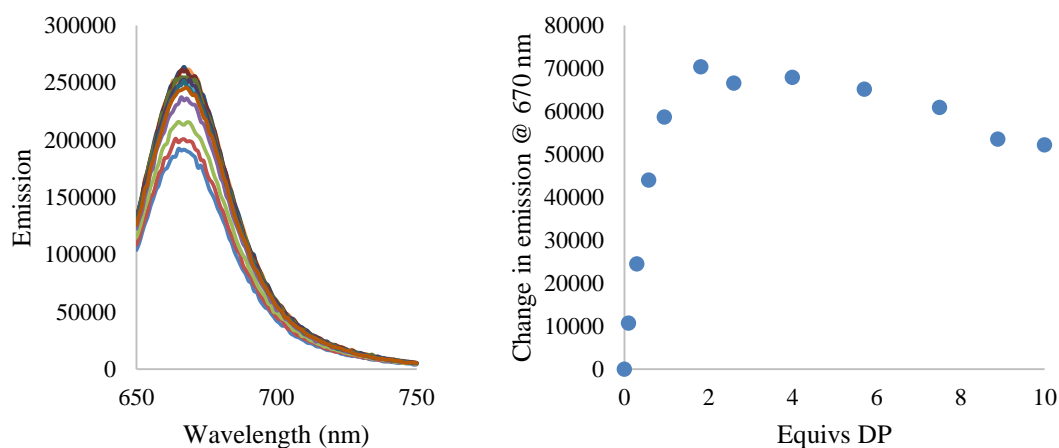


Figure 4.33 Fluorescence titration of DP into BSA/SQ5. Addition of DP (0-100 μM) to BSA (10 μM) and SQ5 (10 μM) in 10 mM phosphate buffer, H₂O, pH 7.00, 0.02% NaN₃, $\lambda_{\text{ex}} = 640 \text{ nm}$, $\lambda_{\text{em}} = 670 \text{ nm}$.

4.5.5 SQ6

BSA was titrated into SQ6, and the fluorescence spectra were recorded (**Figure 4.34**). As with SQ5, the emission increased upon addition of BSA, and the isotherm

implies a 1:1 binding stoichiometry as well. An absorbance titration was not performed, but we did observe that the SQ6 solution decolorized upon the addition of BSA as well.

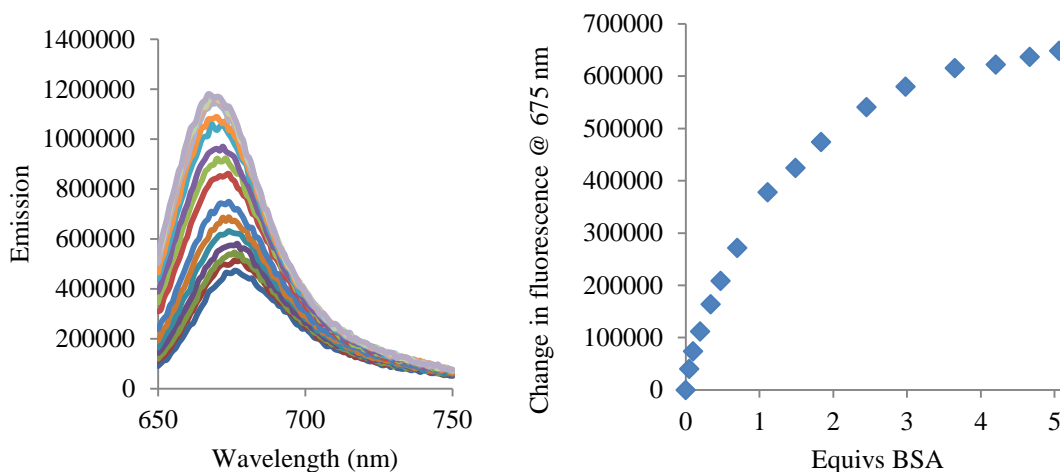


Figure 4.34 *Fluorescence titration of BSA into SQ6.* Addition of BSA (0-50 μM) to SQ6 (10 μM) in 10 mM phosphate buffer, H₂O, pH 7.00, 0.02% NaN₃, < 1% DMSO, $\lambda_{\text{ex}} = 640 \text{ nm}$, $\lambda_{\text{em}} = 675 \text{ nm}$.

4.5.6 SQ7

BSA was titrated into **SQ7**, and the absorbance and fluorescence spectra were recorded (**Figure 4.35** and **4.36**). Again, the same behavior was observed for **SQ7** as for **SQ5** and **6**: the absorbance decreases as BSA is added. However, the shape of the fluorescence binding isotherm is different. The emission increases to about 0.5 equivalents of BSA and then begins to decrease. The shape is similar to the binding isotherm of **SQ4**, but in this titration the emission initially increases rather than decreasing as with **SQ4**. Thus, perhaps either the dimer or two separate monomers can bind to serum albumin. We postulate that in the case of **SQ7**, two monomers bind separately to BSA, which leads to the increase in emission up to 0.5 equivalents. As more BSA is added, there are more

binding sites available and one monomer/protein, and the emission decreases relative to the peak emission when two squaraines are bound.

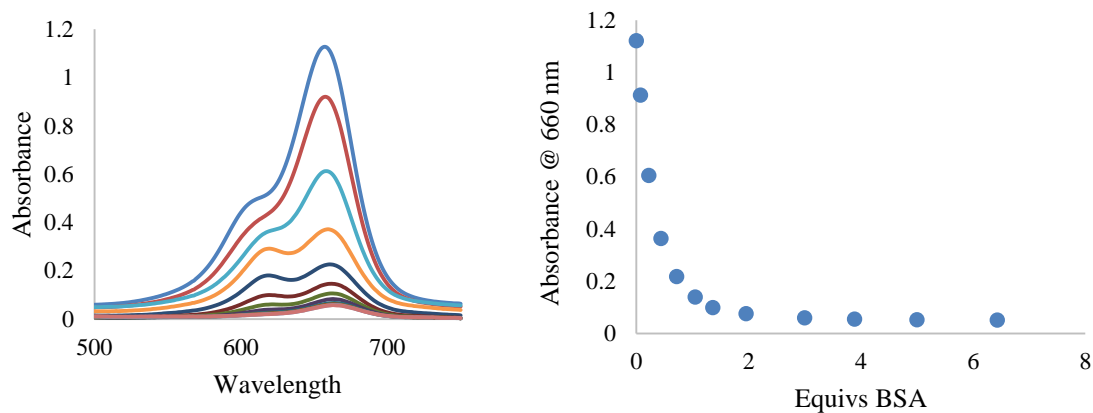


Figure 4.35 Absorbance titration of BSA into **SQ7**. Addition of BSA (0-129 μM) to **SQ7** (20 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 .

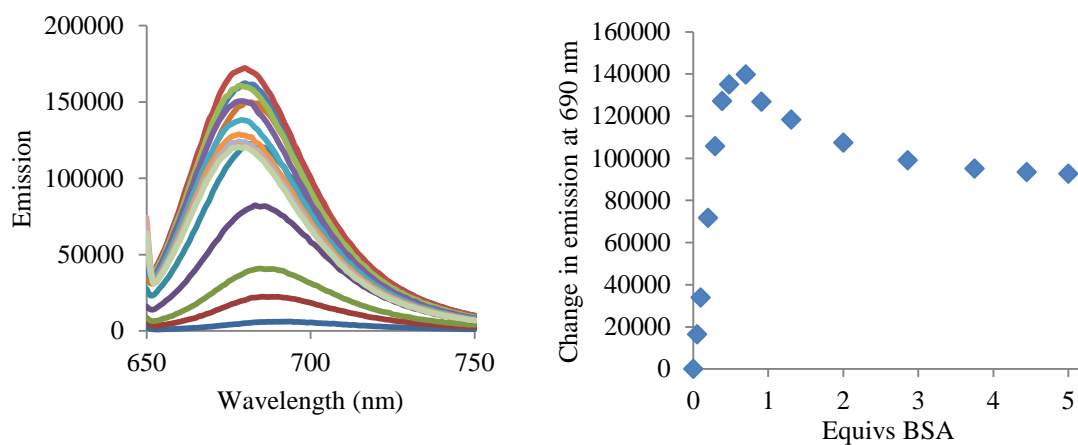


Figure 4.36 Fluorescence titration of BSA into **SQ7**. Addition of BSA (0-100 μM) to **SQ7** (20 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 , $\lambda_{\text{ex}} = 640 \text{ nm}$, $\lambda_{\text{em}} = 690 \text{ nm}$.

4.5.7 SQ8

BSA was titrated into **SQ8**, and the absorbance and fluorescence were monitored (**Figure 4.37** and **4.38**). The results were consistent with **SQ5**. At very small additions of BSA (< 0.2 equivalent), a small blip in the isotherms are observed. This phenomenon is probably due to some higher stoichiometry of binding occurring when very little BSA is present (e.g. dimer or two monomers binding). Clearly, this effect is much smaller for **SQ8** than for **SQ7** or **SQ4**.

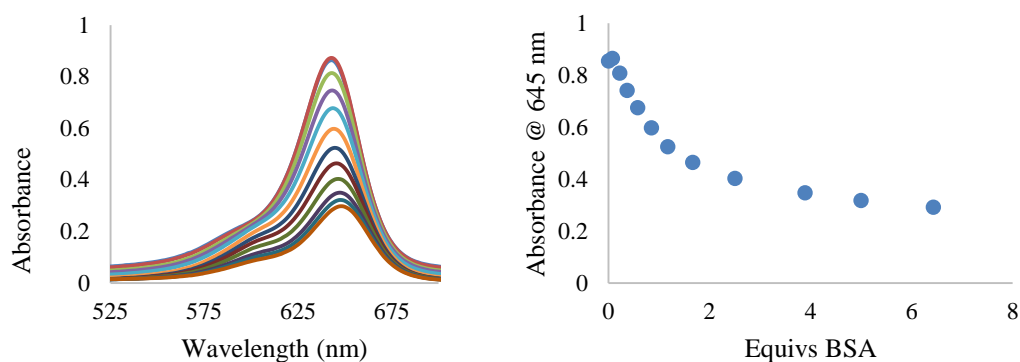


Figure 4.37 *Absorbance titration of BSA into **SQ8**. Addition of BSA (0-129 μM) to **SQ8** (20 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 .*

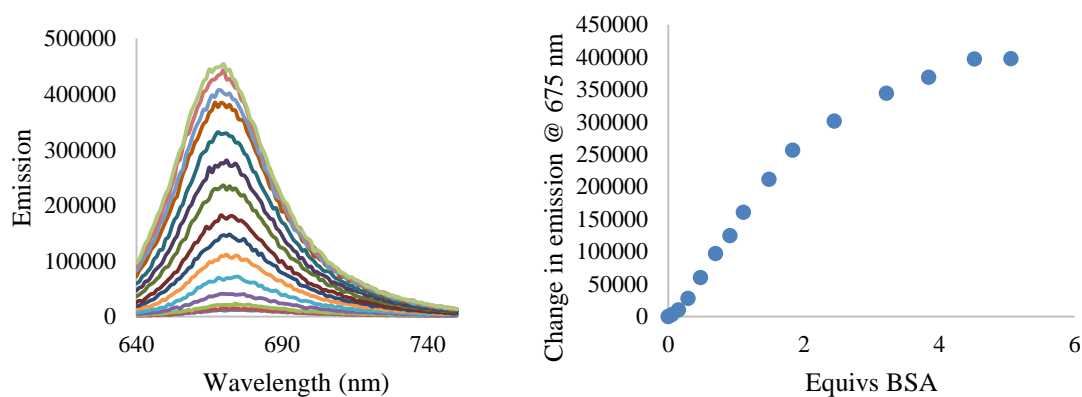


Figure 4.38 *Fluorescence titration of BSA into **SQ8**. Addition of BSA (0-100 μM) to **SQ8** (20 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 , $\lambda_{\text{ex}} = 630 \text{ nm}$, $\lambda_{\text{em}} = 675 \text{ nm}$.*

4.5.8 SQ9

Finally, BSA was titrated into **SQ9**, and the absorbance and fluorescence data was obtained (**Figure 4.39** and **4.40**). From **SQ9**'s absorbance spectrum, we know that it exists primarily as the dimer in solution. Perhaps the amide groups stabilize the dimer by hydrogen bonding, making it the dominant form for **SQ9** but not for **SQ5**, 7, or 8. As BSA was added to **SQ9**, the dimer peak diminished and the monomer peak increased. The two isotherms cross at two equivalents of BSA, although we are not sure what the significance of this crossing point is. This titration implies that the BSA only binds the monomer and so breaks up the dimer. The fluorescence titration supports this notion as the isotherm seems to show a 1:1 binding stoichiometry.

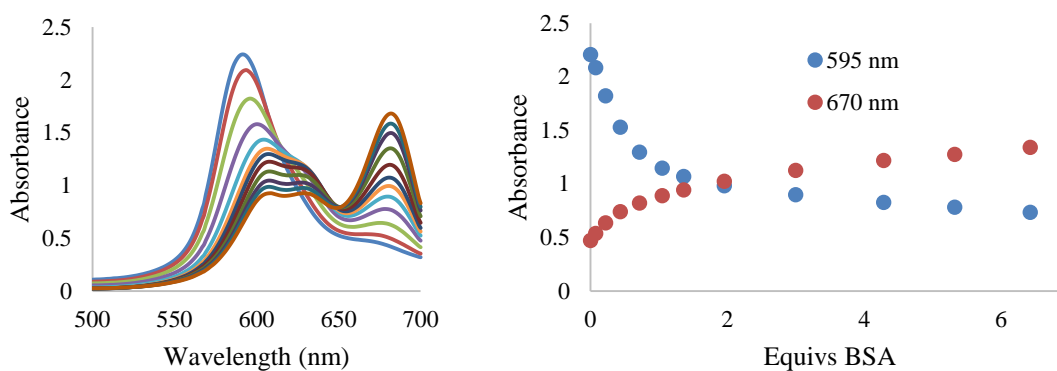


Figure 4.39 Absorbance titration of BSA into **SQ9**. Addition of BSA (0-129 μM) to **SQ9** (20 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 .

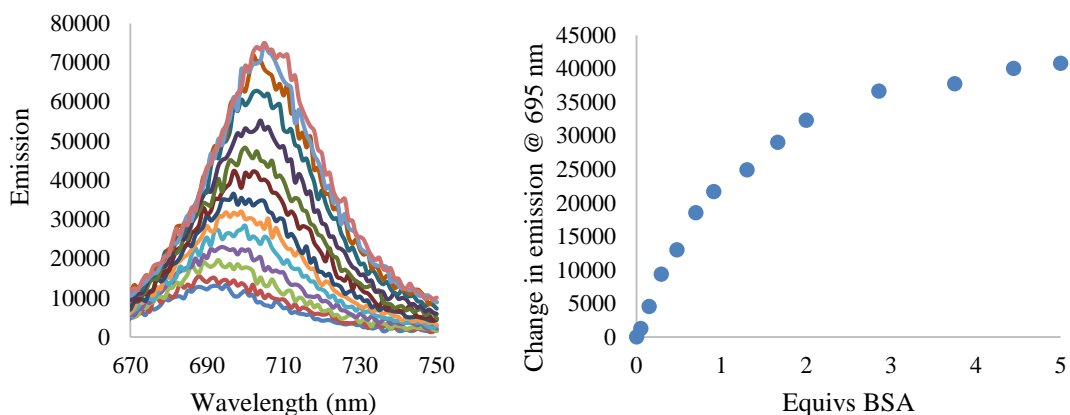


Figure 4.40 *Fluorescence titration of BSA into SQ9.* Addition of BSA (0-100 μM) to SQ9 (20 μM) in 10 mM phosphate buffer, H₂O, pH 7.00, 0.02% NaN₃, λ_{ex} = 650 nm, λ_{em} = 695 nm.

Since SQ9 does not exhibit a decrease in its monomer peak upon BSA addition as SQ5, 7, and 8 do, we wanted to use the same tests as we did for SQ5 to test the effect of displacement by DNSA and DP to compare to SQ5 (Figures 4.41-4.44). When DNSA was added to SQ9/BSA, almost no change in the absorbance spectrum of SQ9 occurred, which matches the small effect of DNSA on SQ5/BSA. A cooperative binding effect to BSA by DNSA and DP was apparent for SQ9 in the fluorescence titrations. The emission of SQ9 increased as DNSA and DP are added. The increase steadily continued at a lower magnitude for DNSA, which was consistent with the allosteric model proposed previously for SQ5. For DP, the binding isotherm implies a 1:1:1 stoichiometry of DP/SQ9/BSA, suggesting that the DP could be sharing Site II with SQ9 rather than displacing the squaraine.

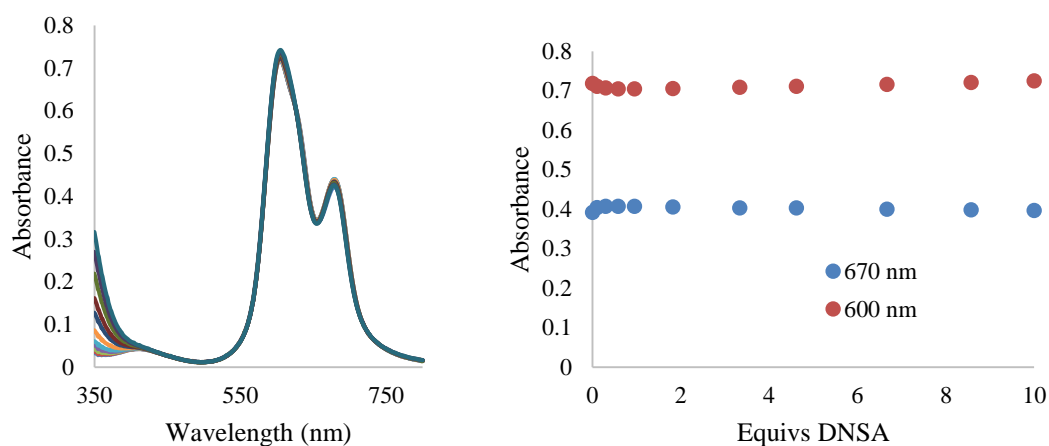


Figure 4.41 *Absorbance titration of DNSA into BSA/SQ9.* Addition of DNSA (0-100 μM) to BSA (10 μM) and **SQ9** (10 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 .

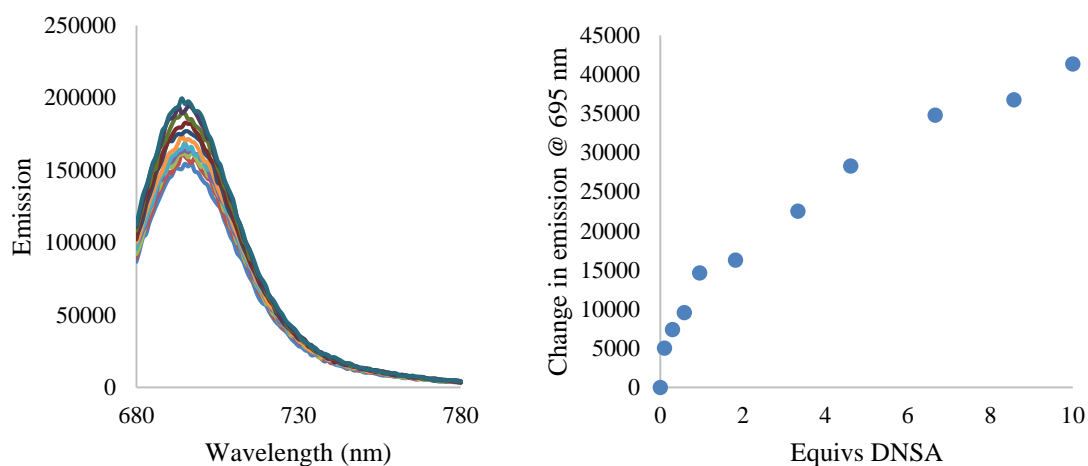


Figure 4.42 *Fluorescence titration of DNSA into BSA/SQ9.* Addition of DNSA (0-100 μM) to BSA (10 μM) and **SQ9** (10 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 , $\lambda_{\text{ex}} = 670 \text{ nm}$, $\lambda_{\text{em}} = 695 \text{ nm}$.

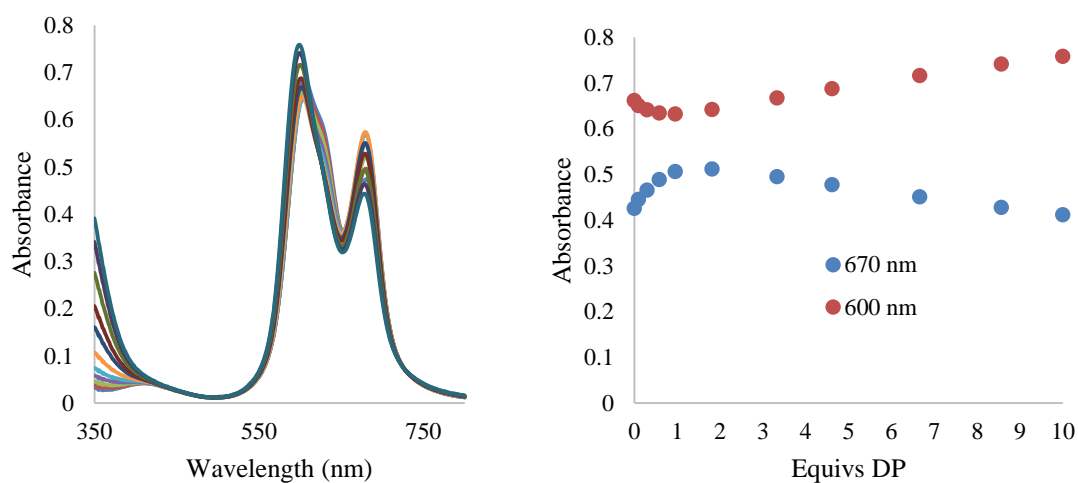


Figure 4.43 *Absorbance titration of DP into BSA/SQ9.* Addition of DP (0-100 μM) to BSA (10 μM) and SQ9 (10 μM) in 10 mM phosphate buffer, H₂O, pH 7.00, 0.02% NaN₃.

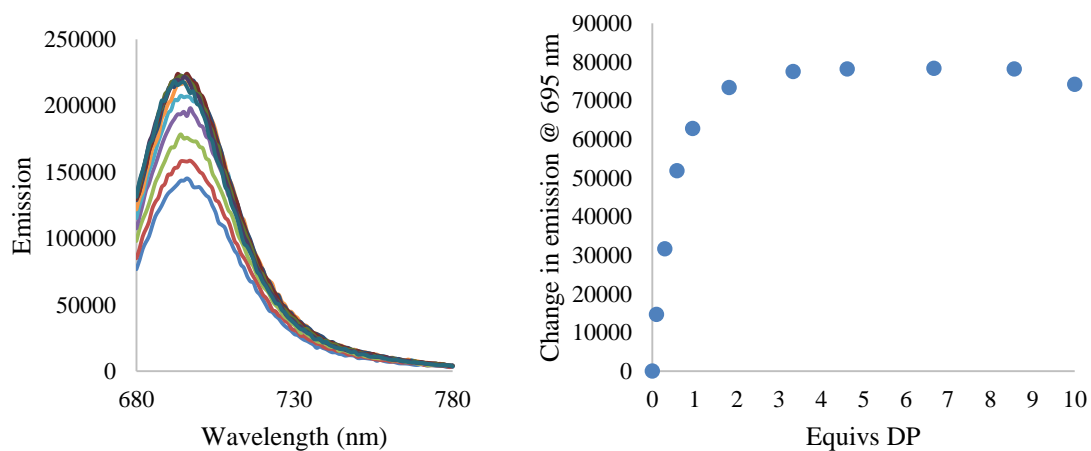


Figure 4.44 *Fluorescence titration of DP into BSA/SQ9.* Addition of DP (0-100 μM) to BSA (10 μM) and SQ9 (10 μM) in 10 mM phosphate buffer, H₂O, pH 7.00, 0.02% NaN₃, $\lambda_{\text{ex}} = 670 \text{ nm}$, $\lambda_{\text{em}} = 695 \text{ nm}$.

4.5.9 Summary

The series of sulfonated squaraines exhibit complex binding behavior to serum albumin, in part due to the monomer-dimer equilibrium. Some of the squaraines bind with greater than 1:1 stoichiometry (**SQ4**, **7**) to BSA. The dominant binding site for these squaraines seems to be Sudlow Site II as evidenced by indicator displacement titrations with DNSA and DP.

For **SQ5**, **6**, **7**, and **8**, the puzzling phenomenon of a decrease in absorbance with a concomitant increase in emission was observed. Our best hypothesis to explain these two apparently opposing effects is that they are the net of multiple effects. The emission of the squaraines increases upon binding to serum albumin due to the hydrophobic environment of the binding pocket. The absorbance of the squaraines decreases for **SQ5**, **7**, and **8** due to reaction with lysines in the protein or perhaps due to steric or electrostatic effects from the binding site.

4.6 THIOL ADDITION TO SQUARAINES IN THE PRESENCE OF SERUM ALBUMIN

Toward the goal of building receptors from serum albumin, squaraines, and thiols, we wanted to test the binding of the squaraine/thiol complex to serum albumin. To do so, we performed indicator displacement experiments with DNSA and DP. A solution of BSA and DNSA (or DP) was titrated with **SQ5**, **SQ5/NAC**, and **SQ5/VVCLKC** (**Figure 4.45** and **4.46**). VVCLKC is a thiol-containing peptide (see Chapter 5 for the synthesis and characterization). The emission of the DNSA or DP was monitored to measure displacement of these probes by the titrants. DP was displaced somewhat less by **SQ5/NAC** and **SQ5/VVCLKC** compared to **SQ5** alone, while DNSA was displaced the same amount by all three titrants. Based on our Site II binding model, this result makes sense.

One interpretation of the result is that the **SQ5**-thiol complexes bind less well to Site II because of the change in shape of the squaraine and added bulk of the thiol that could be extending out of the pocket into solution. Another interpretation is that the apparently lower affinity of the **SQ5**-thiol complex is actually indicative of the complex not binding at all to serum albumin. The thiol adds reversibly to the squaraine, and the squaraine binds reversibly to the protein. These two equilibria may be functioning in parallel such that the SQ-thiol complex itself never binds. This interpretation would explain why **SQ5/VVKLKC** appears to bind better than **SQ5/NAC**. If the complex was binding, the extra bulk of an entire peptide would logically be more cumbersome than cysteine alone. However, VVKLKC is probably less nucleophilic than NAC, so at any given time less squaraine is sequestered as the complex and more is available for binding.

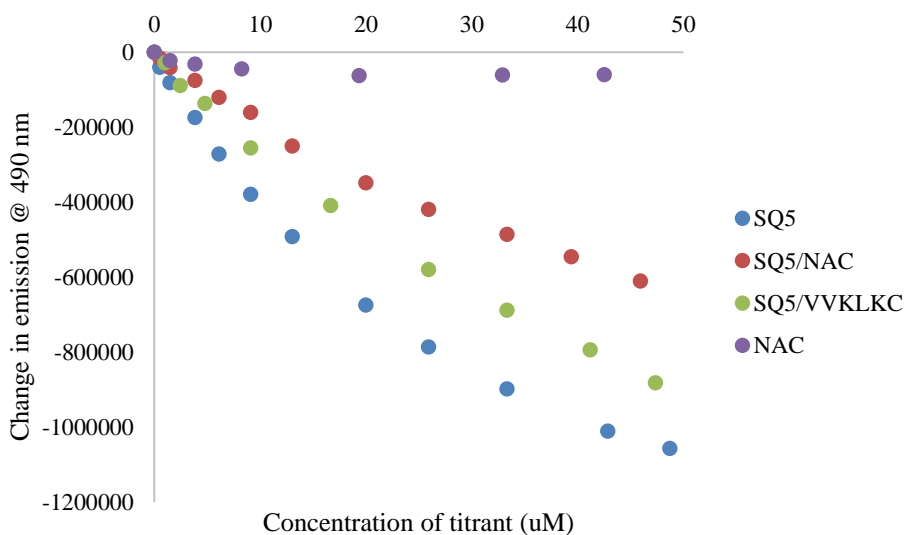


Figure 4.45 Titration of **SQ5**, **SQ5/NAC**, **SQ5/VVKLKC**, and **NAC** into **BSA/DP**. A solution of **BSA** (10 μ M) and **DP** (9 μ M) was titrated with **SQ5** (0-49 μ M), **SQ5/NAC** (0-46 μ M), **SQ5/VVKLKC** (0-47 μ M), and **NAC** (0-43 μ M) in 10 mM phosphate buffer, H₂O, pH 7.00, 0.02% NaN₃, $\lambda_{\text{ex}} = 350$ nm.

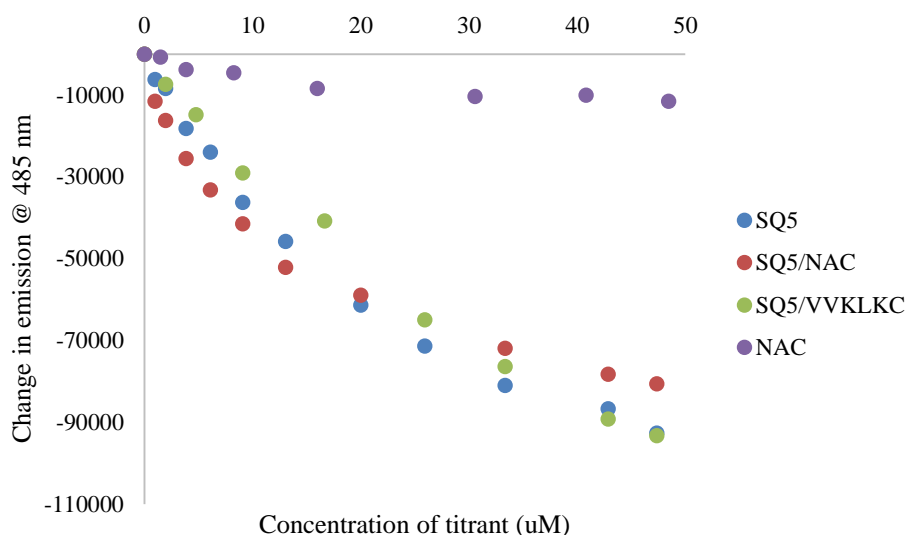


Figure 4.46 Titration of *SQ5*, *SQ5/NAC*, *SQ5/VVKLKC*, and *NAC* into *BSA/DNSA*. A solution of *BSA* ($10\mu\text{M}$) and *DNSA* ($9\mu\text{M}$) was titrated with *SQ5* ($0\text{--}47\mu\text{M}$), *SQ5/NAC* ($0\text{--}47\mu\text{M}$), *SQ5/VVKLKC* ($0\text{--}47\mu\text{M}$), and *NAC* ($0\text{--}48\mu\text{M}$) in 10 mM phosphate buffer, H_2O , $\text{pH } 7.00$, 0.02% NaN_3 , $\lambda_{\text{ex}} = 360\text{ nm}$.

To further explore these phenomena, we investigated whether serum albumin exerts a protective or cooperative effect on thiol addition to squaraines. Ramaiah and coworkers identified a protective effect on phloroglucinol-derived squaraines by β -cyclodextrin, although the cyclodextrin did not protect aniline-derived squaraines from thiol addition.³² To test our squaraines, solutions of *BSA/SQ* were titrated with thiol while monitoring the absorbance. The decrease in absorbance was standardized by expressing it as the percent decrease of the initial absorbance to compare the squaraines to one another (**Figure 4.47**- monomer, **Figure 4.48**- dimer). The original spectra and titration details for the experiments can be found in **4.10.2**. The order of reactivity from **Figure 4.18** is largely retained in the presence of serum albumin.

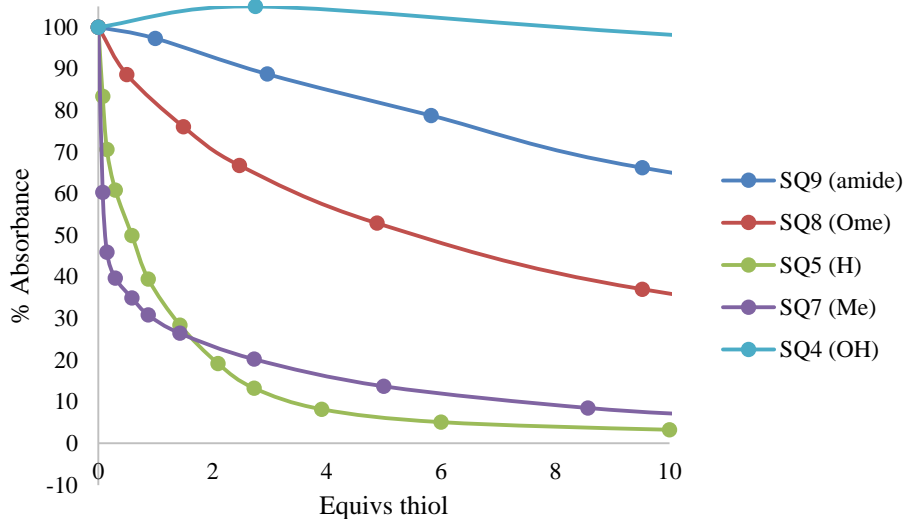


Figure 4.47 Isotherms for thiol addition to all of the SQ/BSA combinations overlaid (*monomer*). The decrease in absorbance for the monomer peak is described as a percentage of the initial absorbance of each squaraine. See **Figures 4.57-4.61** for experimental details and original spectra.

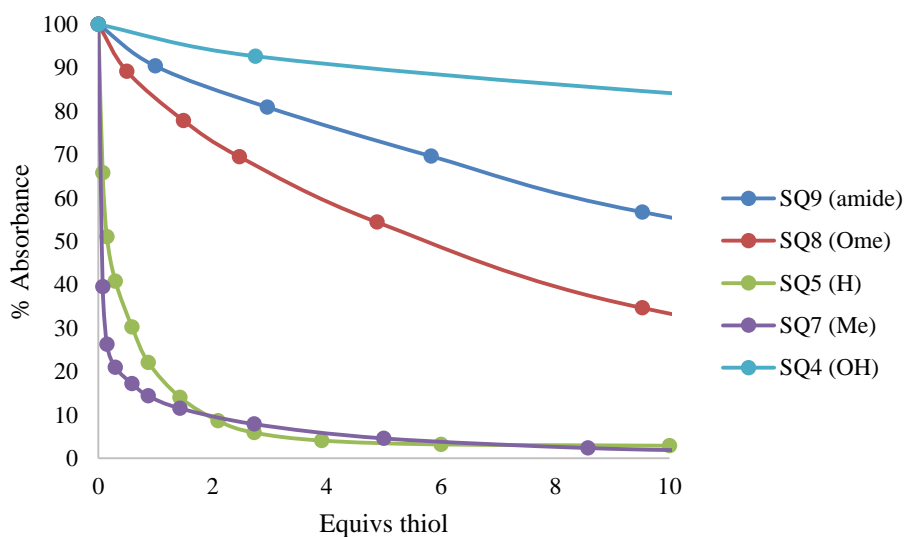


Figure 4.48 Isotherms for thiol addition to all of the SQ/BSA combinations overlaid (*dimer*). The decrease in absorbance for the dimer peak is described as a percentage of the initial absorbance of each squaraine. See **Figures 4.57-4.61** for experimental details and original spectra.

In order to try to detect a protective or cooperative effect by the serum albumin, we plotted the isotherms for the addition of thiol to squaraine together with the addition to BSA/SQ using the percent absorbance (**Figures 4.49-4.53**). In each case, the BSA seems to exert a cooperative effect on thiol addition to the squaraines. One way that BSA could appear to have a cooperative effect is by breaking up the dimers of the squaraines and increasing the relative concentration of monomer. Another explanation for these results is that the squaraines are reacting with lysines in the protein, which decreases the concentration of free squaraine in solution. In that case, the true starting concentration of squaraine is not known, and the equivalents of thiol added cannot be calculated accurately.

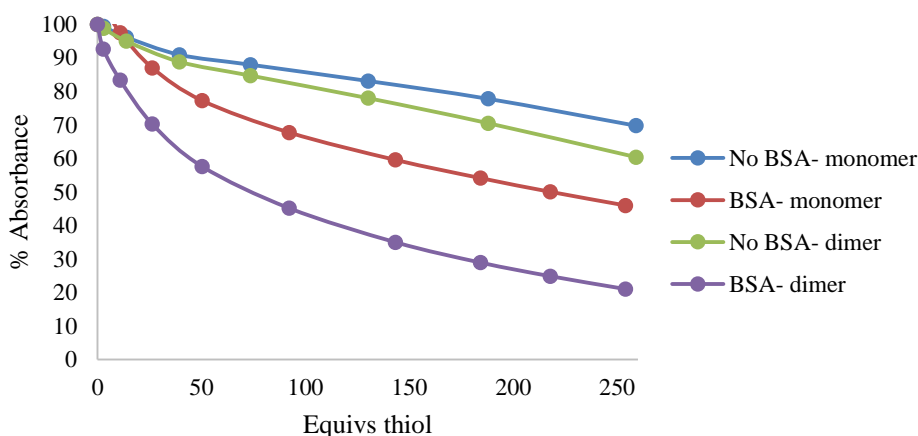


Figure 4.49 *Isotherms for thiol addition to SQ4 and SQ4/BSA.* The decrease in absorbance is described as a percentage of the initial absorbance of SQ4 or SQ4/BSA. See **Figures 4.13** and **4.57** for experimental details and original spectra.

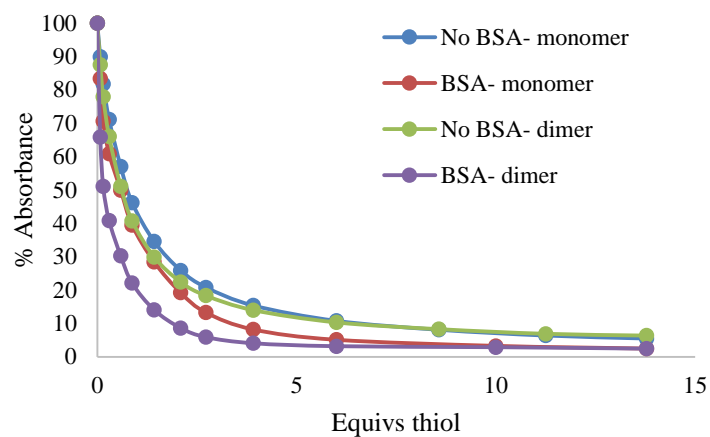


Figure 4.50 *Isotherms for thiol addition to SQ5 and SQ5/BSA.* The decrease in absorbance is described as a percentage of the initial absorbance of SQ5 or SQ5/BSA. See **Figures 4.14** and **4.58** for experimental details and original spectra.

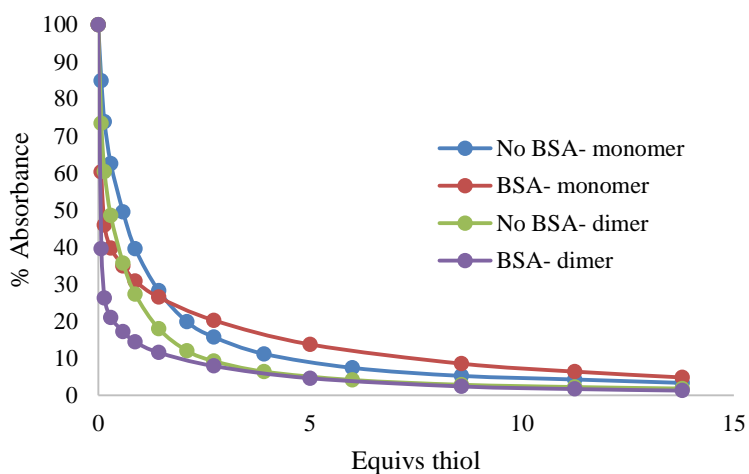


Figure 4.51 *Isotherms for thiol addition to SQ7 and SQ7/BSA.* The decrease in absorbance is described as a percentage of the initial absorbance of SQ7 or SQ7/BSA. See **Figures 4.15** and **4.59** for experimental details and original spectra.

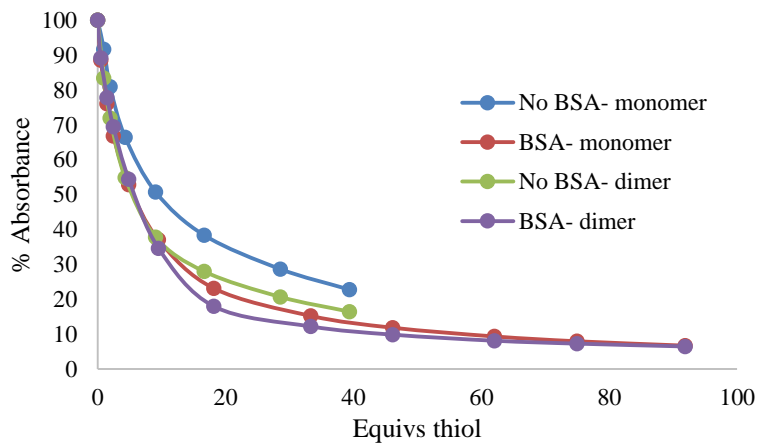


Figure 4.52 *Isotherms for thiol addition to SQ8 and SQ8/BSA.* The decrease in absorbance is described as a percentage of the initial absorbance of SQ8 or SQ8/BSA. See Figures 4.16 and 4.60 for experimental details and original spectra.

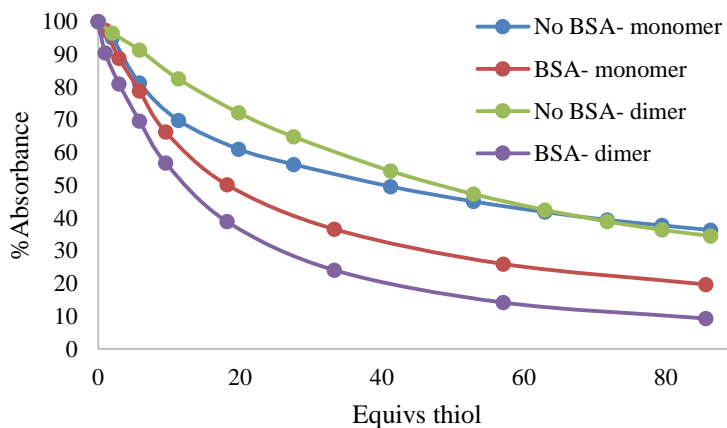


Figure 4.53 *Isotherms for thiol addition to SQ9 and SQ9/BSA.* The decrease in absorbance is described as a percentage of the initial absorbance of SQ9 or SQ9/BSA. See Figures 4.17 and 4.61 for experimental details and original spectra.

4.7 CONCLUSIONS

Squaraines are organic chromophores with absorption and emission in the red to near-IR region of the spectrum, making them useful dyes for a variety of applications. For our purposes, we were interested in the properties of some squaraines to 1) add thiols through a conjugate addition reaction and 2) bind in the hydrophobic pockets of serum albumin. We wanted to exploit these two properties to build dynamic and multivalent receptors from serum albumin, squaraines, and thiol-containing recognition units. As a starting point toward achieving this goal, we synthesized a series of water soluble squaraines and investigated their reactivity to thiols and their affinity for serum albumin.

Aniline-derived squaraines with no *ortho* substitution are known to be very reactive with thiols and other nucleophiles to the point that they are unstable in aqueous solution. On the other hand, squaraines with an *ortho* hydroxyl group are inert to nucleophiles. We hypothesized that we could tune the reactivity of aniline-derived squaraines by the substitution at the *ortho* position of the benzene ring. Thus, we synthesized squaraines with different electron-donating groups at the *ortho* position and studied their reactivity toward cysteine. We found that the different substitution did indeed change the reactivity of the squaraine, and we achieved intermediate reactivity between the H-substituted squaraine and the OH-substituted squaraine. Although the specific order of reactivity that we found did not conform to σ values for these groups. We were able to rationalize these inconsistencies based on additional factors for the squaraines that the σ values would not take into account.

We also investigated the binding properties of these squaraines to serum albumin. The sulfonated squaraines exhibit an increase in their emission upon binding to BSA. Some of the squaraines bind with greater than 1:1 stoichiometry to BSA, which could be due to the dimer binding in one site or multiple monomers in different sites. The dominant

binding site for these squaraines seems to be Sudlow Site II, as evidenced by indicator displacement titrations with DNSA and DP. Finally, we tested the effect of serum albumin on the thiol addition to the squaraines and detected a cooperative effect by the protein.

4.8 FUTURE WORK

While we made significant progress toward understanding the properties of these squaraines, questions remain that could be answered by further investigation. The reactivity trend of the substituted squaraines could be further elucidated by synthesizing more derivatives (e.g. X = C(CH₃)₃, F, or Ph) for testing. Derivatives with substitution at both ortho positions could also be made to see if additive or blended effects could be achieved for further refinement of reactivity. The origin of the decoloration of the squaraines in the presence of serum albumin is still unclear. The squaraines could be titrated with other proteins to test if the decrease in absorbance is due to lysine addition rather than non-covalent effects due to binding, although the exact reactivity of the lysines in serum albumin might not be represented by other proteins. Crystal structures of the serum albumin-squaraine complex could be obtained. Finally, the squaraines could be combined with serum albumin and thiol-containing recognition units in the presence of an analyte of interest to test if receptors will assemble as we had envisioned. A specific signaling or characterization method will need to be devised to measure the purported template effect.

4.9 CONTRIBUTIONS

J. Logan Bachman, an undergraduate student, assisted in the synthesis of several of the squaraines.

4.10 EXPERIMENTAL METHODS

4.10.1 General

Unless otherwise indicated, chemicals and reagents were obtained from Sigma Aldrich and used without further purification. Dansyl amide and dansyl proline piperidinium salt were obtained from TCI America.

HPLC-grade solvents were prepared with 0.1% TFA (v/v) and filtered through 0.2 micron filter. HPLC was performed on a Shimadzu instrument with a preparative C-18 column using a water and acetonitrile mobile phase. Other reverse-phase separation was performed on a Teledyne Isco CombiFlash instrument using a water and acetonitrile mobile phase without TFA. RediSep C-18 columns were used (4.3 gram flash column).

LC-MS analysis was performed on an Agilent 1200 Series HPLC with an Agilent 6130 single quadrupole mass spectrometer (ESI and APCI ionization). NMR spectra were obtained on a 400 MHz Varian instrument.

The fluorescence experiments were performed with a PTI fluorimeter using an 814 photomultiplier detection system and a 75W xenon short arc lamp. The data for the Job plot was obtained from a Biotek Synergy 2 Multimode Microplate Reader. The absorbance experiments were performed with a Beckman Coulter DU 800 Spectrophotometer or an Agilent Cary 100 UV-VIS.

The binding constants were calculated using Open Data Fit by the Centre for Bio-Nano Science (CBNS) at <http://supramolecular.org/>.

4.10.2 Synthesis

The literature source for each compound is indicated next to the compound name. If there is no source given, it is a new compound. For these squaraines that had not previously been synthesized, ¹H NMR and high-resolution MS data were obtained. These

squaraines have relatively low solubility in any solvent (< 1 mM) and are unstable in their best solvent (water), decomposing over the course of several hours. Thus, it was not possible to obtain ¹³C NMR data due to the high concentrations and long scan times necessary.

4.10.2.1 SQ1

*Di-tert-butyl 2,2'-((3-(benzyloxy)phenyl)azanediyl)diacetate (4.2)*⁴⁰

To 10 mL of acetonitrile was added 3-benzyloxyaniline (500 mg, 2.5 mmol, 1 eq) and proton sponge (1.18 g, 5.5 mmol, 2.2 eq). Then t-butylbromoacetate (0.78 mL, 5.3 mmol, 2.1 eq) was added, and the reaction was refluxed for 20 hr. After 20 hr, another 0.4 mL of t-butylbromoacetate (2.7 mmol) and 0.59 g of proton sponge (2.75 mmol) were added, and the reaction was refluxed for an additional 5 hr. The reaction mixture was allowed to cool and was filtered to remove insoluble material. Water was added and the solution was extracted four times with ethyl acetate and washed three times with brine. The organic layer was dried with Na₂SO₄, and the solvent was removed under reduced pressure. The crude material was purified by silica chromatography with 10-50% ethyl acetate in hexanes. The compound was a yellow oil (1.0 g, 94% yield). ¹H-NMR (400 MHz, CDCl₃, ppm): δ 1.44 (s, 18H, CH₃), 3.97 (s, 4H, CH₂), 6.19 (d, 1H, Ar-H), 6.22 (s, 1H, Ar-H), 6.39 (d, 1H, Ar-H), 7.09 (t, 1H, Ar-H), 7.36 (m, 5H, Ar-H of benzyl).

*Di-tert-butyl 2,2'-((3-hydroxyphenyl)azanediyl)diacetate (4.3)*⁴⁰

To a flask was added **4.2** (1g, 2.34 mmol, 1 eq), Pd/C (1.51 g), cyclohexene (17.45 mL, 271 mmol, 116 eq), and 100 mL of ethanol. The reaction was refluxed for 2 hr. The reaction mixture was cooled and filtered through Celite. The solvent was removed under reduced pressure to yield a clear oil. The product required no purification (0.75 g, 95%

yield). ¹H-NMR (400 MHz, CDCl₃, ppm): δ 1.42 (s, 18H, CH₃), 3.98 (s, 4H, CH₂), 6.13 (s, 1H, Ar-H), 6.16 (d, 1H, Ar-H), 6.25 (d, 1H, Ar-H), 7.09 (t, 1H, Ar-H).

t-Butyl-Protected **SQ1** (**4**)⁴⁰

To a flask was added squaric acid (70 mg, 0.61 mmol, 1 eq), **4.3** (440 mg, 1.30 mmol, 2.1 eq), 16 mL benzene, and 8 mL n-butanol (2:1). A Dean-Stark trap was attached, and the reaction was refluxed at 100°C for 18 hr. The reaction was allowed to cool, and the reaction mixture was concentrated under reduced pressure. The crude material was purified on a silica column with DCM/MeOH. The product was a blue solid (300 mg, 65% yield). ¹H-NMR (400 MHz, CDCl₃, ppm): δ 1.46 (s, 36H, CH₃), 4.08 (s, 8H, CH₂), 6.06 (s, 2H, Ar-H), 6.24 (d, 2H, Ar-H), 7.90 (d, 2H, Ar-H). MS (ESI): 753.3 (M+1).

SQ1⁴⁰

The *t*-butyl-protected squararine (100 mg, 1.33 mmol) was placed in 3.75 mL DCM, 11 mL TFA, and 150 μL water. The solution was stirred for 18 hr at room temperature. The solvent was removed under reduced pressure. The product required no purification and was a green iridescent solid (66 mg, 92% yield). ¹H-NMR (400 MHz, d-DMSO, ppm): δ 4.35 (s, 8H, CH₂), 6.08 (s, 2H, Ar-H), 6.54 (d, 2H, Ar-H), 7.76 (d, 2H, Ar-H). MS (ESI): 528.4 (M+1).

4.10.2.2 SQ2

3-(1,1,2-trimethyl-1*H*-benzo[*e*]indol-3-ium-3-yl)propane-1-sulfonate (**4.1**)³⁹

To 1,1,2-trimethyl-1-benz[*e*]indole (2g, 9.56 mmol, 1 eq) was added 15 mL of *p*-xylene and 1,3-propane sultone (1 mL, 11.17 mmol, 1.2 eq). The reaction was refluxed for 6 hr and then cooled. The reaction mixture was filtered to collect the solid. The solid was washed with DCM and diethyl ether and dried. The product required no further purification (2.66 g, 84% yield). ¹H-NMR (400 MHz, d-DMSO, ppm): δ 1.75 (s, 6H, CH₃), 2.21 (m,

2H, CH₂), 2.67 (t, 2H, CH₂), 2.93 (s, 3H, CH₃), 4.77 (t, 2H, CH₂), 7.72 (t, 1H, Ar-H), 7.78 (t, 1H, Ar-H), 8.24 (m, 3H, Ar-H), 8.35 (m, 1H, Ar-H).

SQ2³⁹

To a flask was added squaric acid (25 mg, 2.19 mmol, 1 eq), **4.1** (1.5 g, 4.52 mmol, 2.1 eq), pyridine (3 mL), and n-butanol (10 mL). A Dean-Stark trap was attached, and the reaction was refluxed for 5 hr. After the reaction cooled, ether was added to precipitate the product. The product was isolated by filtration and recrystallized three times in methanol. The product was a green solid (236 mg, 12% yield, pyridinium salt). ¹H-NMR (400 MHz, d-DMSO, ppm): δ 1.95 (s, 12H, CH₃), 2.06 (m, 4H, CH₂), 2.58 (t, 4H, CH₂), 4.35 (m, 4H, CH₂), 5.88 (s, 2H, CH), 7.44 (t, 2H, Ar-H), 7.61 (t, 2H, Ar-H), 7.81 (d, 2H, Ar-H), 7.98 (d, 8H, Ar-H), 8.28 (m, 2H, Ar-H), 8.50 (m, 2H, Ar-H), 8.79 (s, 4H, Ar-H). MS (ESI): 739.2 (M+1).

4.10.2.3 ***SQ3Br***

SQ3¹

To a flask was added squaric acid (250 mg, 2.19 mmol, 1 eq), phloroglucinol (552 mg, 4.38, 2 eq), and 5 mL of acetic acid. The reaction was refluxed for 4 hr. The reaction was cooled, and the product was collected by filtration. The crude product was recrystallized from acetic acid. The product was a purple solid (300 mg, 41% yield). MS (ESI): 329.2 (M-1).

SQ3Br⁴⁶

SQ3 (100 mg, 0.3 mmol, 1 eq) was placed in 70 mL of acetic acid. The solution was stirred and gently heated (up to 50°C) to dissolve **SQ3**. To this solution was added a solution of bromine (0.07 mL, 13.6 mmol, 4.5 eq) in acetic acid dropwise over 1 hr. The

reaction turned from purple to green. The precipitate was collected. The product required no further purification (68 mg, 35% yield). MS (ESI): 644.7 (M-1).

4.10.2.4 SQ4

*3-(ethylamino)phenol (4.5)*⁴⁷

To a flask was added 3-aminophenol (4.0 g, 37 mmol, 1 eq), potassium carbonate (5.0 g, 37 mmol, 1 eq), and 20 mL of DMF. Then iodoethane (2.9 mL, 37 mmol, 1 eq) was added, and the reaction was heated to 100°C for 2 hr. The reaction was cooled to room temperature and filtered to remove any solid. Then water (20 mL) was added, and the solution was extracted three times with ethyl acetate. The organic layer was washed three times with water to remove DMF. The product was purified on a silica column with 0-30% ethyl acetate in hexanes (1.25 g, 25% yield). ¹H-NMR (400 MHz, CDCl₃, ppm): δ 1.25 (t, 3H, CH₃), 3.14 (m, 2H, CH₂), 6.13 (t, 1H, Ar-H), 6.18 (m, 1H, Ar-H), 6.23 (m, 1H, Ar-H), 7.02 (t, 1H, Ar-H).

*3-((3-hydroxyphenyl)(ethyl)ammonio)propane-1-sulfonate (4.8)*⁴¹

To a flask was added **4.5** (1.25 g, 9.11 mmol, 1 eq), 1,3-propanesultone (0.88 mL, 10.02 mmol, 1.2 eq), and 10 mL of isopropanol. The reaction was refluxed for 3 hr. The precipitate was collected and washed with isopropanol. The product required no further purification (1.68 g, 71% yield). ¹H-NMR (400 MHz, d-DMSO, ppm): δ 0.97 (t, 3H, CH₃), 1.73 (m, 2H, CH₂), 2.58 (t, 2H, CH₂), 3.60 (m, 4H, CH₂), 6.90 (broad s, 1H, Ar-H), 7.05 (broad s, 2H, Ar-H), 7.37 (broad s, 1H, Ar-H), 10.23 (broad s, 1H, OH), 11.44 (broad s, 1H, NH).

SQ4⁴¹

To a flask was added squaric acid (103 mg, 0.9 mmol, 1 eq), **4.8** (490 mg, 1.89 mmol, 2.1 eq), sodium bicarbonate (159 mg, 1.89 mmol, 2.1 eq), 10 mL butanol, 5 mL

benzene, and 3 mL DMSO. The reaction was refluxed for 3 hr. Ether was added to precipitate the product. The product was very wet, so it was taken up in methanol. The methanol was filtered to remove solid, and then the methanol was removed under reduced pressure to give a green solid. The crude product was purified by RP-HPLC in water/acetonitrile with 0.1% TFA. The product was a green iridescent solid (7.4 mg, 1.4% yield). MS (ESI): 596.2, 297.1 (M-1). HRMS (ESI): expected m/z 297.07, found m/z 297.07 ((M-1)/2); expected m/z 617.12, found m/z 617.13 (M-1).

4.10.2.5 SQ5

*Sodium 3-(ethyl(phenyl)amino)propane-1-sulfonate (4.9)*⁴⁸

To a flask was added N-ethylaniline (2.52 mL, 20.0 mmol, 1 eq), 1,3-propanesultone (1.93 mL, 22.0 mmol, 1.2 eq), and 20 mL of isopropanol. The reaction was refluxed for 24 hr. The product was isolated by filtration. The product was taken up in water, and the solution was carefully adjusted to pH 7 with 1M NaOH (aq). The water was removed by lyophilization. The sodium salt was obtained (3.18 g, 60% yield). ¹H-NMR (400 MHz, d-DMSO, ppm): δ 0.97 (t, 3H, CH₃), 1.74 (m, 2H, CH₂), 2.60 (t, 2H, CH₂), 3.65 (m, 4H, CH₂), 7.60 (broad s, 5H, Ar-H). MS (ESI): 242.2 (M-1).

SQ5

To a flask was added squaric acid (64 mg, 0.6 mmol, 1 eq), **4.9** (303 mg, 1.2 mmol, 2.1 eq), 16 mL benzene, and 8 mL n-butanol. A Dean Stark trap was attached, and the reaction was refluxed at 100°C for 30 hr. The reaction was filtered, and the blue solid was washed with isopropanol. The crude product was purified on a C-18 RediSep column with a CombiFlash instrument eluting with water/acetonitrile. The product was a purple solid (5.51 mg, 1.6% yield). ¹H-NMR (400 MHz, d-DMSO, ppm): δ 1.13 (t, 6H, CH₃), 1.85 (m, 4H, CH₂), 2.45 (m, 4H, CH₂), 3.53 (m, 4H, CH₂), 3.60 (t, 4H CH₂), 6.98 (d, 4H, Ar-H),

8.06 (d, 4H, Ar-H). LRMS (ESI): 281.2 ((M-1)/2). HRMS (ESI): expected m/z 281.07, found m/z 281.07 ((M-1)/2); expected m/z 585.13, found m/z 585.14 (M-1).

4.10.2.6 SQ6

*SQ6*²⁵

To a flask was added squaric acid (114 mg, 1.0 mmol, 1 eq), 2-(N-ethyl-anilino)ethanol (0.3 mL, 1.9 mmol, 1.9 eq), 16 mL benzene, and 8 mL n-butanol. A Dean-Stark trap was attached, and the reaction was refluxed at 100°C for 18 hr. The product was filtered from the reaction and washed with DCM. The product was a blue solid (88.4 mg, 22% yield). ¹H-NMR (400 MHz, d-DMSO, ppm): δ 1.14 (t, 6H, CH₃), 3.27 (m, 4H, CH₂), 3.59 (m, 4H, CH₂), 4.98 (s, 1H, OH), 6.78 (d, 4H, Ar-H), 8.07 (d, 4H, Ar-H). MS (ESI): 409.2 (M+1).

4.10.2.7 SQ7

*Sodium 3-(ethyl(m-tolyl)amino)propane-1-sulfonate (4.10)*⁴⁸

To a flask was added N-ethyl-m-toluidine (2.1 mL, 14.8 mmol, 1 eq), 1,3-propanesultone (1.8 g, 14.8 mmol, 1 eq), and 20 mL of isopropanol. The reaction was refluxed for 18 hr. The product did not precipitate from the reaction, so the solvent was removed under reduced pressure. The residue was taken up in water and neutralized to pH 7 with 1M NaOH (aq). The aqueous solution became cloudy and was washed two times with ether. The aqueous solution was lyophilized. The product was a cream-colored solid (2.46 g, 65% yield). ¹H-NMR (400 MHz, d-DMSO, ppm): δ 1.06 (t, 3H, CH₃), 1.76 (m, 2H, CH₂), 2.21 (s, 3H, CH₃), 2.43 (m, 2H, CH₂), 3.30 (m, 4H, CH₂), 6.36 (d, 1H, Ar-H), 6.47 (m, 2H, Ar-H), 6.99 (t, 1H, Ar-H). MS (ESI): 256.2 (M-1).

SQ7

To a flask was added squaric acid (100 mg, 0.9 mmol, 1 eq), **4.10** (489 mg, 1.8 mmol, 2 eq), 16 mL benzene, and 8 mL n-butanol. A Dean-Stark trap was attached, and the reaction was refluxed at 100°C for 24 hr. The solvent was removed under reduced pressure. The crude product was purified on a C-18 RediSep column with a CombiFlash instrument eluting with water/acetonitrile. The product was a purple solid (6.8 mg, 1.2% yield). ¹H-NMR (400 MHz, d-DMSO, ppm): δ 1.15 (t, 6H, CH₃), 1.88 (m, 4H, CH₂), 2.45 (m, 4H, CH₂), 2.76 (s, 6H, CH₃), 3.53 (m, 4H, CH₂), 3.58 (t, 4H CH₂), 6.77 (d, 2H, Ar-H), 7.09 (m, 2H, Ar-H), 8.06 (d, 2H, Ar-H). LRMS (ESI): 295.1 ((M-1)/2), 591.1 (M-1). HRMS (ESI): expected m/z 295.09, found m/z 295.09 ((M-1)/2); expected m/z 613.17, found m/z 613.17 ((M-1)+Na).

4.10.2.8 SQ8

*N-ethyl-3-methoxyaniline (4.6)*⁴⁹

To a flask was added m-anisidine (4.2 mL, 37.4 mmol, 1 eq), iodoethane (2.3 mL, 29.9 mmol, 0.8 eq), potassium carbonate (4.0 g, 29.9 mmol, 0.8 eq), and 7 mL DMF. The reaction was heated to 60°C for 24 hr. Water was added to the reaction mixture, and the solution was extracted two times with ethyl acetate. The organic layer was washed with water twice and with brine and then dried with MgSO₄. The solvent was removed under reduced pressure. The crude product was purified on a basic alumina with 10-20% ethyl acetate in hexanes. The product was a brown oil (1.74 g, 31% yield). ¹H-NMR (400 MHz, CDCl₃, ppm): δ 1.24 (t, 3H, CH₃), 3.15 (m, 2H, CH₂), 3.76 (s, 3H, CH₃), 6.18 (s, 1H, Ar-H), 6.25 (m, 2H, Ar-H), 7.09 (t, 1H, Ar-H).

*Sodium 3-(ethyl(3-methoxyphenyl)amino)propane-1-sulfonate (4.11)*⁴⁸

To a flask was added **4.6** (1.74 g, 11.5 mmol, 1 eq), 1,3-propanesultone (1.5 mL, 17.1 mmol, 1.5 eq), and 10 mL isopropanol. The reaction was refluxed for 3 hr. The

reaction mixture was concentrated and neutralized with 1M NaOH (aq). The solvent was removed under reduced pressure and lyophilized. The crude product was recrystallized from acetone. The product was a white solid (1.86 g, 55% yield). ¹H-NMR (400 MHz, d-DMSO, ppm): δ 1.01 (t, 3H, CH₃), 1.72 (m, 2H, CH₂), 2.41 (m, 2H, CH₂), 3.27 (m, 4H, CH₂), 3.65 (s, 3H, CH₃), 6.10 (d, 1H, Ar-H), 6.15 (s, 1H, Ar-H), 6.22 (d, 2H, Ar-H), 6.97 (t, 1H, Ar-H). MS (ESI): 272.2 (M-1).

SQ8

To a flask was added squaric acid (70 mg, 0.61 mmol, 1 eq), **4.11** (381 mg, 1.29 mmol, 2.1 eq), 10 mL benzene, and 5 mL n-butanol. A Dean-Stark trap was attached, and the reaction was refluxed at 100°C for 48 hr. The reaction mixture was filtered to collect the product, washing with ethyl acetate and allowing to dry thoroughly. The crude product was purified on a C-18 RediSep column with a CombiFlash instrument eluting with water/acetonitrile. The product was a purple solid (26.5 mg, 6% yield). ¹H-NMR (400 MHz, d-DMSO, ppm): δ 1.12 (t, 6H, CH₃), 1.82 (m, 4H, CH₂), 2.45 (m, 4H, CH₂), 3.50 (m, 4H, CH₂), 3.65 (t, 4H CH₂), 3.83 (s, 6H, CH₃), 6.35 (s, 2H, Ar-H), 6.51 (d, 2H, Ar-H), 8.53 (d, 2H, Ar-H). LRMS (ESI): 311.2 ((M-1)/2), 623.2 (M-1). HRMS (ESI): expected m/z 311.08, found m/z 311.08 ((M-1)/2); expected m/z 645.16, found m/z 645.15 ((M-1)+Na).

4.10.2.9 *SQ9*

N-(3-(ethylamino)phenyl)acetamide (**4.7**)⁵⁰

To a flask was added 3-aminoacetanilide (4 g, 26.6 mmol, 1 eq), iodoethane (1.65 mL, 21.3 mmol, 0.8 eq), potassium carbonate (3.59 g, 26.6 mmol, 1 eq), and 7 mL DMF. The reaction was heated to 60°C for 20 hr. The reaction was filtered to remove solids. Water was added to the solution, and it was extracted three times with ethyl acetate. The

organic layer was washed brine twice and then dried with MgSO₄. The solvent was removed under reduced pressure. The crude product was purified on basic alumina with 20-100% ethyl acetate in hexanes. The product was a brown oil (1.42 g, 30% yield). ¹H-NMR (400 MHz, CDCl₃, ppm): δ 1.23 (t, 3H, CH₃), 2.13 (s, 3H, CH₃), 3.13 (m, 2H, CH₂), 6.44 (s, 1H, Ar-H), 6.73 (m, 1H, Ar-H), 7.08 (t, 2H, Ar-H), 7.35 (s, 1H, NH). MS (ESI): 179.2 (M+1).

Sodium 3-((3-acetamidophenyl)(ethyl)amino)propane-1-sulfonate (4.12)

To a flask was added **4.7** (0.5 g, 2.8 mmol, 1 eq), 1,3-propanesultone (0.3 mL, 3.4 mmol, 1.2 eq), and 10 mL acetonitrile. The reaction was refluxed for 18 hr. The reaction mixture was concentrated. Water was added to the residue, and the solution was neutralized with 1M NaOH (aq). The solution was washed with DCM and then lyophilized. The product was a white solid (620 mg, 68% yield). ¹H-NMR (400 MHz, d-DMSO, ppm): δ 1.02 (t, 3H, CH₃), 1.74 (m, 2H, CH₂), 1.96 (s, 3H, CH₃), 2.35 (m, 2H, CH₂), 3.23 (m, 4H, CH₂), 6.33 (d, 1H, Ar-H), 6.84 (m, 2H, Ar-H), 6.97 (t, 1H, Ar-H), 9.65 (s, 1H, NH). ¹³C-NMR (400 MHz, d-DMSO, ppm): δ 11.09 (CH₃), 22.55 (CH₃), 22.68 (CH₂), 27.82 (CH₂), 48.67 (CH₂), 60.41 (CH₂), 104.06 (Ar-C), 107.75 (Ar-C), 108.57 (Ar-C), 129.07 (Ar-C), 139.37 (Ar-C), 148.02 (Ar-C), 170.18 (C=O). LRMS (ESI): 299.2 (M-1). HRMS (ESI): expected m/z 299.11, found m/z 299.11 (M-1)

SQ9

To a flask was added squaric acid (70 mg, 0.61 mmol, 1 eq), **4.12** (416 mg, 1.3 mmol, 2.1 eq), 5 mL benzene, and 5 mL n-butanol. A Dean-Stark trap was attached, and the reaction was refluxed at 100°C for 48 hr. The reaction mixture was filtered to collect the product, washing with ethyl acetate and isopropanol and allowing to dry thoroughly. The product was a green solid and required no further purification (212 mg, 49% yield). ¹H-NMR (400 MHz, d-DMSO, ppm): δ 1.24 (t, 6H, CH₃), 1.86 (m, 4H, CH₂), 2.17 (s, 3H,

CH₃), 2.39 (m, 4H, CH₂), 3.36 (m, 4H, CH₂), 3.55 (m, 4H CH₂), 6.76 (d, 2H, Ar-H), 8.06 (s, 2H, Ar-H), 8.28 (d, 2H, Ar-H) 11.94 (s, 1H, NH). LRMS (ESI): 338.2 ((M-1)/2). HRMS (ESI): expected m/z 338.09, found m/z 338.09 ((M-1)/2); expected m/z 699.18, found m/z 699.18 ((M-1)+Na).

4.10.2.10 c-BSA

BSA (332 mg, 5 μ mol, 1 eq) was dissolved in 15 mL of 50 mM phosphate buffer, 10 mM EDTA, H₂O, pH 7.5. Iodoacetamide (9.24 mg, 50 μ mol, 10 eq) was added, and the reaction was stirred for 5 hr. The reaction mixture was filtered to remove any insoluble components and purified against 10 mM phosphate buffer, H₂O, pH 7.0, 0.02% NaN₃ with 10 kDa MWCO centrifugal filter units. The concentration of the protein in the solution was measured by the Bradford method (**Section 5.9.6.1**). The thiol content was measured by Ellman's test (**Section 5.9.6.2**) and was found to be zero.

4.10.3 Titrations

The procedure for each titration is described in the caption of the figure that displays the results.

4.11 ADDITIONAL FIGURES AND EXPERIMENTS

4.11.1 ^1H NMR Monitoring of Thiol Addition

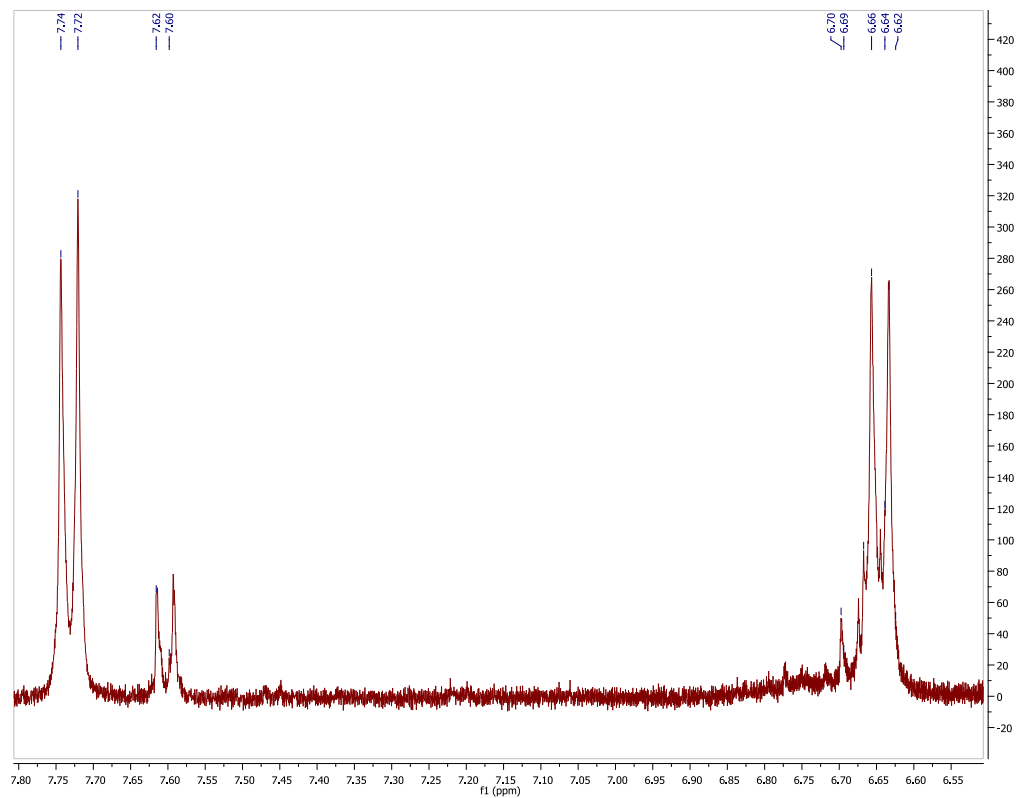


Figure 4.54 ^1H NMR of *SQ5* in D_2O . *SQ5* $800\ \mu\text{M}$.

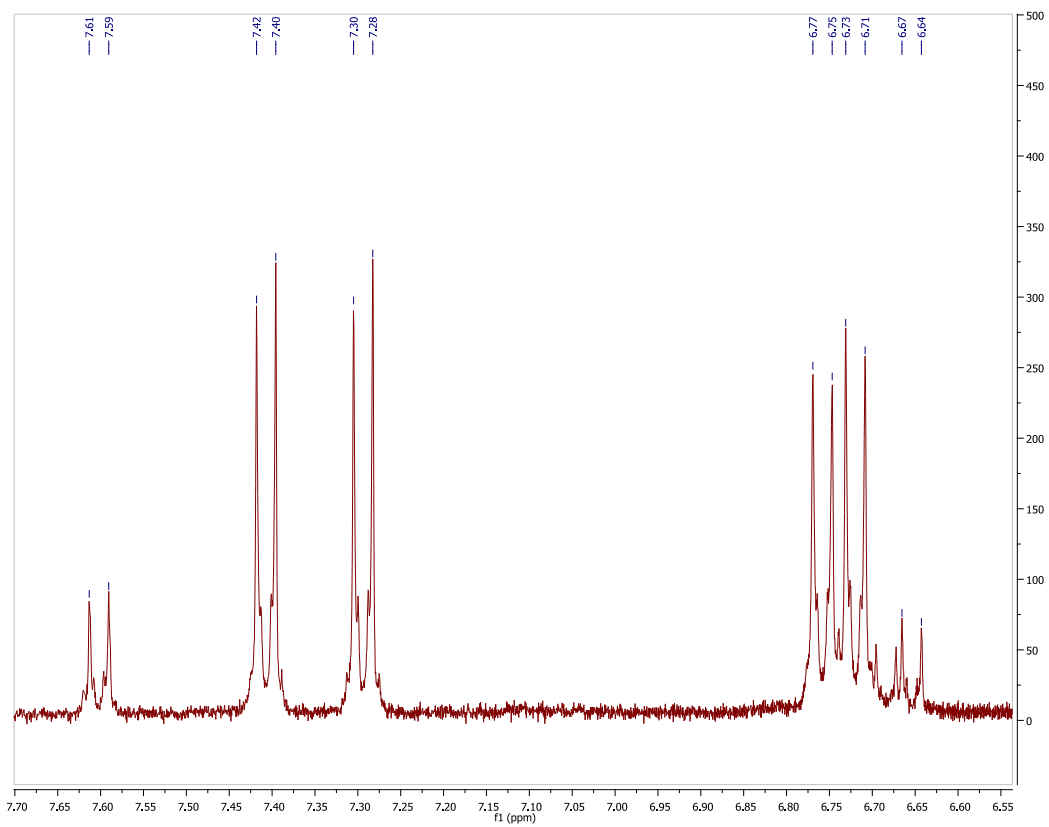


Figure 4.55 ^1H NMR of **SQ5** and thiol in D_2O taken immediately after preparation. **SQ5** 800 μM , NAC 800 μM .

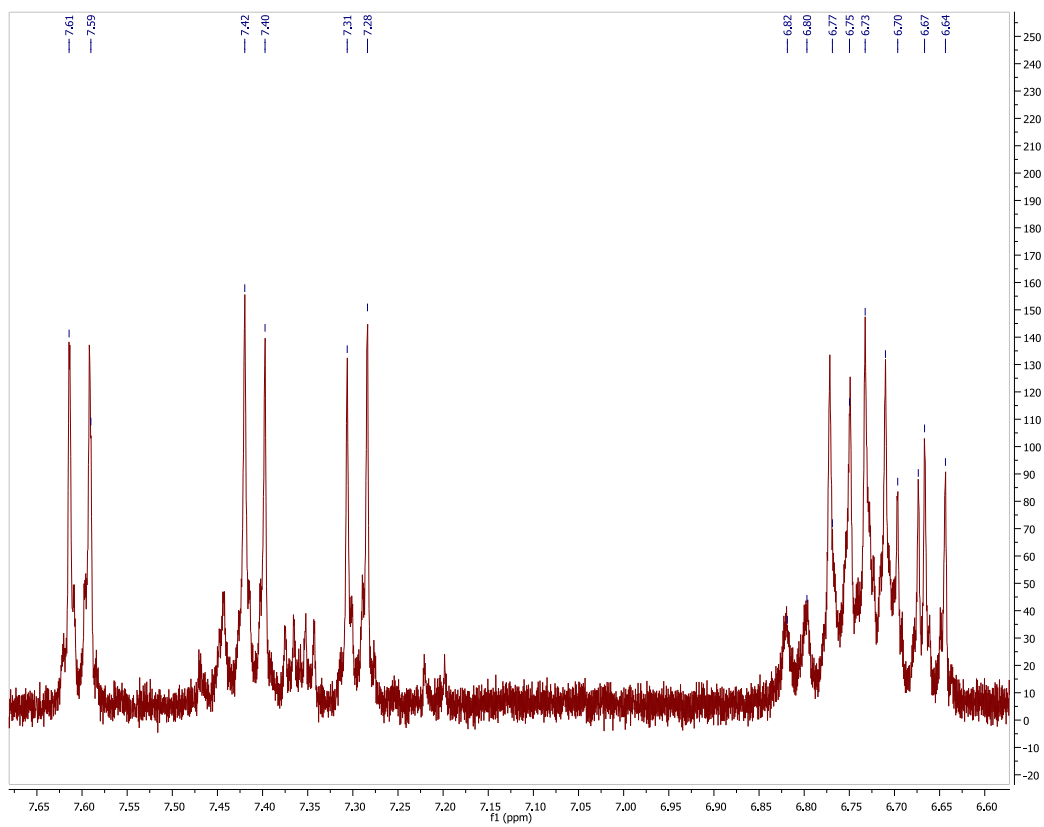


Figure 4.56 ^1H NMR of **SQ5** and thiol in D_2O .taken 24 hr after the sample was prepared. **SQ5** 800 μM , NAC 800 μM

4.11.2 Titrations of Thiol into BSA/SQ

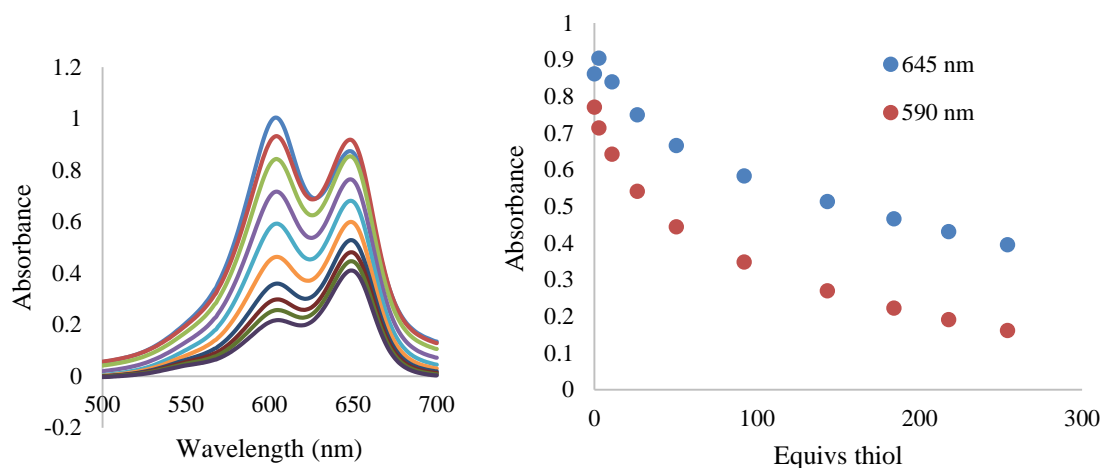


Figure 4.57 *Titration of thiol into SQ4/BSA.* A solution of BSA (10 μM) and SQ4 (10 μM) was titrated with NAC (0-254 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 .

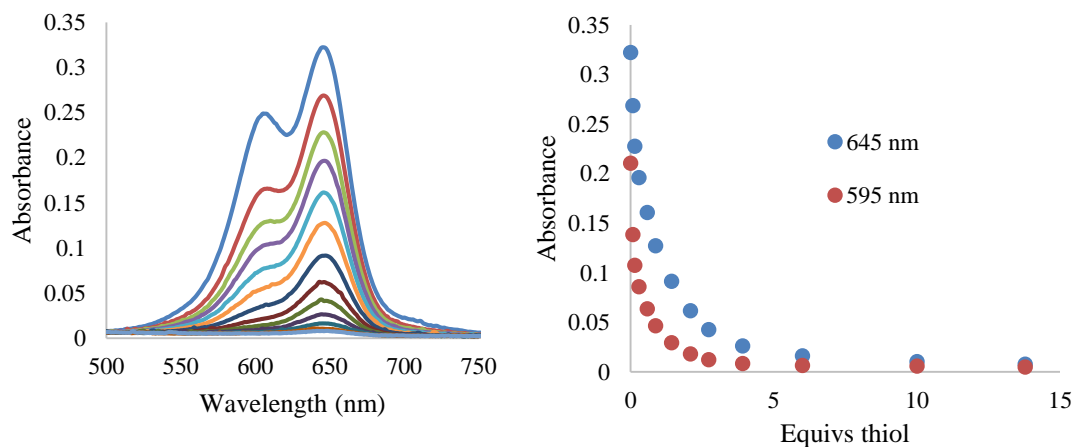


Figure 4.58 *Titration of thiol into SQ5/BSA.* A solution of BSA (10 μM) and SQ5 (10 μM) was titrated with NAC (0-14 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 .

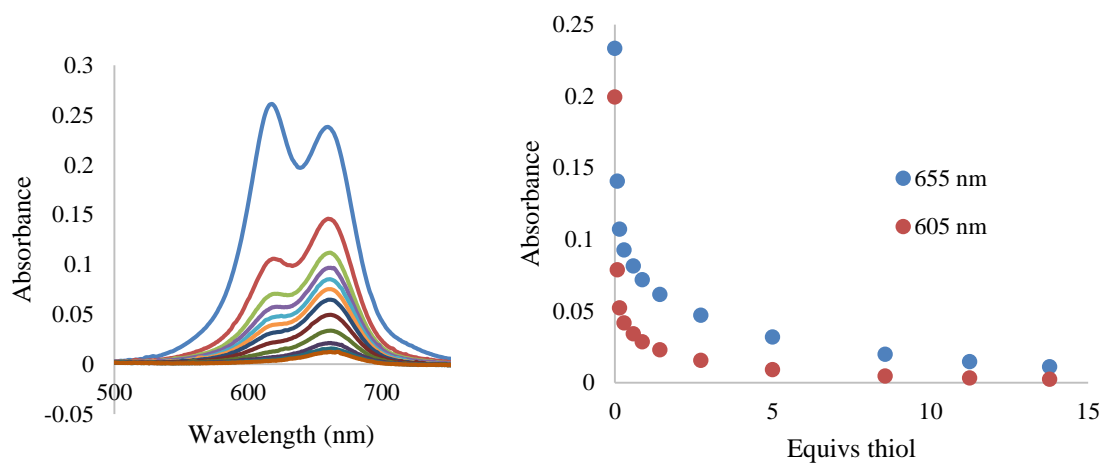


Figure 4.59 *Titration of thiol into SQ7/BSA.* A solution of BSA (10 μM) and SQ7 (10 μM) was titrated with NAC (0-14 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 .

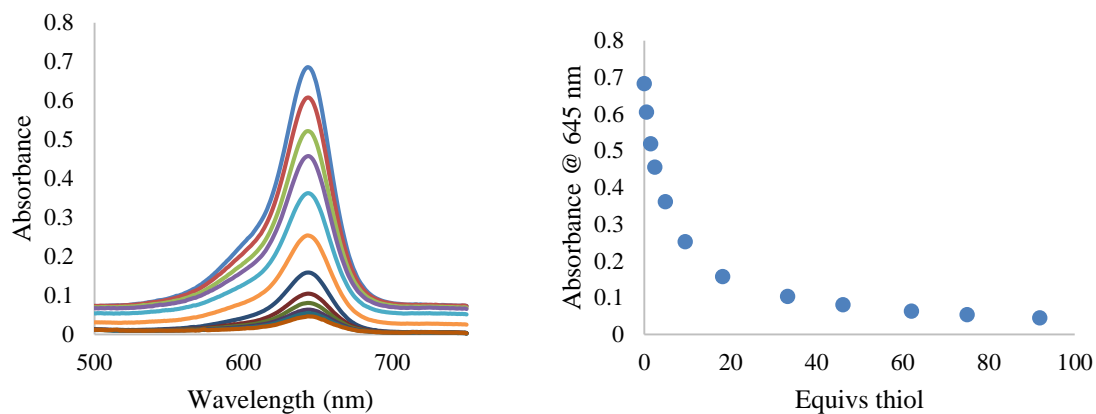


Figure 4.60 *Titration of thiol into SQ8/BSA.* A solution of BSA (10 μM) and SQ8 (10 μM) was titrated with NAC (0-92 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 .

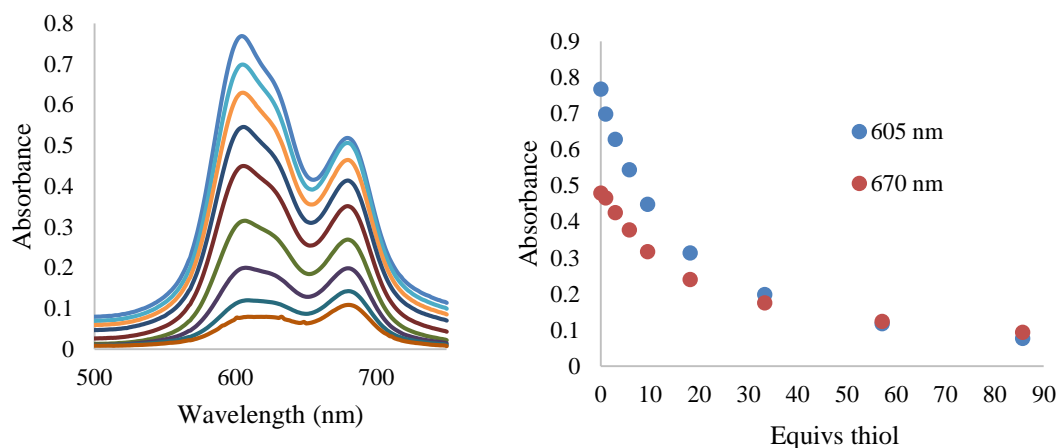


Figure 4.61 *Titration of thiol into SQ9/BSA.* A solution of BSA (10 μM) and SQ9 (10 μM) was titrated with NAC (0-86 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 .

4.11.3 Titrations of G5 PAMAM Dendrimer into SQ5

Amine-terminated PAMAM dendrimers are known to bind numerous anionic dyes, so we tested such a dendrimer (G5) for binding to SQ5 as a potential alternative scaffold to serum albumin.^{51,52} Ultimately, we were concerned, as we were with serum albumin, that the squaraine could be reacting with amines in the dendrimer.

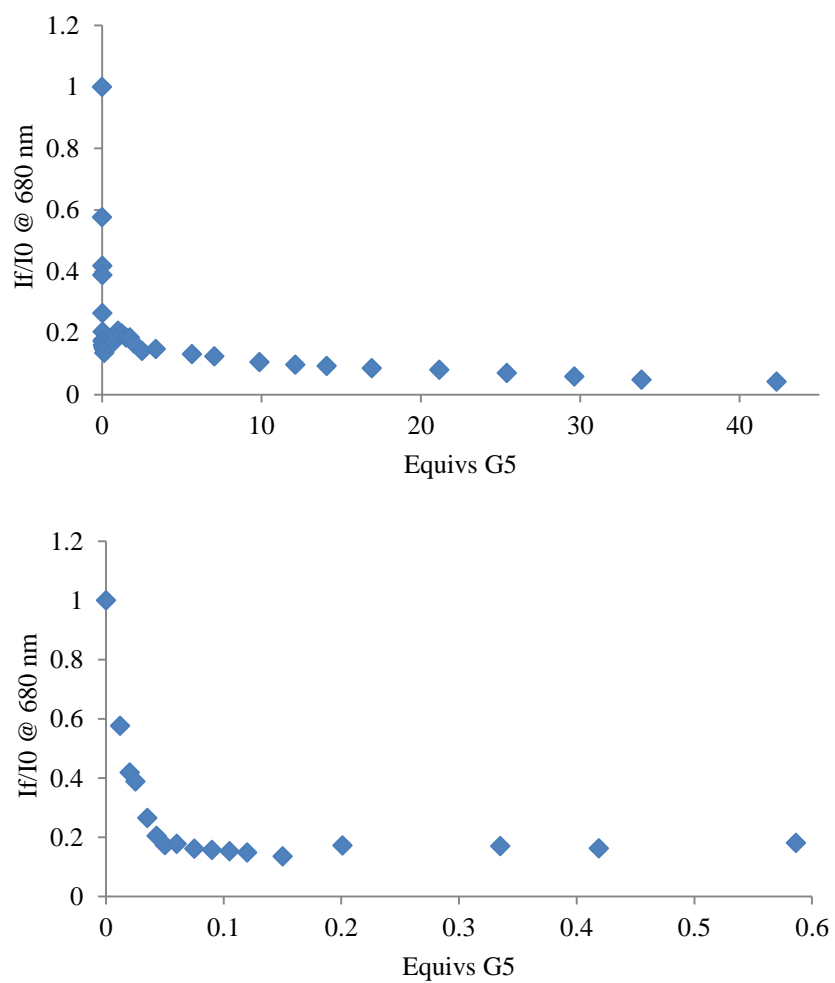


Figure 4.62 Fluorescence titration of G5 dendrimer into **SQ5**. Addition of G5 (0-85 μM) to **SQ5** (2 μM) in 50 mM HEPES buffer, H_2O , pH 7.4, $\lambda_{\text{ex}} = 640 \text{ nm}$.

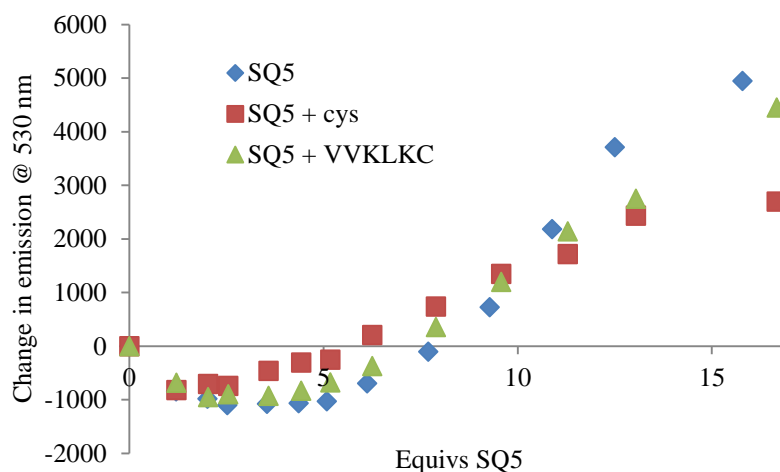


Figure 4.63 Displacement of carboxyfluorescein from G5 dendrimer by **SQ5**, **SQ5/NAC**, and **SQ/VVKLKC**. To carboxyfluorescein (2 μM) and G5 dendrimer (0.67 μM) was added **SQ5** (0–32 μM), **SQ5/NAC** (0–33 μM), and **SQ5/VVKLKC** (0–33 μM) in 50 mM HEPES buffer, H_2O , pH 7.4, $\lambda_{\text{ex}} = 485 \text{ nm}$.

4.12 REFERENCES

- (1) Treibs, A.; Jacob, K. *Angew. Chemie Int. Ed. English* **1965**, *4*, 694–694.
- (2) Law, K. Y. *Chem. Rev.* **1993**, *93*, 449–486.
- (3) Law, K.-Y.; Bailey, F. C. *Dye. Pigment.* **1988**, *9*, 85–107.
- (4) Cole, E. L.; Arunkumar, E.; Xiao, S.; Smith, B. A.; Smith, B. D. *Org. Biomol. Chem.* **2012**, *10*, 5769–5773.
- (5) Silvestri, F.; Irwin, M. D.; Beverina, L.; Facchetti, A.; Pagani, G. A.; Marks, T. J. *J. Am. Chem. Soc.* **2008**, *130*, 17640–17641.
- (6) Xiao, X.; Wei, G.; Wang, S.; Zimmerman, J. D.; Renshaw, C. K.; Thompson, M. E.; Forrest, S. R. *Adv. Mater.* **2012**, *24*, 1956–1960.
- (7) Wei, G.; Wang, S.; Renshaw, K.; Thompson, M. E.; Forrest, S. R. *ACS Nano* **2010**, *4*, 1927–1934.
- (8) Yagi, S.; Hyodo, Y.; Hirose, M.; Nakazumi, H.; Sakurai, Y.; Ajayaghosh, A. *Org. Lett.* **2007**, *9*, 1999–2002.

- (9) Arunkumar, E.; Chithra, P.; Ajayaghosh, A. *J. Am. Chem. Soc.* **2004**, *126*, 6590–6598.
- (10) Collins, C. G.; Peck, E. M.; Kramer, P. J.; Smith, B. D. *Chem. Sci.* **2013**, *4*, 2557.
- (11) Basheer, M. C.; Alex, S.; George Thomas, K.; Suresh, C. H.; Das, S. *Tetrahedron* **2006**, *62*, 605–610.
- (12) Ajayaghosh, A. *Acc. Chem. Res.* **2005**, *38*, 449–459.
- (13) Ramaiah, D.; Eckert, I.; Arun, K. T.; Weidenfeller, L.; Epe, B. *Photochem. Photobiol.* **2004**, *79*, 99–104.
- (14) Dirk, C. W.; Herndon, W. C.; Cervantes-Lee, F.; Selnau, H.; Martinez, S.; Kalamegham, P.; Tan, A.; Campos, G.; Velez, M. *J. Am. Chem. Soc.* **1995**, *117*, 2214–2225.
- (15) Bigelow, R. W.; Freund, H.-J. *Chem. Phys.* **1986**, *107*, 159–174.
- (16) Sreejith, S.; Carol, P.; Chithra, P.; Ajayaghosh, A. *J. Mater. Chem.* **2008**, *18*, 264.
- (17) Das, S.; Kamat, P. V.; De la Barre, B.; Thomas, K. G.; Ajayaghosh, A.; George, M. V. *J. Phys. Chem.* **1992**, *96*, 10327–10330.
- (18) Das, S.; Thanulingam, T. L.; Thomas, K. G.; Kamat, P. V.; George, M. V. *J. Phys. Chem.* **1993**, *97*, 13620–13624.
- (19) Chen, H.; Farahat, M. S.; Law, K.-Y.; Whitten, D. G. *J. Am. Chem. Soc.* **1996**, *118*, 2584–2594.
- (20) Sauer, M.; Hofkens, J.; Enderlein, J. In *Handbook of Fluorescence Spectroscopy and Imaging*; Wiley-VCH: Weinheim, Germany, 2011; pp. 1–30.
- (21) *J-Aggregates, Volume 2*; Kobayashi, T., Ed.; World Scientific Publishing Co. Pte. Ltd.: Singapore, 2012.
- (22) Ros-Lis, J. V.; Martínez-Máñez, R.; Soto, J. *Chem. Commun.* **2002**, 2248–2249.
- (23) Ros-Lis, J. V.; García, B.; Jiménez, D.; Martínez-Máñez, R.; Sancenón, F.; Soto, J.; Gonzalvo, F.; Valldecabres, M. C. *J. Am. Chem. Soc.* **2004**, *126*, 4064–4065.
- (24) Ros-Lis, J. V.; Martínez-Máñez, R.; Soto, J.; Villaescusa, L. A.; Rurack, K. *J. Mater. Chem.* **2011**, *21*, 5004.

- (25) Yan, Z.; Guang, S.; Xu, H.; Liu, X. *Analyst* **2011**, *136*, 1916–1921.
- (26) Fan, J.; Wang, Z.; Zhu, H.; Fu, N. *Sensors Actuators B Chem.* **2013**, *188*, 886–893.
- (27) Hewage, H. S.; Anslyn, E. V. *J. Am. Chem. Soc.* **2009**, *131*, 13099–13106.
- (28) Shafeekh, K. M.; Rahim, M. K. A.; Basheer, M. C.; Suresh, C. H.; Das, S. *Dye. Pigment.* **2013**, *96*, 714–721.
- (29) Luo, C.; Zhou, Q.; Zhang, B.; Wang, X. *New J. Chem.* **2011**, *35*, 45.
- (30) Fan, J.; Chen, C.; Lin, Q.; Fu, N. *Sensors Actuators B Chem.* **2012**, *173*, 874–881.
- (31) Liao, S.; Han, W.; Ding, H.; Xie, D.; Tan, H.; Yang, S.; Wu, Z.; Shen, G.; Yu, R. *Anal. Chem.* **2013**, *85*, 4968–4973.
- (32) Arun, K. T.; Jayaram, D. T.; Avirah, R. R.; Ramaiah, D. *J. Phys. Chem. B* **2011**, *115*, 7122–7128.
- (33) Liu, X.-D.; Sun, R.; Ge, J.-F.; Xu, Y.-J.; Xu, Y.; Lu, J.-M. *Org. Biomol. Chem.* **2013**, *11*, 4258–4264.
- (34) Terpetsching, E.; Szmecinski, H.; Lakowicz, J. R. *Anal. Chim. Acta* **1993**, *282*, 633–641.
- (35) Volkova, K. D.; Kovalska, V. B.; Tatarets, A. L.; Patsenker, L. D.; Kryvorotenko, D. V.; Yarmoluk, S. M. *Dye. Pigment.* **2007**, *72*, 285–292.
- (36) Volkova, K. D.; Kovalska, V. B.; Losytskyy, M. Y.; Bento, A.; Reis, L. V.; Santos, P. F.; Almeida, P.; Yarmoluk, S. M. *J. Fluoresc.* **2008**, *18*, 877–882.
- (37) Jisha, V. S.; Arun, K. T.; Hariharan, M.; Ramaiah, D. *J. Am. Chem. Soc.* **2006**, *128*, 6024–6025.
- (38) Jisha, V. S.; Arun, K. T.; Hariharan, M.; Ramaiah, D. *J. Phys. Chem. B* **2010**, *114*, 5912–5919.
- (39) Zhang, Y.; Yue, X.; Kim, B.; Yao, S.; Bondar, M. V.; Belfield, K. D. *ACS Appl. Mater. Interfaces* **2013**, *5*, 8710–8717.
- (40) Isgor, Y. G.; Akkaya, E. U. *Tetrahedron Lett.* **1997**, *38*, 7417–7420.

- (41) Ghazarossian, V.; Pease, J. S.; Hu, M. W.; Laney, M.; Tarnowski, T. L. *Fluorescent Dyes*. EP89308412, 1989.
- (42) West, R.; Powell, D. L. *J. Am. Chem. Soc.* **1963**, *85*, 2577–2579.
- (43) Law, K.-Y.; Bailey, F. C.; Bluett, L. J. *Can. J. Chem.* **1986**, *64*, 1607–1619.
- (44) Ritchie, C. D.; Sager, W. F. *Progress in Physical Organic Chemistry*; Cohen, S. G.; Streitwieser, A.; Taft, R. W., Eds.; Progress in Physical Organic Chemistry; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1964; Vol. 2.
- (45) Peters, T. *All About Albumin: Biochemistry, Genetics, and Medical Applications*; Academic Press: San Diego, CA, 1996.
- (46) Ramaiah, D.; Joy, A.; Chandrasekhar, N.; Eldho, N. V.; Das, S.; George, M. V. *Photochem. Photobiol.* **1997**, *65*, 783–790.
- (47) Jose, J.; Ueno, Y.; Burgess, K. *Chemistry* **2009**, *15*, 418–423.
- (48) Tamaoku, K.; Murao, Y.; Akiura, K.; Ohkura, Y. *Anal. Chim. Acta* **1982**, *136*, 121–127.
- (49) Sun, Q.; Qian, J.; Tian, H.; Duan, L.; Zhang, W. *Chem. Commun. (Camb)*. **2014**, *50*, 8518–8521.
- (50) Sadao, A. Production of 3-(N-alkylamino)-acylanilide. JPH08143523, 1996.
- (51) Bonizzoni, M.; Long, S. R.; Rainwater, C.; Anslyn, E. V. *J. Org. Chem.* **2012**, *77*, 1258–1266.
- (52) Rainwater, J. C.; Anslyn, E. V. *Chem. Commun. (Camb)*. **2010**, *46*, 2904–2906.

Chapter 5: Development of Dynamic Receptors from Serum Albumin and Thia-Michael Addition Chemistry

5.1 INTRODUCTION

As another strategy for building dynamic receptors, we sought to construct large, multivalent receptors from serum albumin by reversible covalent attachment of recognition units. Rather than relying on non-covalent association of the units with serum albumin, as was the case with the fatty acid-peptides (**Chapter 3**) and squaraines (**Chapter 4**), the units in this design would be covalently anchored to the protein. Thus, the bulk of our work went toward developing a suitable protein-labeling protocol that incorporated a reversible covalent linkage.

5.1.1 Dynamic Combinatorial Chemistry

Dynamic combinatorial chemistry (DCC) uses reversible covalent bonding to generate libraries of compounds under thermodynamic control.¹⁻³ In the presence of a guest molecule, the distribution of the species in the dynamic combinatorial library (DCL) will change as a function of the affinity of each species for the guest, according to Le Chatelier's principle. In this way, the species with the highest affinity for the guest will be selected for, while poor binders will be sacrificed (**Figure 5.1**). The best receptors for the guest are not only identified but also amplified in this process. This technique has been used to identify receptors for a variety of molecules, most notably biologically significant ones.⁴⁻¹⁰

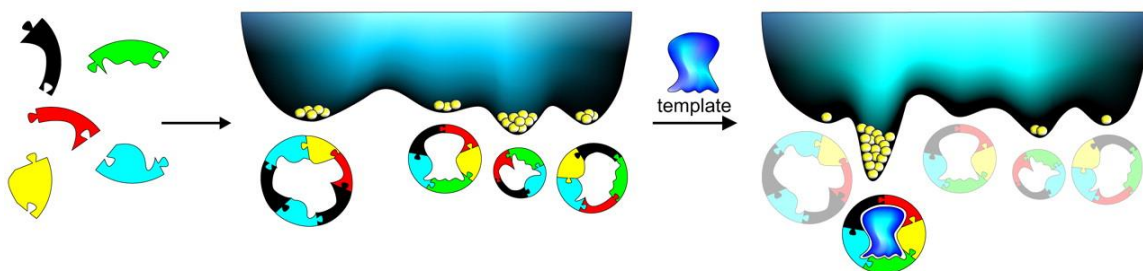


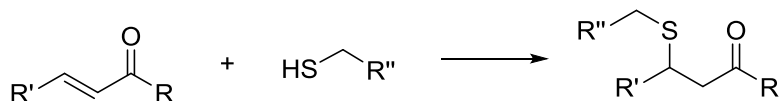
Figure 5.1 *Illustration of a small dynamic combinatorial library (DCL) and the effect of a template on the free energy of library species.* (Reprinted from Ref-Otto Science paper: © Association for the Advancement of Science, 2002).

5.1.1.1 Reversible Covalent Bonding

Chemical reactions that have been used for building DCLs for biological targets include C=N, S-S, and S-C bond-forming reactions.^{3,11} Chemical reactions that are suitable for DCC must be readily reversible and thus be under thermodynamic control. Many organic reactions require heating to high temperatures and extended reaction times to achieve thermodynamic control.¹² DCLs for biological targets such as proteins must be able to function under physiological or near physiological conditions (e.g. neutral pH, ambient pressure, mild temperatures). Other biologically-relevant targets such as peptides or natural products could theoretically survive a more extreme environment. However, the context of receptor identification is important for future utility of the receptor or receptors identified. For instance, a good receptor for a target compound at pH 12 might not be such a good receptor for that compound at pH 7.4, depending on the binding interactions at work. Thus, there are relatively few reactions suitable for constructing DCLs for biologically-relevant targets.

In this work, we sought to develop the use of thia-Michael reactions as the basis for DCLs (**Scheme 5.1**). Conjugate addition reactions are relative newcomers to the DCC tool box,¹¹ and studies are still ongoing with respect to developing specific conjugate acceptor

substrates and reaction conditions that meet the criteria described above.^{13,14} Therefore, we chose the thia-Michael reaction for our work because this type of reversible covalent reaction has not been as thoroughly studied or widely used as other reactions such as disulfide exchange. By incorporating thia-Michael chemistry into our protein receptors, we aimed to expand our basic understanding of these reactions for use in DCLs.



Scheme 5.1 *Thia-Michael reaction.*

5.1.2 Protein Modification and Conjugation

Proteins are common targets for modification and conjugation for a variety of purposes.¹⁵ Proteins can be modified to introduce a specific functionality (e.g. thiol, amine, carboxylate), to block a specific native functionality, or even to link the protein to another species (e.g. a fluorescent tag, another protein). The amino acids within proteins that are typically used for modification and conjugation purposes are the ones with ionizable side chains: aspartic acid, glutamic acid, lysine, arginine, cysteine, histidine, and tyrosine. When deprotonated, these side chains can be sufficiently nucleophilic to add to electrophiles. A variety of electrophilic reagents are commercially available for reacting with these side chains. Some considerations in protein modification include the selectivity of the reaction, the reaction conditions necessary (e.g. temperature, pH, solvent), the solvent accessibility of the targeted side chains, and the effect of the modification on the secondary, tertiary, and quaternary structure of the protein.

For protein conjugates, the characterization can be case-specific. For instance, if a protein is conjugated to a fluorophore, the emission of that moiety is measured to detect the modification. However, if the conjugated species does not provide a simple and selective detection method, then more general procedures are applied. In order to measure the approximate molecular weight of the modified protein, SDS-PAGE or gel-filtration chromatography with molecular weight markers are used.²¹

Tandem mass spectrometry has found extensive use as a bottom-up proteomics tool for detecting and characterizing modified amino acids in proteins.^{22–25} In bottom-up protein analysis, the protein is first digested by a protease into peptides (**Figure 5.2**). The peptides are separated by liquid chromatography (LC) and then ionized, usually by electrospray (ESI). Then the peptides are subjected to tandem mass spectrometry with one or more dissociation steps (MS^X , where $X \geq 2$ since one mass spectrometer is always necessary for the ion analysis at the end of the fragmentation step/s). Each peptide is fragmented along the amide backbone. Collision-induced dissociation by a neutral gas is the most common fragmentation method (**Figure 5.3**).^{26,27} In this process, the amide backbone is broken to yield b- and y-type cations (**Figure 5.4**). These fragmentation patterns are used to identify modified amino acids by comparison to the peptide fragments expected from the native protein. In this way, both the type and location of the conjugated groups are determined. It is also possible to obtain some quantitative or semi-quantitative information about the number of modifications of each type at each location. This methodology is a powerful tool for obtaining detailed information about protein conjugates.

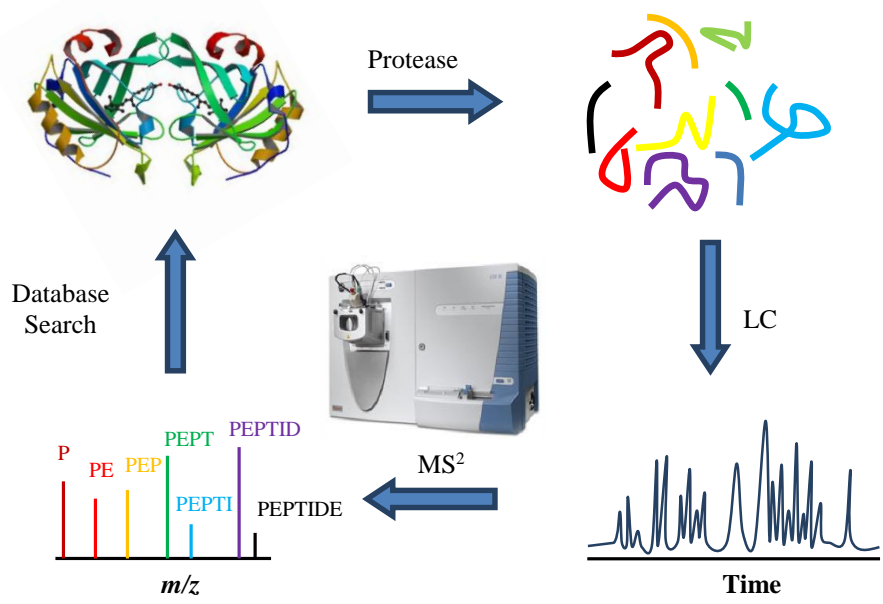


Figure 5.2 *Bottom-up mass spectrometric analysis of proteins.* (Image by Drs. Jared Shaw and Jennifer Brodbelt)

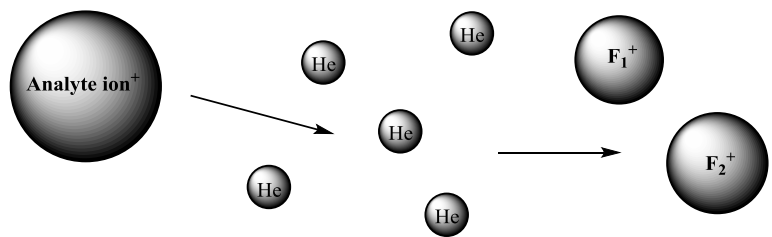


Figure 5.3 *Collision-induced dissociation by helium gas.* (Image by Drs. Jared Shaw and Jennifer Brodbelt)

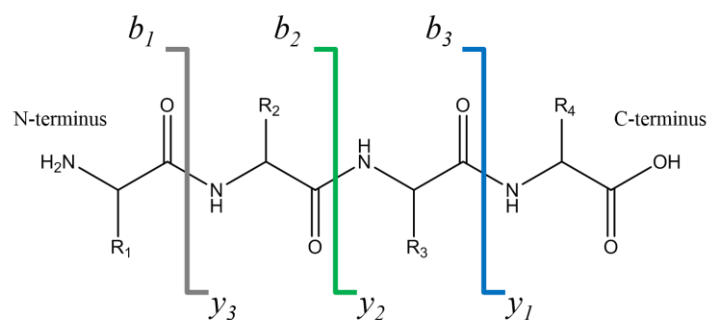
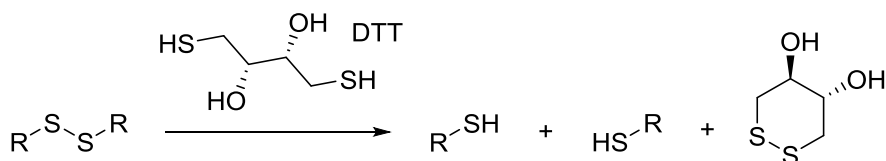


Figure 5.4 Collision-induced dissociation yields *b*- and *y*-ions. (Image by Drs. Jared Shaw and Jennifer Brodbelt)

5.2 PARTIAL REDUCTION OF SERUM ALBUMIN

Bovine serum albumin contains 35 cysteine residues in 17 disulfide bonds and one additional free cysteine.²⁸ As a way of generating more thiols on serum albumin, some of these disulfide bonds were reduced. Following the procedure of Grinnell *et al.*,²¹ dithiothreitol (DTT) was used as a reducing agent (**Scheme 5.4**). The buffer was deoxygenated by sparging with nitrogen to prevent reoxidation of the thiols. A sample of BSA was exposed to DTT under nitrogen, and then any excess DTT was removed by dialysis.



Scheme 5.4 Reduction of disulfides by dithiothreitol (DTT).

The thiol was quantified by Ellman's test.¹⁷ An average of 20 thiols per BSA molecule were obtained from the reduction. The reduced BSA (r-BSA) was prone to reoxidation in air, readily precipitating or forming a gel due to disulfide crosslinking between proteins.²⁹ Nevertheless, r-BSA was reacted with alkylating agents, including iodoacetamide and maleimide, to cap the thiols. Eventually we wanted to use conjugate

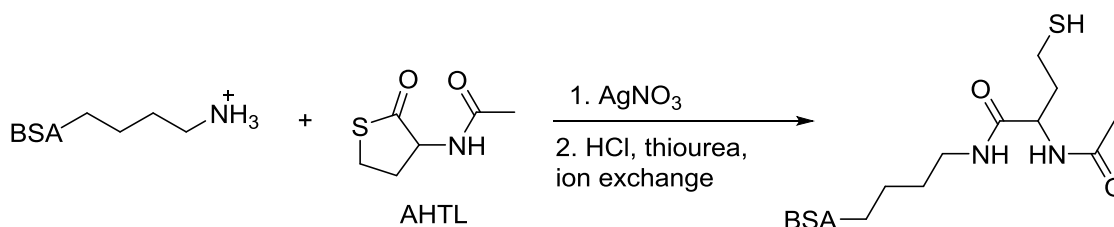
acceptor-appended recognition units, but at this point we just needed to cap the thiols with something to obtain a stable sample. However, the alkylated r-BSA became insoluble in water. Such problems with reduced serum albumin samples have been reported,^{21,29,30} and we decided to abandon this approach.

5.3 THIOLATION OF SERUM ALBUMIN

As an alternative strategy to reduction of serum albumin, we next explored thiolation of the protein. The ϵ -amine of lysine residues can be thiolated using *N*-acetylhomocysteine thiolactone (AHTL) or *N*-succinimidyl *S*-acetylthioacetate (SATA).¹⁵ We investigated the use of both thiolating agents, and also synthesized our own derivative of SATA with an *S*-trityl rather than an *S*-acetyl group (STTA).

5.3.1 *N*-acetylhomocysteine thiolactone (AHTL)

AHTL is activated by silver ion coordination.^{31,32} The amine attacks the activated carbonyl of the thiolactone, opening the ring to yield a free thiolate (**Scheme 5.5**). The silver then coordinates to this free thiolate and must be removed by acidification and ion exchange chromatography to yield the thiol.

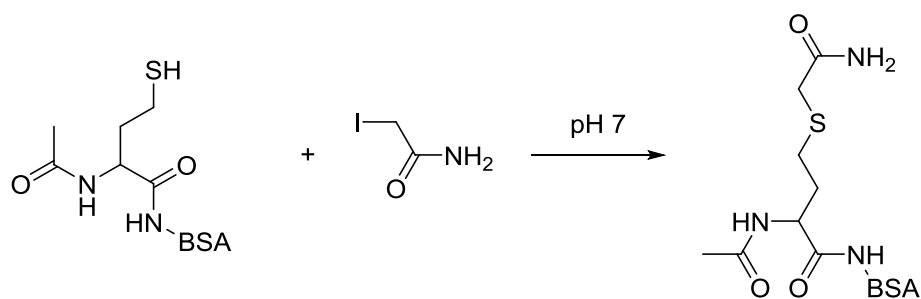


Scheme 5.5 *Thiolation of BSA by AHTL and silver nitrate.*

5.3.1.1 Capping with Iodoacetamide

A sample of BSA was thiolated with AHTL and silver nitrate (t-BSA), and the thiol content was measured by Ellman's test (18 thiols per BSA). The pH of the sample was

adjusted back up to 7, and iodoacetamide was added to alkylate the t-BSA (c-t-BSA) (**Scheme 5.6**). Alkylation was confirmed by a negative Ellman's test. With the c-t-BSA sample in hand, we wanted to determine if the protein had been significantly denatured by the chemical modification. Thus, we titrated several fluorescent indicators with native BSA and c-t-BSA to compare (**Figures 5.5-5.7**). We used NBD-FA (Chapter 2, 3, and 4), carboxyfluorescein (CF), and 2-anthracenecarboxylate (AC, Chapter 2 and 3). In each case, we observed that c-t-BSA binds the indicators less strongly than native BSA, indicating that some denaturation of the protein occurred upon modification. However, the indicators did still bind to c-t-BSA, so we knew that we had not completely denatured the protein by modifying it.



Scheme 5.6 Alkylation of thiol by iodoacetamide.

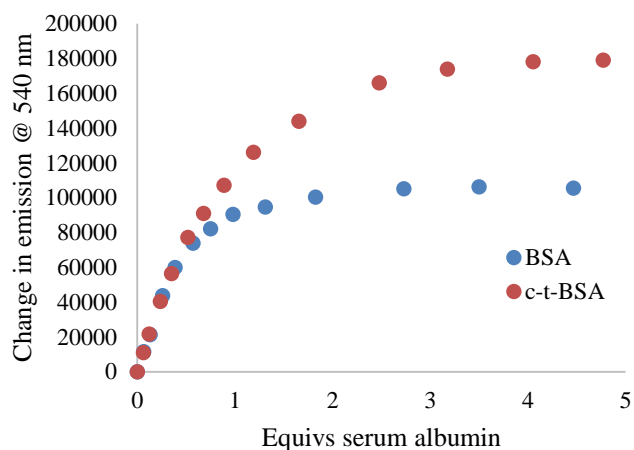


Figure 5.5 *Titration of BSA and c-t-BSA into NBD-FA.* BSA (0-68 μM) and c-t-BSA (0-73 μM) were added to NBD-FA (15.3 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 , $\lambda_{\text{ex}} = 470 \text{ nm}$.

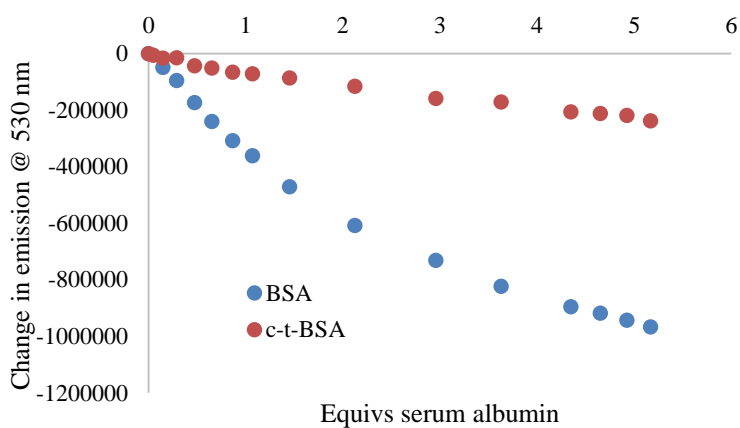


Figure 5.6 *Titration of BSA and c-t-BSA into CF.* BSA (0-52 μM) and c-t-BSA (0-52 μM) were added to CF (10 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 , $\lambda_{\text{ex}} = 490 \text{ nm}$.

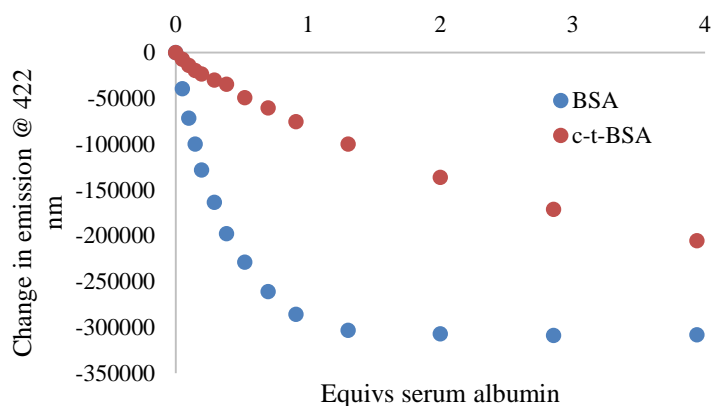
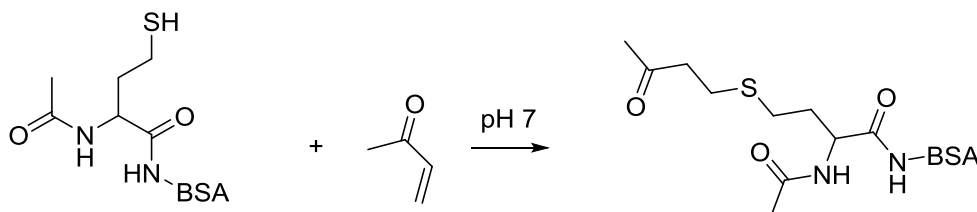


Figure 5.7 Titration of BSA and c-t-BSA into AC. BSA (0-39 μM) and c-t-BSA (0-39 μM) were added to AC (10 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 , $\lambda_{\text{ex}} = 386 \text{ nm}$.

5.3.1.1 Capping with Methyl Vinyl Ketone

Since we ultimately wanted to react the new thiols that we were installing on BSA with conjugate acceptors, we switched to methyl vinyl ketone (MVK) for our capping agent. A sample of BSA was thiolated with AHTL (17 thiol per BSA by Ellman's test), and the thiols were reacted with MVK (**Scheme 5.7**). We titrated this sample (k-t-BSA) into 1,8-anilinonaphthalene sulfonate (ANS, Chapter 2 and 3) in order to study the state of the serum albumin's structure (**Figure 5.8**). ANS binds in the crevices between domains of serum albumin, so it is a good indicator of denaturation of the protein.³³ From our titration, we observed that the k-t-BSA was unfolded somewhat compared to native BSA. However, some ANS could still bind, so the protein was not completely unfolded.



Scheme 5.7 Alkylation of thiol by methyl vinyl ketone (MVK).

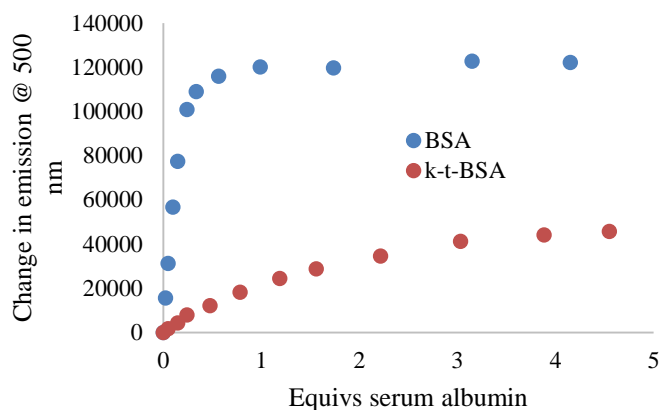
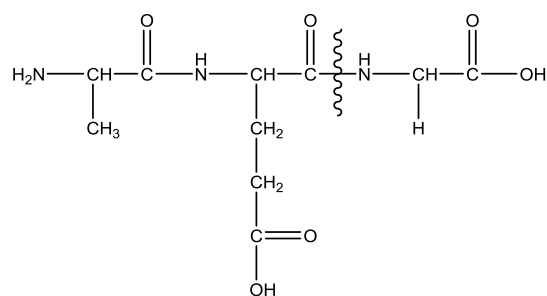


Figure 5.8 *Titration of BSA and k-t-BSA into ANS.* BSA (0-42 μM) and c-t-BSA (0-46 μM) were added to ANS (10 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 , $\lambda_{\text{ex}} = 400 \text{ nm}$.

At this point we began to explore characterization of the modified protein samples by mass spectrometry. To prepare the sample for analysis, the protein was fully reduced with DTT, and the exposed thiols were capped with iodoacetamide. Then, the protein was digested with endoproteinase Glu-C, which cleaves primarily at glutamic acid residues but also, to a lesser extent, at aspartic acid (**Scheme 5.8**).³⁴ The resultant peptide mixture was separated on a C-18 column and then introduced to the mass spectrometer by electrospray ionization (ESI). Collision-induced dissociation was used to fragment the peptides in order to sequence them (MS^2). In this way, the exact lysines that had been modified were identified by a protein database search.

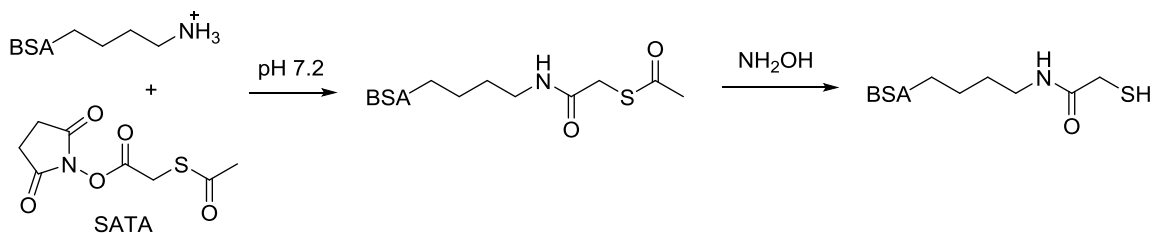


Scheme 5.8 *Cleavage of endoproteinase Glu-C.*

Thirteen lysines were identified as having been thiolated and then capped with MVK (Lys-12, 41, 159, 173, 232, 239, 280, 294, 322, 350, 465, 544, and 573). Two lysines had been thiolated but were alkylated with iodoacetamide during the MS sample preparation (Lys-41, 285), and several lysines were alkylated directly by MVK (Lys-57, 93, 132, 136, 159, 294, 499, and 544). Only 60% sequence coverage was achieved, probably due to incomplete digestion, so this data likely represents only some of the modifications present.

5.3.2 *N*-succinimidyl *S*-acetylthioacetate (SATA)

We also tested SATA for thiolating serum albumin. The amine of lysine adds to the activated ester of SATA, forming an amide bond and installing an *S*-acetyl protected thiol (**Scheme 5.9**).³⁵ The thiol is deprotected by treatment with hydroxylamine.



Scheme 5.9 *Thiolation of BSA by SATA and hydroxylamine.*

BSA was reacted with 30 equivalents of SATA. The resultant sample (p-t-BSA) was subjected to a quantitative ninhydrin test to determine how many lysine residues had been thiolated. The test indicated that 19 lysines, per BSA had been thiolated. The solubility of this p-t-BSA sample was problematic, presumably because many of the charged groups on the protein surface had been blocked with a neutral species. However, using a small amount of DMSO (4% v/v) helped the protein to fully dissolve in buffer.

Before we deprotected the thiol and created the opportunity for disulfide formation and aggregation of the proteins, the p-t-BSA sample was titrated into NBD-FA to probe its structure (**Figure 5.9**). From this experiment, we could see that the modified protein was somewhat denatured compared to native BSA, which was consistent with our t-BSA samples prepared from AHTL.

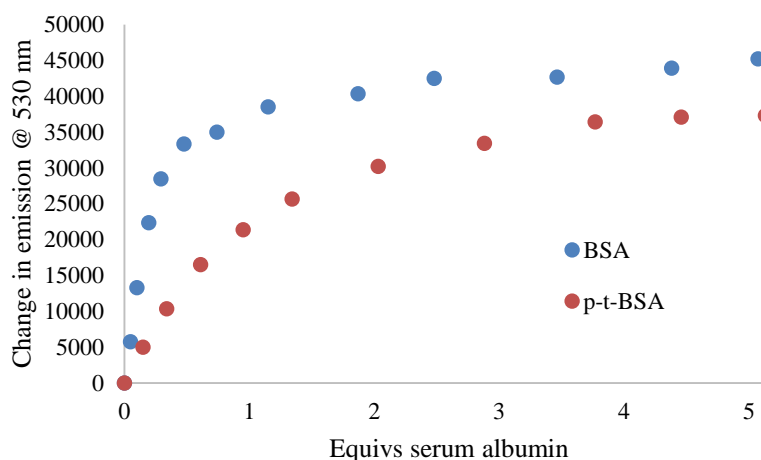


Figure 5.9 Titration of BSA and p-t-BSA into NBD-FA. BSA (0-51 μ M) and p-t-BSA (0-51 μ M) were added to NBD-FA (10 μ M) in 10 mM phosphate buffer, H₂O, pH 7.00, 0.02% NaN₃, $\lambda_{\text{ex}} = 470$ nm.

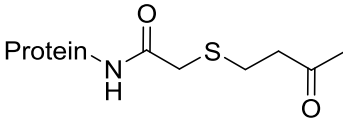
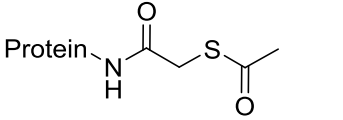
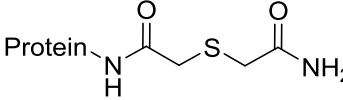
The p-t-BSA was deprotected with hydroxylamine, but the Ellman's test of the t-BSA indicated only 4 thiols per BSA, which is much lower than the ninhydrin test for

amine indicated. Furthermore, the MS analysis found only that some lysines had been acetylated by SATA, and none had the expected *S*-acetylthioglycolate modification.

A new batch of SATA was purchased from a different vendor, and the procedure was repeated with BSA and 50 equivalents of SATA to yield a new p-t-BSA sample. A portion of the sample was submitted for MS analysis, and the rest was deprotected with hydroxylamine to free the thiols. An Ellman's test of this t-BSA sample indicated 13 thiols per BSA. Furthermore, the MS analysis of p-t-BSA identified 11 lysines that had an *S*-acetylthioglycolate moiety (Lys- 41, 64, 93, 131, 159, 239, 499, 520, 537, 544, and 573). The t-BSA sample was reacted with MVK to alkylate the thiols. This k-t-BSA sample was also analyzed by MS: 21 lysines were found to have been thiolated and then reacted with MVK (Lys- 12, 41, 64, 93, 127, 131, 132, 136, 159, 211, 221, 224, 294, 322, 350, 414, 520, 523, 524, 544, and 573). This result, while positive, was rather inconsistent with the MS analysis and Ellman test of the corresponding p-t-BSA sample, which indicated fewer modifications.

Thus, new k-t-BSA samples were prepared in the same way to study the reproducibility of the modification procedure. The Ellman test indicated only 6 thiols per BSA and 10 thiols per BSA for these t-BSA samples; however, the MS results for the corresponding k-t-BSA samples found many of the same modifications as the first k-t-BSA (**Table 5.1**). One of the replicate samples (#2) was used for binding studies (**Figures 5.10-5.12**). As with the other modified serum albumin samples, k-t-BSA did not bind the fluorescent indicators as strongly as native BSA, indicating some degree of denaturation due to the modification.

Table 5.1 Summary of the lysine modifications of the *p-t*-BSA replicate samples.

Modification Type	Sample 1	Sample 2	Sample 3
 (lysine was thiolated, deprotected, and reacted with MVK)	Lys-12, 41, 64, 93, 127, 131, 132, 136, 159, 211, 221, 224, 294, 322, 350, 413, 520, 523, 524, 544, 573	Lys-12, 41, 64, 93, 131, 132, 136, 159, 211, 221, 224, 294, 322, 350, 413, 520, 523, 524, 544, 573	Lys- 12, 41, 64, 93, 127, 131, 132, 136, 159, 294, 322, 350, 465, 499, 520, 523, 524, 544, 573
 (lysine was thiolated but not deprotected)	280, 537	280, 322	12, 173, 285, 520, 533, 535, 573
 (lysine was thiolated and deprotected but did not react with MVK)	136	322, 544	51, 180

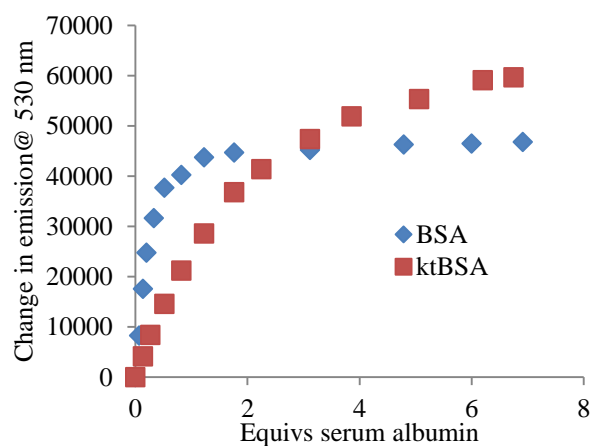


Figure 5.10 Titration of BSA and *k-t*-BSA into NBD-FA. BSA (0-69 μ M) and *k-t*-BSA (0-68 μ M) were added to NBD-FA (10 μ M) in 10 mM phosphate buffer, H₂O, pH 7.00, 0.02% NaN₃, $\lambda_{\text{ex}} = 470$ nm.

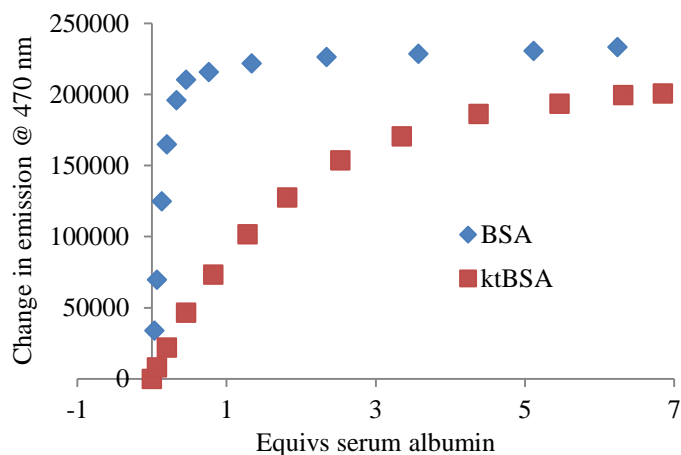


Figure 5.11 *Titration of BSA and k-t-BSA into ANS.* BSA (0-62 μM) and k-t-BSA (0-68 μM) were added to ANS (10 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 , $\lambda_{\text{ex}} = 400 \text{ nm}$.

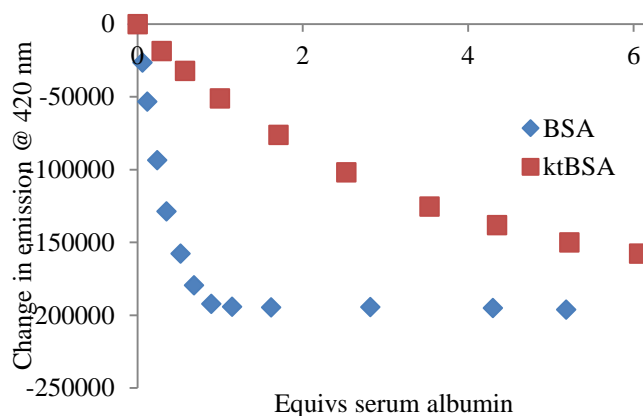


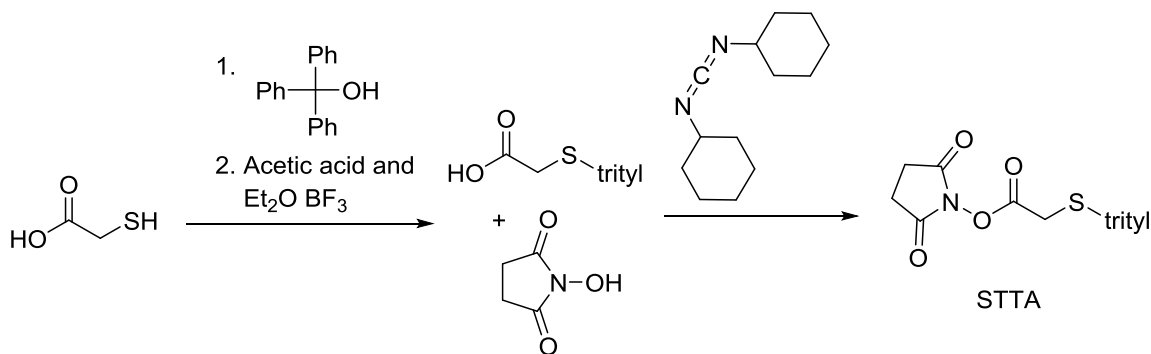
Figure 5.12 *Titration of BSA and k-t-BSA into AC.* BSA (0-43 μM) and k-t-BSA (0-50 μM) were added to ANS (8.3 μM) in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 , $\lambda_{\text{ex}} = 386 \text{ nm}$.

As we continued to repeat the thiolation with SATA and subsequent reaction with MVK, we observed a significant amount of acetylation of lysines (around 20 lysines) and fewer of the desired modifications (around 8 lysines) with new samples. We also detected many more lysines that had been thiolated and deprotected but had not reacted with MVK,

presumably because these thiols were tied up in disulfide bonds during the alkylation procedure (around 17 lysines).

5.3.3 *N*-succinimidyl *S*-tritylthioacetate (STTA)

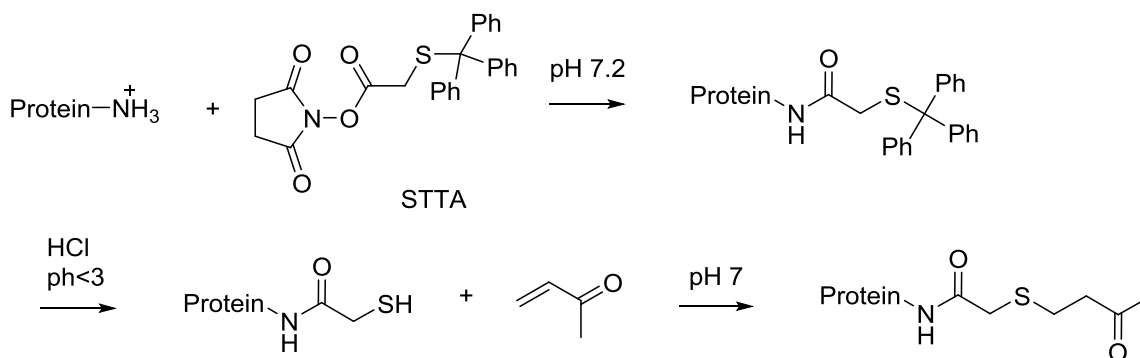
We synthesized our own variant of SATA with a trityl protecting group that could be removed by acid (**Scheme 5.10**). This modification was desirable because we were concerned with disulfide bond formation during the deprotection of SATA, which was performed as neutral pH. If the deprotection could be done under acidic conditions, then disulfide bond formation would be inhibited, at least until the pH was brought back up for the alkylation of t-BSA by MVK. Thioglycolic acid was trityl protected according to a literature procedure.³⁶ Then the acid was esterified with *N*-hydroxysuccinimide.³⁵



Scheme 5.10 *Synthesis of STTA.*

BSA was reacted with 100 equivalents of STTA, and the protected sample was submitted for MS analysis (**Scheme 5.11**). The p-t-BSA had only a few of the desired modifications (Lys-20, 41, 64, 159, 242, 294, 350, 544, and 573). The removal of the trityl group was attempted by acidifying the p-t-BSA solution with HCl to a pH of 2. However, the protein began to precipitate at this low pH, and the trityl group did not appear to have been removed according to a negative Ellman's test. This procedure was repeated with

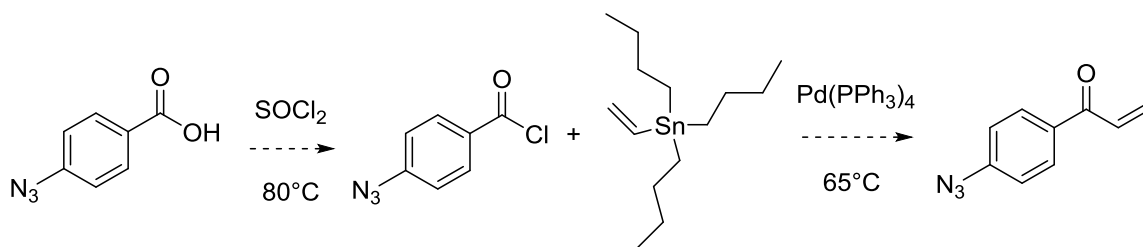
TFA instead of HCl, but again no Ellman-positive thiol was generated. We concluded that the acidity necessary to remove the trityl group precipitates the protein, so this reagent is not suitable for use with proteins.



Scheme 5.11 *Thiolation of BSA by STTA and subsequent deprotection and reaction with MVK.*

5.3.4 Synthesis of a Conjugate Acceptor

While most of our efforts went toward generating more thiol groups on serum albumin, we did begin to synthesize a conjugate acceptor partner to react with these thiols. We designed the synthesis of an α,β -unsaturated ketone with an azide (**Scheme 5.12**). We wanted to use an α,β -unsaturated ketones because they are more reactive than their corresponding ester or amide.¹³ The desired recognition units (e.g. peptides, boronic acids) could be attached to the conjugate acceptor by copper-catalyzed alkyne-azide cycloaddition (CuAAC).^{37,38} However, the Stille coupling to install the vinyl group was unsuccessful. The azide was reduced to the amine instead, probably by a Staudinger reduction.



Scheme 5.12 Attempted synthesis of the conjugate acceptor.

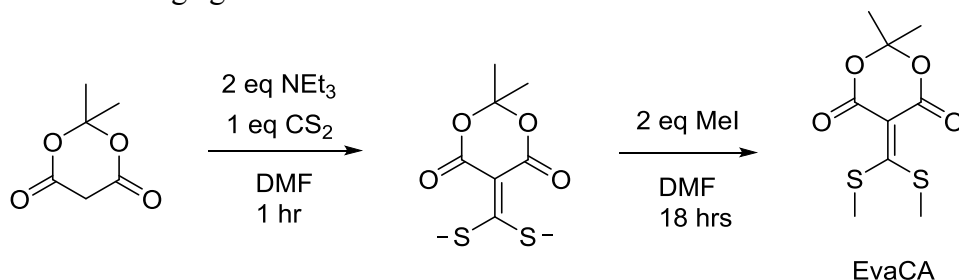
5.3.5 Summary

While our initial strategy to partially reduce serum albumin resulted in unstable and insoluble samples, we were able to use both AHTL and SATA to thiolate serum albumin, adding as many as twenty thiols per protein. We further demonstrated that these thiols could be alkylated with a conjugate acceptor. We characterized these modified BSA samples using chemical tests (Ellman, ninhydrin) and by mass spectrometry. However, we became concerned that the thiolated proteins were aggregating and then crosslinking through disulfide bond formation. Additionally, the thiolation reactions with SATA became inconsistent in later samples. Finally, we were struggling with the synthesis of an α,β -unsaturated ketone that could be easily conjugated to recognition units. Thus, we switched to an alternative strategy.

5.4 LABELLING OF SERUM ALBUMIN BY A CONJUGATE ACCEPTOR (EVA CA)

As an different approach to labeling serum albumin, we became interested in the chemistry of 5-alkylidene-2,2-dimethyl-1,3-dioxane-4,6-dione derivatives such as the one shown in **Scheme 5.13**. The particular derivative shown, which we dubbed EvaCA, was reported by Wentrup and coworkers to add amines.^{39,40} EvaCA was synthesized according to their procedure from Meldrum's acid, carbon disulfide, and methyl iodide (**Scheme 5.13**). While Wentrup *et al.* observed that EvaCA added two amines, we hypothesized that the addition of one amine deactivates the conjugate acceptor to further addition due to

resonance donation by the nitrogen. Then, a thiol would add more readily than a second amine to EvaCA, due to its greater nucleophilicity. Thus, EvaCA could be used as a selective crosslinking agent for amines and thiols.

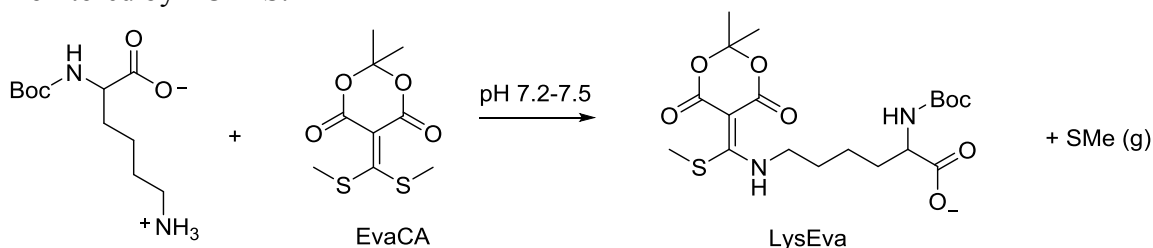


Scheme 5.13 *Synthesis of EvaCA.*

5.4.1 Reactions of EvaCA with Lysine and Cysteine

5.4.1.1 Lysine and Cysteine

First, we wanted to test our hypothesis about the reactivity of EvaCA. EvaCA was reacted with α -Boc-lysine (**Scheme 5.14**). Nitrogen was bubbled through the solution to aid in removing methyl mercaptan gas from the solution. The reaction progress was monitored by LC-MS.

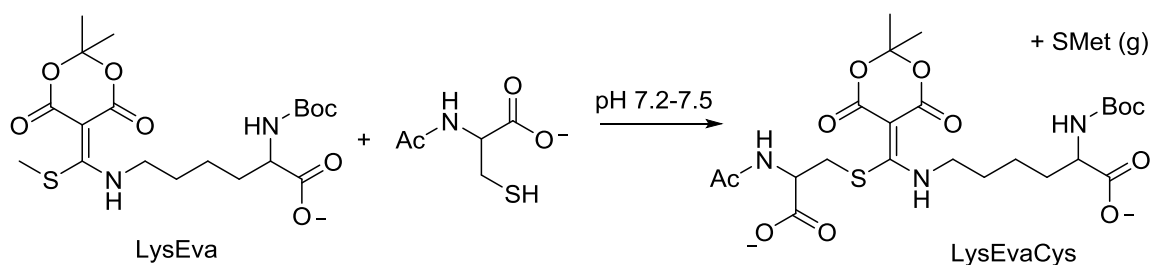


Scheme 5.14 *Reaction of EvaCA with α -Boc-lysine to form LysEva.*

For these reactions, aqueous phosphate buffer at pH 7.2 or 7.5 was used, and there was no measurable difference in the speed of the reaction between the two pHs. We found a strong concentration dependence on the speed of the reaction. For example, a reaction that was 37 mM in EvaCA and α -Boc-lysine completed in about 18 hours, while a reaction

that was 8 mM in EvaCA and α -Boc-lysine completed in about 72 hours. One equivalent of lysine was used, and no diaddition product (LysEvaLys) was observed under these conditions, which supports our hypothesis about EvaCA's reactivity.

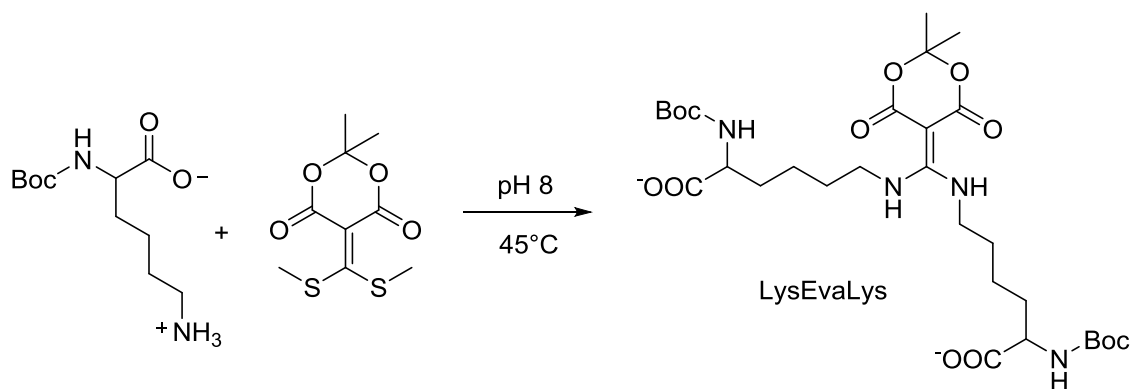
Next, NAC was added to LysEva at the same pH (**Scheme 5.15**). The nitrogen flow was continued, and the reaction was again monitored by LC-MS. The reaction took over about a week to reach 90% conversion of LysEva to LysEvaCys.



Scheme 5.15 Reaction of LysEva with *N*-acetylcysteine (NAC) to form LysEvaCys.

5.4.1.2 Lysine

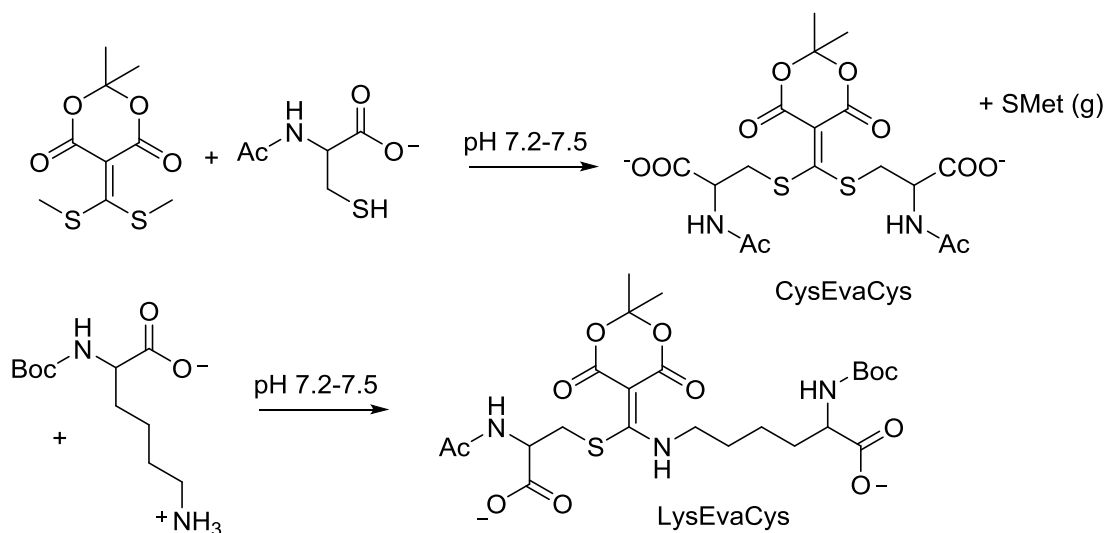
Looking toward reacting EvaCA with serum albumin, we wanted to know if two amines would be able to add to EvaCA under our conditions. If two amines could readily add, then there could be crosslinking between proteins, which would be undesirable. EvaCA was reacted with five equivalents of α -Boc-lysine at pH 8 and 45 °C (**Scheme 5.16**). After 18 hours, an aliquot of the reaction was taken and analyzed by LC-MS. The analysis showed that only a small amount of LysEvaLys had formed. Thus, under the conditions we would be using for BSA labeling, we expected only one lysine to add to each EvaCA.



Scheme 5.16 Reaction of *EvaCA* with excess α -Boc-lysine to form *LysEvaLys*.

5.4.1.2 Cysteine

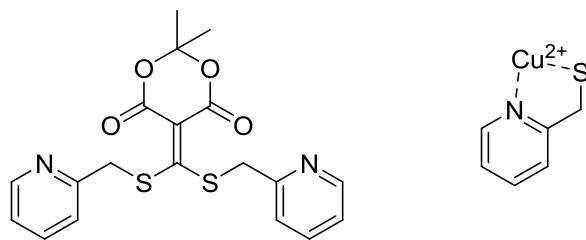
Due to the relatively slow addition of cysteine to LysEva, we wanted to explore other approaches to using *EvaCA* to label the protein. If we could essentially exchange out methyl mercaptan for other thiols, then we could use the resultant CysEvaCys to label the lysines of serum albumin in one step. To test this hypothesis, *EvaCA* was reacted with two equivalents of NAC at pH 7.5 and room temperature (**Scheme 5.17**). After 18 hours, the *EvaCA* was converted entirely to CysEvaCys. Then, one equivalent of α -Boc-lysine was added to the reaction. After three days, the CysEvaCys had largely been converted to LysEvaCys. While the approach worked, it is wasteful in terms of thiol because one of the thiols from each CysEvaCys is lost, and the reaction with an amine takes several days to complete.



Scheme 5.17 Reaction of *EvaCa* with NAC to form *CysEvaCys* and then with α -Boc-lysine to form *LysEvaCys*.

5.4.2 Synthesis of an EvaCA Derivative

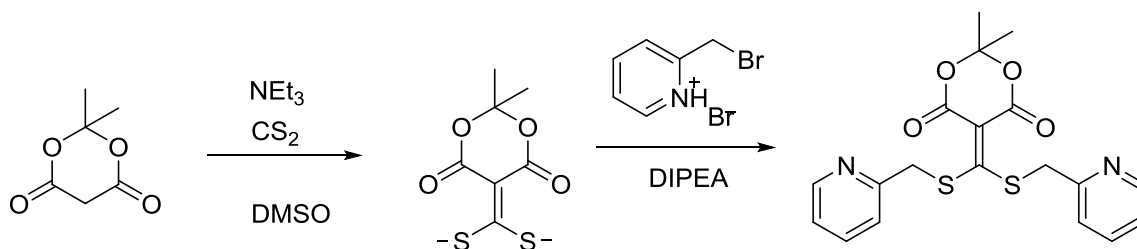
Since *LysEva* reacted slowly with cysteine, we postulated that a derivative of *EvaCA* from pyridin-2-ylmethanethiol (**Scheme 5.18**, *PyrCA*) could be synthesized and used for a faster alternative to *EvaCA*. If an ion such as copper(II) or zinc(II) was present during the reaction of *PyrCA* with lysine and cysteine, then the pyridin-2-ylmethanethiol would be a better leaving group by coordination to the ion (**Scheme 5.18**).



Scheme 5.18 Structure of *PyrCA* (left) and pyridin-2-ylmethanethiol coordinated to copper(II) (right).

5.4.2.1 Synthesis of PyrCA via 2-(Bromomethyl)pyridine

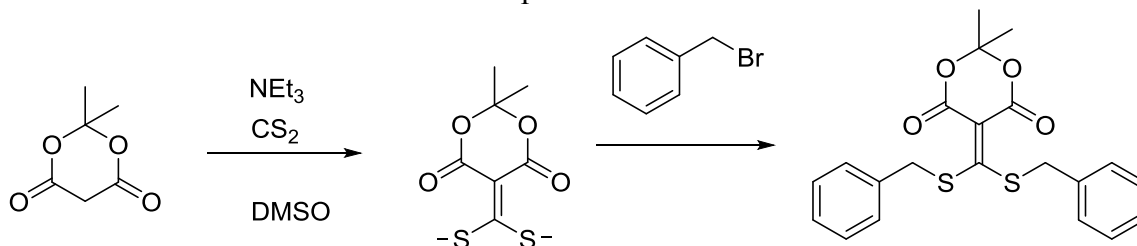
Meldrum's acid was reacted with carbon disulfide and then with 2-(bromomethyl)pyridine (BMP) (**Scheme 5.19**). However, the reaction did not produce PyrCA. From the TLC and LC-MS analysis, several products were formed, but we were unable to identify most of them from their masses. The reaction was attempted again with the free base of BMP. This time, the mass of the product was identifiable from LC-MS analysis of the reaction mixture. We attempted to isolate the PyrCA by column chromatography but were never successful. The compound seemed to be oxidizing or otherwise degrading on the column.



Scheme 5.19 Synthesis of PyrCA via 2-(bromomethyl)pyridine.

5.4.2.2 Synthesis of BenzCA via Benzylbromide

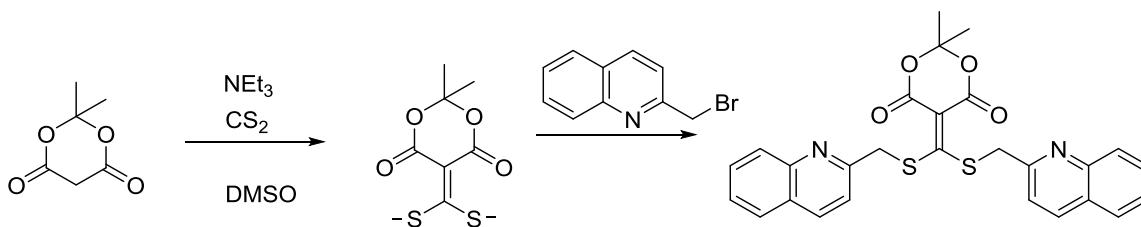
We synthesized BenzCA from Meldrum's acid and benzylbromide (BB) (**Scheme 5.20**) to confirm that the problem with synthesizing PyrCA had to do with the pyridine. The crude product was not purified, but ^1H NMR and LCMS analysis of the reaction mixture indicated the BenzCA had been produced.



Scheme 5.20 *Synthesis of BenzCA via benzylbromide.*

5.4.2.3 Synthesis of QuinCA via 2-(Bromomethyl)quinolone

Since PyrCA was unstable under ambient conditions, we decided to try to make QuinCA from 2-(bromomethyl)quinolone instead (**Scheme 5.21**). This derivative also has a nitrogen atom for coordinating metal ions, but we hoped that the quinolone would be less prone to oxidation. First, 2-(chloromethyl)quinolin-1-ium chloride was converted to the free base and then to the bromide with LiBr.⁴¹ The 2-(bromomethyl)quinolone was obtained in 39% yield. Then Meldrums' acid and 2-(bromomethyl)quinolone were reacted together to make QuinCA. ¹H NMR and LC-MS analysis of the crude material were not indicative of product formation. As with PyrCA, a variety of products were formed, but we could not identify them. Curiously, the ¹H NMR had no peaks in the methylene region of the spectrum. We tried the reaction again according to another procedure for making EvaCA that used potassium fluoride as the base instead of triethylamine. Again, the ¹H NMR and LC-MS of the crude material did not show product formation. We attempted to isolate the species in the crude mixture to better characterize them, but none of the components that we were able to isolate were identifiable as QuinCA.

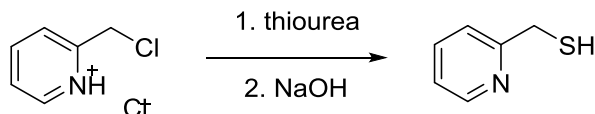


Scheme 5.21 *Synthesis of QuinCA via 2-(bromomethyl)quinolone.*

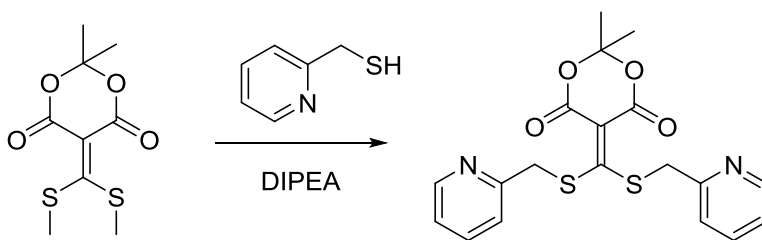
5.4.2.4 Synthesis of PyrCA via Pyridin-2-ylmethanethiol

As an alternative approach to accessing PyrCA, we made 2-pyridinemethanethiol (PMT) and then reacted it with EvaCA. PMT was synthesized according to a literature

procedure in 6.3% yield (**Scheme 5.22**).⁴² Then, the PMT was reacted with EvaCA (**Scheme 5.23**). However, ¹H NMR and LC-MS analysis of the crude material did not show any product. Thus, we abandoned our attempts to synthesize PyrCA or QuinCA and continued forward with EvaCA.



Scheme 5.22 *Synthesis of 2-pyridinethanethiol.*



Scheme 5.23 *Synthesis of PyrCA via 2-pyridinethanethiol.*

5.4.3 Synthesis of a Model BSA Peptide and Cysteine-Containing Peptides

For MS analysis of modified BSA samples, the protein had to be digested, and the resultant peptides fragmented by CID. However, since we found that EvaCA can be fragmented more easily than the peptide backbone by CID, a model peptide was necessary to characterize this fragmentation in order to develop an appropriate method for analysis of BSA samples. Additionally, the model peptide allowed for further studies of EvaCA's reactivity.

A peptide sequence from BSA that would appear in Glu-C digestion and that contained one lysine was selected (IAHRFKDLGE) and synthesized via solid-phase peptide synthesis (SPPS). The N-terminus was acetylated to prevent any reaction with

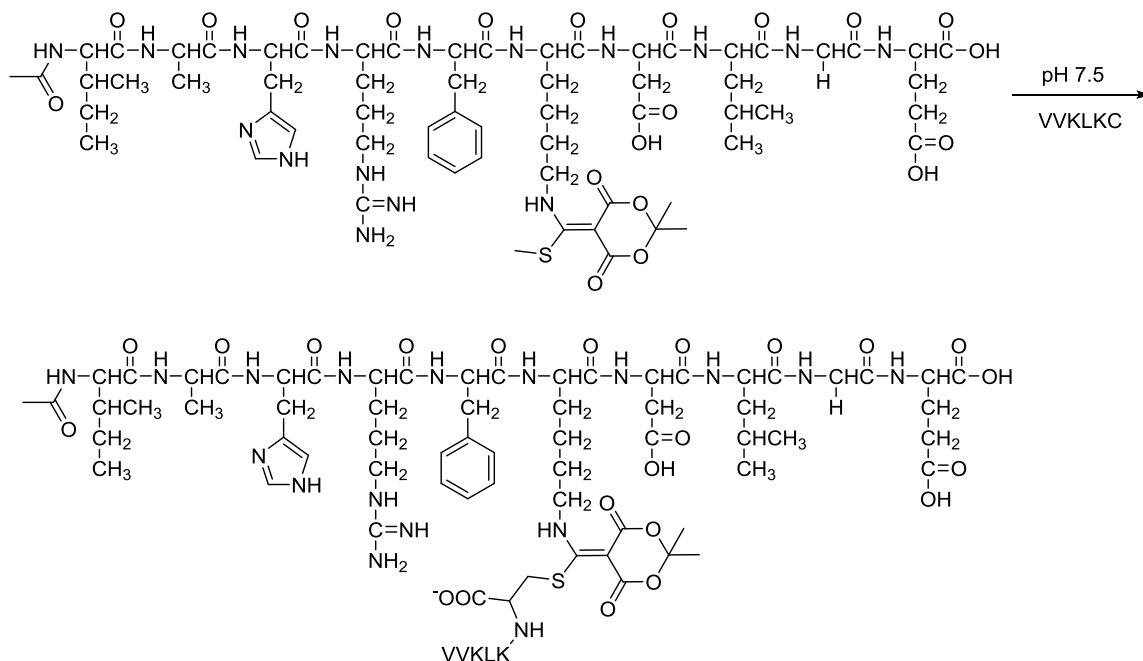
EvaCA (acIAHRFKDLGE). Then, the crude peptide material was reacted with EvaCA to yield acIAHRFK(Eva)DLGE. The peptide was purified by RP-HPLC, but acIAHRFK(Eva)DLGE was not resolved from acAHRFK(Eva)DLGE. Thus, a mixture of the two peptides was obtained and used.

A series of cysteine-containing peptides were also synthesized for attaching to BSA through EvaCA. Previous work in our group toward the development of receptors built from a DNA scaffold used pentapeptides conjugated to a DNA intercalator.⁴³ The same peptide motifs from the DNA study were used for initial testing of the serum albumin labeling procedure. These motifs were selected in the DNA study because they were known in the literature to bind to specific cancer cell types or cancer cell surface features, and the goal of that project was to make receptors to target cancer cell surfaces. In this case, we wanted to start by simply showing that cysteine-containing peptides could be appended to serum albumin through EvaCA. Thus, we used the pentapeptide motifs plus a cysteine residue. The nine peptides were synthesized via SPPS and purified by RP-HPLC: VVKLKC, KGGRKC, KRGSKC, KRSSKC, KAGLKC, KYPYKC, PRGDKC, KDGRKC, and KGARCC.

5.4.4 Reaction of the Model Peptide with a Cysteine-Containing Peptide

The model peptides (acIAHRFK(Eva)DLGE and acAHRFK(Eva)DLGE) were reacted with VVKLKC (**Scheme 5.24**). The resultant sample was analyzed by MS and was found to contain the expected crosslinked species. The reaction took about two weeks to complete. The sample was also used to determine the protocol necessary to analyze the peptides from a digest of a sample of BSA functionalized with EvaCA and cysteine-containing peptides. It was determined that an MS³ experiment was necessary. In the first mass spectrometer, EvaCA was fragmented. In the second mass spectrometer, both the

peptide from BSA and the cysteine-containing peptide were fragmented to allow for sequencing and identification of modified lysines. The third mass spectrometer analyzed the peptide fragments.

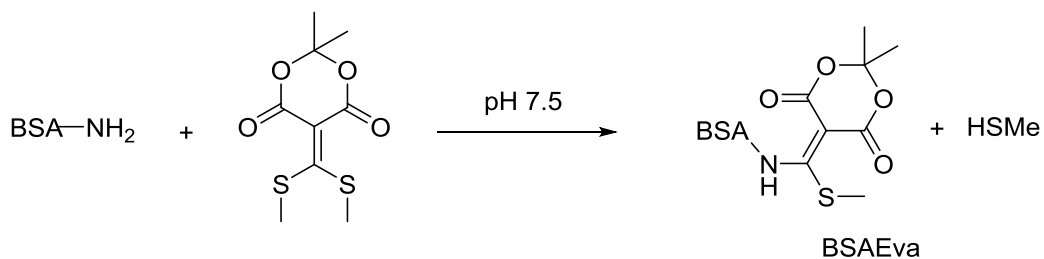


Scheme 5.24 *Synthesis of MPEvaVVKLKC.*

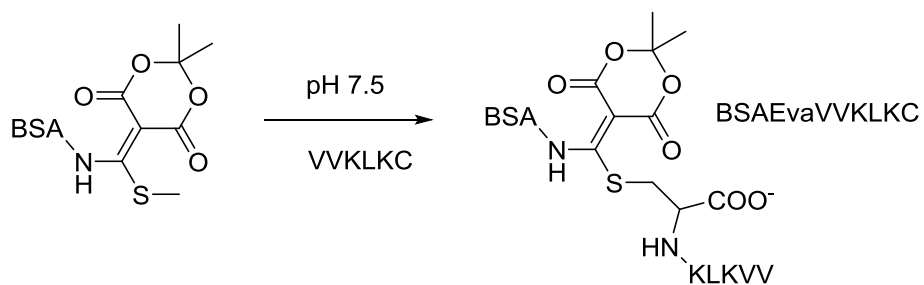
5.4.5 Reaction of EvaCA with BSA

Two samples of BSA were reacted with EvaCA to make BSAEva (**Scheme 5.25**). One of the samples was submitted for MS analysis at this stage, and the other sample was further reacted with VVKLKC to make BSAEvaVVKLKC (**Scheme 5.26**). The BSAEvaVVKLKC was also submitted for MS analysis. The BSAEva sample was also subjected to a ninhydrin test, and it was found that 12 lysines per BSA had been reacted. The MS analysis of BSAEva indicated that 14 lysines had been modified with EvaCA (Lys-12, 20, 51, 64, 131, 159, 242, 261, 273, 312, 350, 362, 535, and 573). BSAEvaVVKLKC

had 14 lysines modified with EvaCA (Lys-12, 20, 51, 64, 131, 173, 180, 273, 350, 362, 377, 413, 535, and 573) but only two of those modifications also had VVLKLC (Lys-64, 350). **Figure 5.13** shows the modified lysines of the samples.



Scheme 5.25 *Synthesis of BSAEva.*



Scheme 5.26 *Synthesis of BSAEvaVVKLKC.*

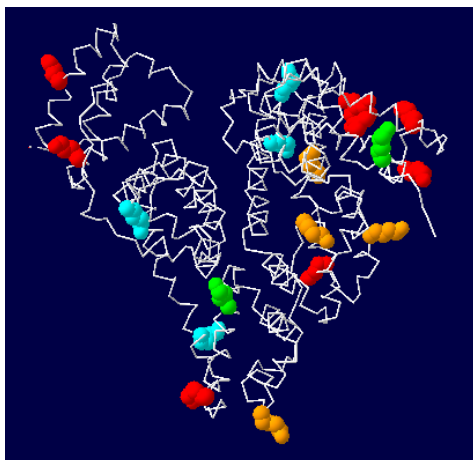


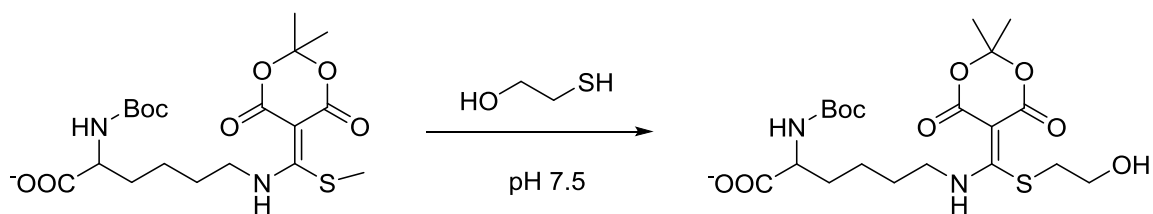
Figure 5.13 *Crystal structure of native BSA with the modified lysines in BSAEva and BSAEvaVVKLKC highlighted. Key: red- EvaCa in both, orange- EvaCA in BSAEva only, blue- EvaCA in BSAEvaVVKLKC only, green- EvaCA and peptide in BSAEvaVVKLKC.*⁴⁴

5.4.6 Alternative Thiols

Since we were only able to modify two lysines in BSA with VVKLKC, we next investigated other thiols.

5.4.6.1 2-Mercaptoethanol

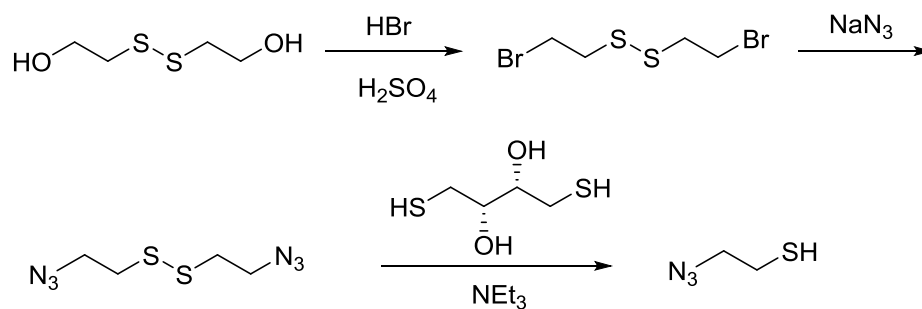
We postulated that a smaller thiol would be more nucleophilic, so we tried reacting 2-mercaptoethanol with LysEva (LysEvaBME, **Scheme 5.27**). The reaction was monitored by LC-MS and was found to be complete after five hours. As hypothesized, this reaction was much faster than the cysteine reaction, which took about a week to complete. We observed additional species that we could not identify based on their masses, which indicated that side reactions were occurring.



Scheme 5.27 *Synthesis of LysEvaBME.*

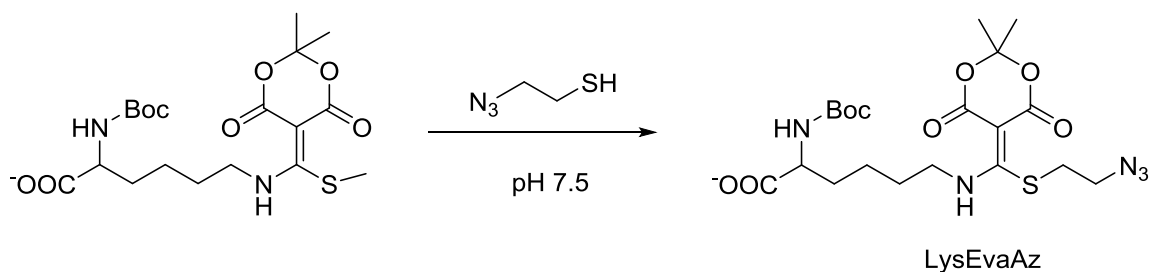
5.4.6.2 2-Azidoethanethiol

We chose a small thiol with an azide group (**Scheme 5.28**) that we could use CuAAC to attach additional units to in a second step. Thus, 2-azidoethanethiol was synthesized according to literature procedure starting from the disulfide of 2-mercaptoethanol.^{38,45}



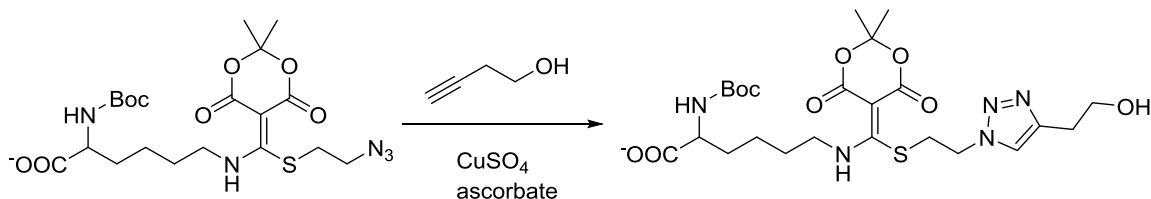
Scheme 5.28 *Synthesis of 2-azidoethanethiol.*

The thiol was reacted with LysEva to form LysEvaAz (**Scheme 5.29**). LC-MS analysis of the reaction showed the desired product, but there were also other species present that could not be identified. As with 2-mercaptoethanol, side reactions were occurring.



Scheme 5.29 *Synthesis of LysEvaAz.*

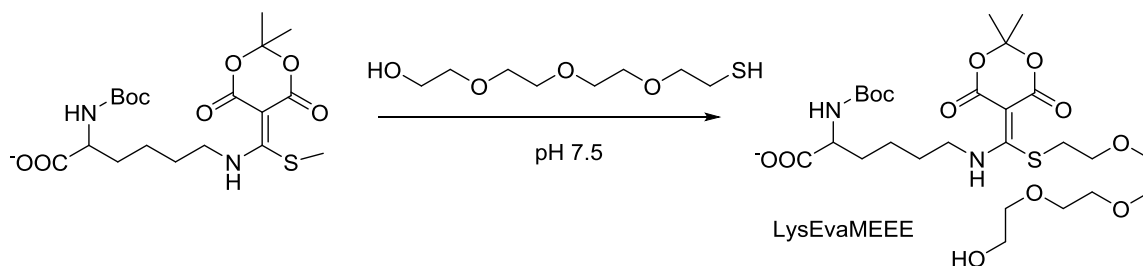
Next, the LysEvaAz sample was subjected to CuAAC conditions to test if this reaction would work (**Scheme 5.30**). However, the desired product was not observed. We did observe the mass of the CuAAC product of 2-azidoethanethiol and 3-butyne-1-ol, which indicates that the CuAAC reaction was working. We also did not observe the mass of LysEvaAz after the sample was subjected to the CuAAC conditions.



Scheme 5.30 *Click reaction of 3-butyne-1-ol with LysEvaAz.*

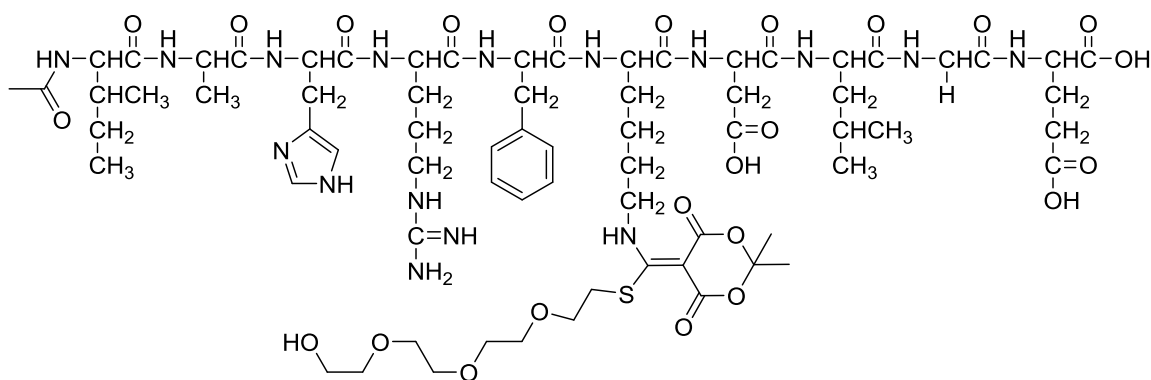
5.4.6.3 2-{2-[2-(2-Mercaptoethoxy)ethoxy]ethoxy}ethanol

Due to the side reactions observed with the small thiols, we wanted to try a slightly larger thiol. LysEva was reacted with 2-{2-[2-(2-mercaptoethoxy)ethoxy]ethoxy}ethanol (MEEEE) to make LysEvaMEEEE (**Scheme 5.31**). The reaction proceeded cleanly within a day.



Scheme 5.31 Synthesis of *LysEvaMEEEE*.

Next, MEEEEE was reacted with the model peptide (acIAHRFK(Eva)DLGE and acAHRFK(Eva)DLGE) to develop the MS method necessary for characterizing a protein modified in this way (**Scheme 5.32**).



Scheme 5.32 Synthesis of *MPEvaMEEEE*.

5.4.7 Reaction of EvaCA with Myoglobin

In order to simplify the MS analysis, we switched to myoglobin, a smaller protein than BSA, that has 19 lysines. Three replicate reactions of myoglobin with EvaCA and then MEEEEE were performed to form MyoEvaMEEEE. The resultant samples were digested and analyzed by MS³. There were five lysines that had been modified with EvaCA and MEEEEE in all three samples (Lys-17, 46, 56, 134, 148, **Figure 5.14**). Thus, using

MEEEE we were able to achieve much more efficient labeling of a protein compared to cysteine.

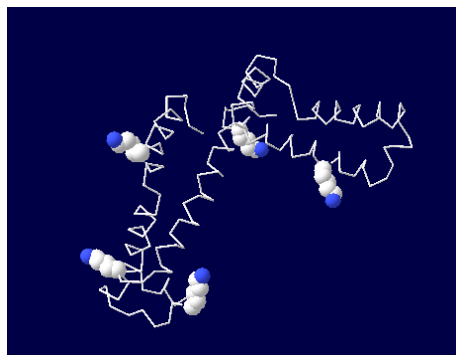
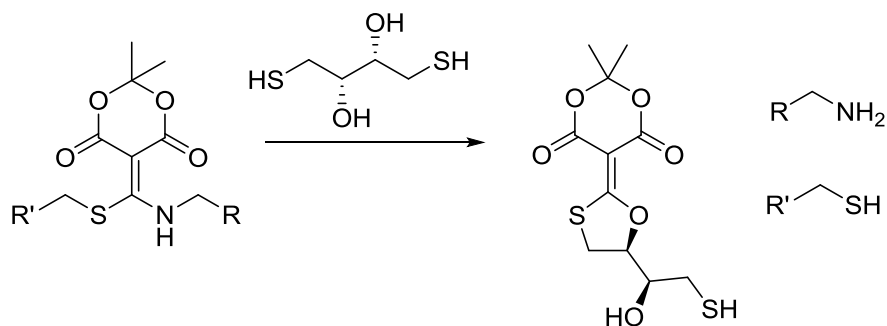


Figure 5.14 *Crystal structure of native myoglobin with the modified lysines highlighted.*⁴⁶

5.4.8 Removal of Amine and Thiol from EvaCA with Dithiothreitol

Based on the reactivity of EvaCA that we had observed, we hypothesized that dithiothreitol (DTT) could be used to remove the amine and thiol from EvaCA (**Scheme 5.33**). Indeed, when LysEvaCys was reacted with DTT, α -Boc-lysine and NAC were liberated, and the DTT-EvaCA adduct shown in **Scheme 5.33** was observed by MS. A sample of LysEvaMEEEE was also reacted with DTT, and again we observed liberation of α -Boc-lysine and MEEEE and the presence of the DTT-EvaCA adduct. However, a sample of LysEvaLys was inert to DTT, indicating that at least one of the species on EvaCA must be a thiol for DTT to effect removal of the groups.



Scheme 5.33 Removal of amine and thiol from *EvaCA* by DTT.

5.4.8.1 Structure of DTT-*EvaCA* Adduct and Mechanism of Formation

We further confirmed the structure of the DTT-*EvaCA* adduct by ^{13}C NMR to be the five-membered ring shown in **Scheme 5.33**. We propose that a seven-membered ring forms first by attack of both thiols, and then a rearrangement quickly occurs to the more stable five-membered ring. We tested both 2-mercaptoethanol and 1,2-ethanedithiol to probe this proposed mechanism. Neither compound formed a five-membered ring with *EvaCA*. Thus, we propose that DTT is essential because it contains two thiols. The alcohol of 2-mercaptoethanol is insufficiently nucleophilic to add to *EvaCA* on the time scale that we measured (up to 48 hours). Furthermore, the thiols in DTT are more nucleophilic due to the adjacent electron-withdrawing alcohols compared to the thiols in 1,2-ethanedithiol.

5.4.8.2 Removal of Label from *MyoEvaMEEEE*

With this knowledge in hand, we wanted to demonstrate that DTT could be used to remove the *EvaCA* label from a protein. The *MyoEvaMEEEE* samples were reacted with DTT overnight, and then the protein sample was purified by a molecular-weight cutoff (MWCO) filter that retains the protein but through which small molecules can pass freely. The flow-through solution containing small molecule components of the sample was analyzed by MS and was found to contain both MEEEE and the DTT-*EvaCA* adduct. Furthermore, the protein portion of the sample was also analyzed by MS and compared to

the spectrum for MyoEvaMEEEE. The sample that had been subjected to DTT showed the mass for unmodified myoglobin (**Figure 5.15**) and did not match the spectrum for MyoEvaMEEEE (**Figure 5.16**), which showed higher molecular weight species than native myoglobin.

MYO_AfterDTT_Full scan 100 uscans_XT_00001_M_#2 RT: 2.00 AV: 1 NL: 7.61E3
T: FTMS + p ESI Full ms [400.00-2000.00]

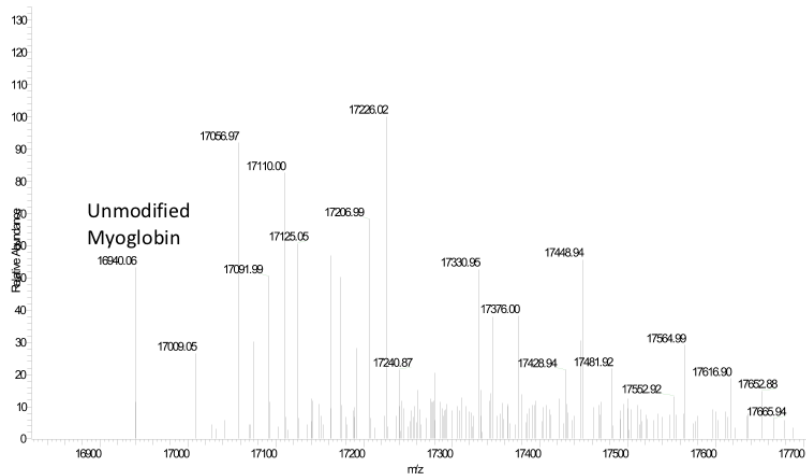


Figure 5.15 ESI spectrum of MyoEvaMEEEE after DTT treatment.

MYO_MEEB_Full scan 100 uscans_XT_00001_M_#2 RT: 2.00 AV: 1 NL: 7.64E3
T: FTMS + p ESI Full ms [400.00-2000.00]

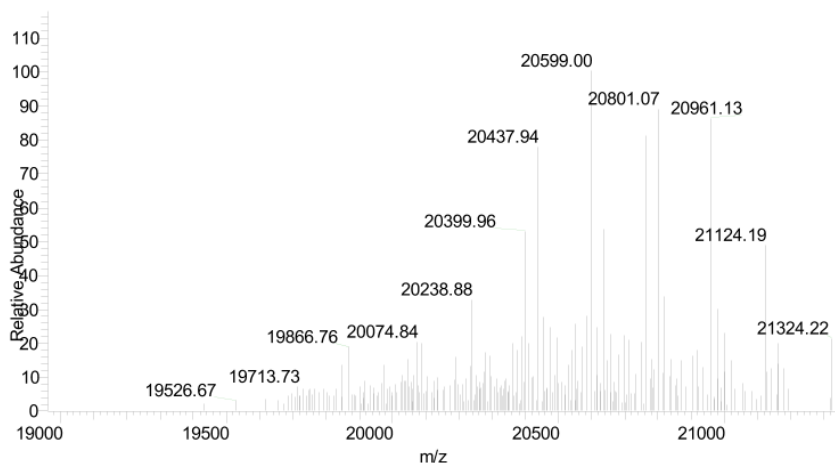
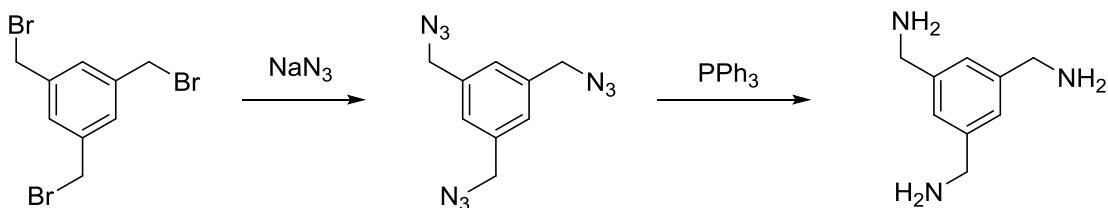


Figure 5.16 ESI spectrum of MyoEvaMEEEE.

5.5 TRIAMINE COMPOUND AS A SCAFFOLD

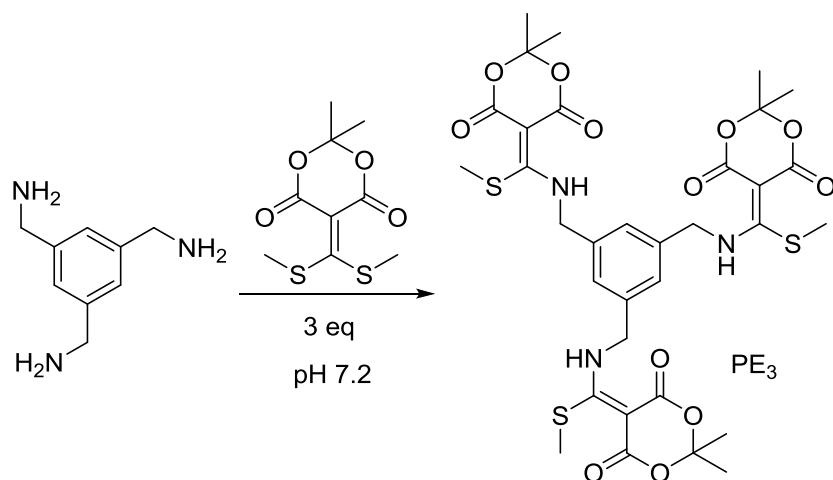
Due to our success using EvaCA as a reversible protein label, we wanted to explore the use of EvaCA in the generation of DCLs on another scaffold. We synthesized a triamine compound (1,3,5-benzenetrimethanamine) from the corresponding bromide according to literature procedure (**Scheme 5.34**).^{47,48} The spacing of the amine on the benzene ring was intended to prevent the possibility of two amines on the same molecule adding to EvaCA.



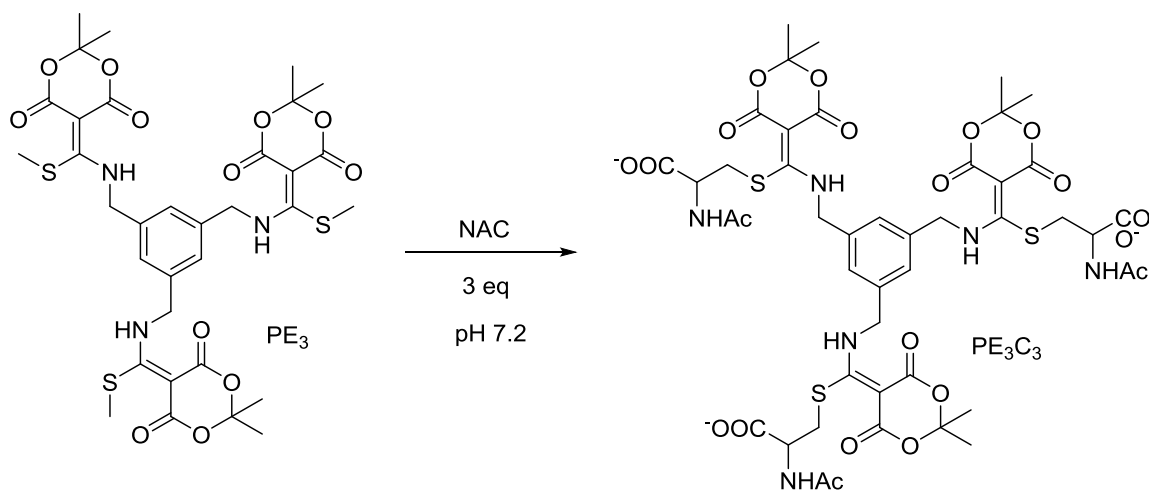
Scheme 5.34 *Synthesis of 1,3,5-benzenetrimethanamine.*

5.5.1 Reaction of the Triamine Compound with EvaCA and Cysteine

The triamine compound was reacted with EvaCA to form PE_3 (**Scheme 5.35**). The solution was sparged gently with nitrogen to remove methyl mercaptan. The product was not very soluble in water, so acetonitrile was added to help solubilize PE_3 . Next, PE_3 was reacted with NAC (**Scheme 5.36**). The reaction was monitored over time by LC-MS. After about two days, the reaction had equilibrated, converging on PE_3C_3 .



Scheme 5.35 *Synthesis of PE₃.*



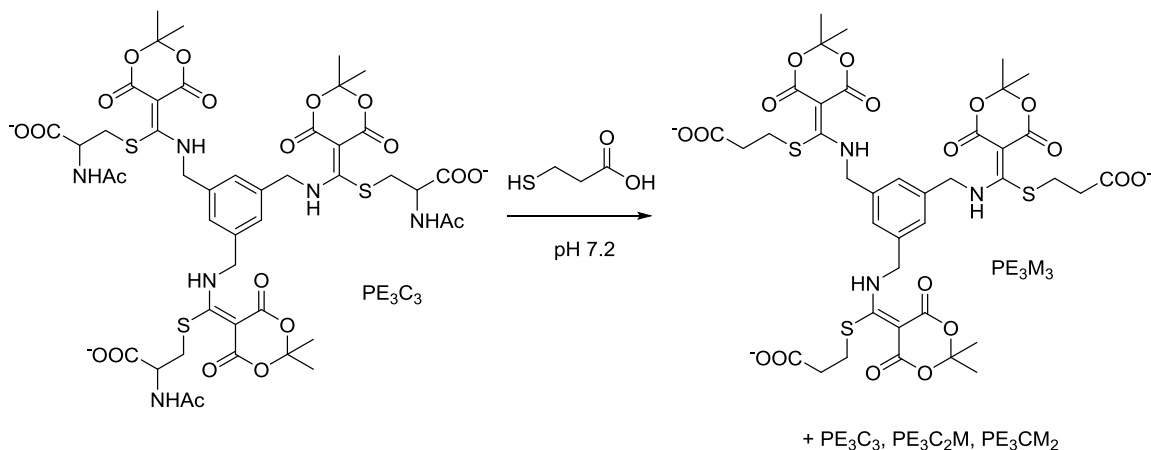
Scheme 5.36 *Synthesis of PE₃C₃.*

PE₃C₃ was further reacted with DTT to clip the triamine and the cysteine from EvaCA. LC-MS analysis showed no masses corresponding to any of the PE_xC_x compounds. Also, the mass of the DTT-EvaCA adduct was observed.

5.5.2 Reversibility of the Thiol Addition to PWEva

In order to use EvaCA in a dynamic combinatorial library, the addition of the thiol must be reversible. Thus, we investigated the reversibility of the thiol addition to PE₃ by

first synthesizing PE_3C_3 and then reacting it with another thiol of comparable nucleophilicity (2-mercaptopropionic acid, **Scheme 5.37**). The reaction was monitored by LC-MS and was found to equilibrate after about five days. As expected, at equilibrium a distribution of the possible species was observed: PE_3C_3 , PE_3C_2M , PE_3CM_2 , and PE_3M_3 .



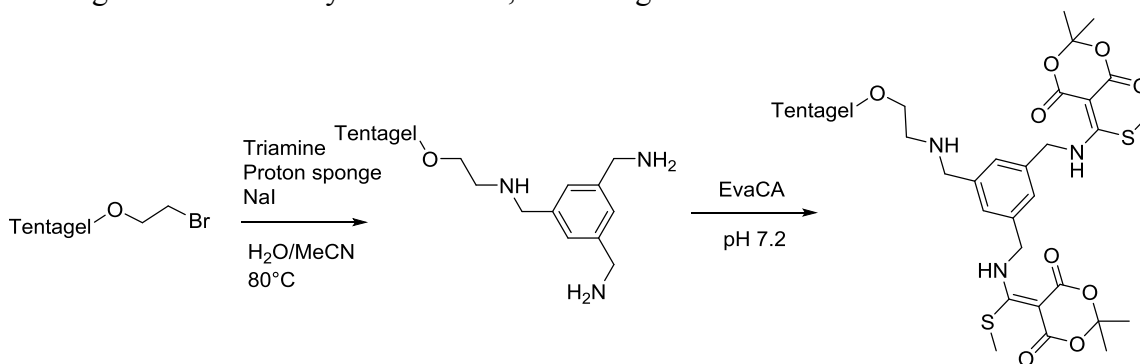
Scheme 5.37 *Thiol exchange on PE_3C_3 with 2-mercaptopropionic acid.*

5.5.3 Resin-Bound Triamine Compound

To aid in the analysis of DCLs generated from PE_3 , we decided to immobilize the triamine compound on a resin via one of the three amines.⁴⁹ Then we could build a bivalent receptor via the remaining two amines using EvaCA and thiol-containing recognition units. This immobilization step would allow us to wash away any unattached recognition units as well as the analyte of interest after equilibration of the library. In this way, the recognition units that were selected for in the equilibration of the library would remain attached to the resin. These recognition units could then be removed from the resin by DTT and analyzed by LC-MS.

Thus, we began by attaching the triamine compound to a Tentagel-Br resin⁵⁰ according to a literature procedure (**Scheme 5.38**).⁵¹ The functionalization of the resin was

characterized by a Kaiser test, which is a colorimetric test for primary amines that uses ninhydrin. After attachment of the triamine compound, the resin tested positive for amine by a Kaiser test. Next, EvaCA was reacted with the resin. After this reaction, the resin was negative for amine by a Kaiser test, indicating that the reaction had worked.



Scheme 5.38 *Immobilization of the triamine on a Tentagel S Br resin and subsequent reaction with EvaCA.*

A portion of the functionalized resin was taken and reacted with NAC for about three days. Then the resin was washed thoroughly with water and methanol to remove any unattached species. Finally, the resin was reacted with DTT, and the reaction mixture was analyzed by LC-MS. We found NAC and the DTT-EvaCA adduct, indicating that the reactions had worked as anticipated. Furthermore, the resin tested positive for amine by a Kaiser test.

5.5.4 Dynamic Combinatorial Libraries

We hypothesized that we could use the resin-bound EvaCA and thiol-containing peptides to make DCLs to template against proteins.⁵²⁻⁵⁴ To test our hypothesis, DCLs were set up from the functionalized resin and the thiol-containing peptides (KGGRKC, KRGSKC, KRSSKC, KAGLKC, KYPYKC, PRGDKC, KDGRKC, and KGARCC). One library contained BSA as a test analyte, and the other library contained no protein as a

comparison. The library was allowed to equilibrate for six days, at which time the solution was removed from the resin, and the resin was washed thoroughly with water and methanol and then was reacted with DTT overnight. The solution was collected for analysis. The resin was subjected to a Kaiser test and was found to be positive for amine.

The same peptide mixture that was used in the libraries was made and used to develop a RP-HPLC method for resolving the peptides (**Figure 5.17**). The resolution on the instrument used was not very good, but the libraries were analyzed with this method anyway as a preliminary determination of the content of each sample (**Figure 5.18** and **5.19**). Some differences between the peaks in the two libraries were apparent; however, the poor resolution and inconsistent retention times made it difficult to draw any solid conclusions from this data about the templating effect of the BSA. However, it was clear from the data that the peptides were immobilized on the resin, which confirmed that the reaction was working.

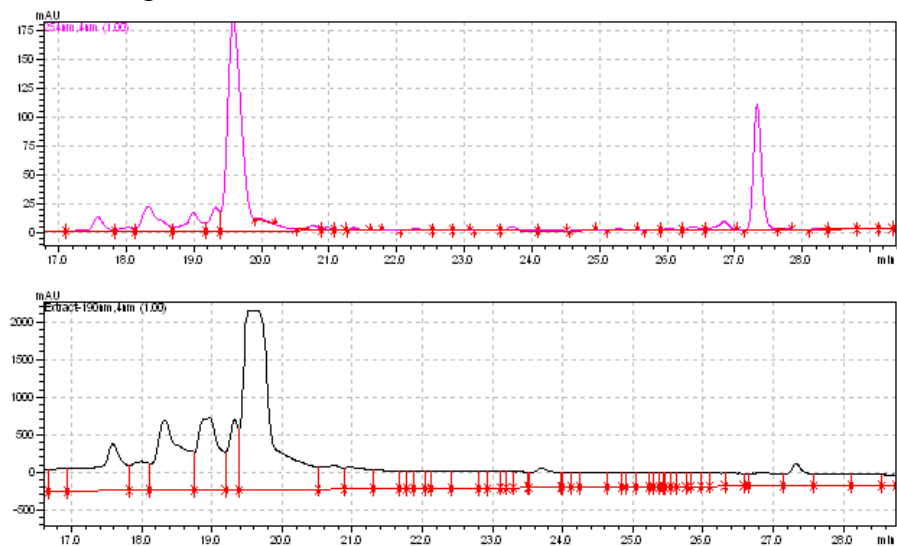


Figure 5.17 Separation of the peptide mixture. Top: $\lambda = 254$ nm, bottom: $\lambda = 190$ nm.

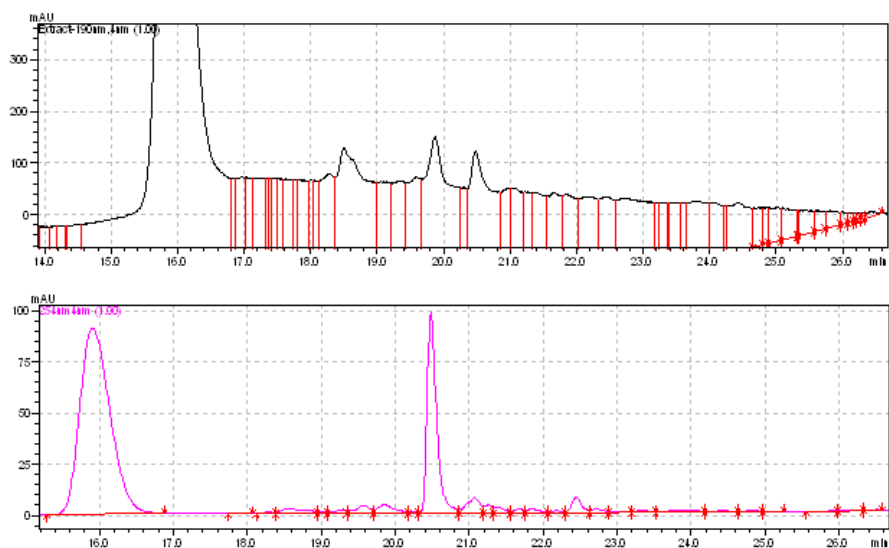


Figure 5.18 Separation of the untemplated library. Top: $\lambda = 190$ nm, bottom: $\lambda = 254$ nm.

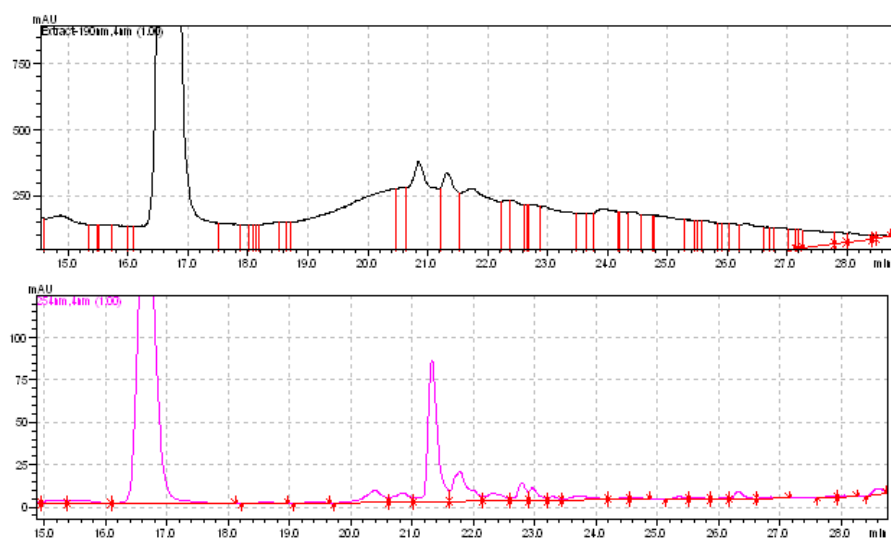


Figure 5.19 Separation of the library templated against BSA. Top: $\lambda = 190$ nm, bottom: $\lambda = 254$ nm.

5.6 CONCLUSIONS

In this work, we developed protocols for chemically modifying proteins to contain either thiols or conjugate acceptors. These functionalities were able to react via thia-Michael chemistry to attach units to the protein. We had the most success using EvaCA for this purpose. Before working with proteins, we studied the reactivity of EvaCA and found that it reacts with an amine and then is rendered less electrophilic so that the second addition is selective for thiol. Thus, EvaCA can be used as a general crosslinker of an amine and a thiol. Furthermore, we found that DTT can be used to remove the amine and thiol from EvaCA, which makes the crosslinking a reversible process. We used EvaCA to reversibly label a sample of myoglobin with a thiol with good reproducibility of the modifications.

Furthermore, using a triamine scaffold, we demonstrated that the thiol addition to EvaCA is a dynamic reaction by observing the scrambling of different thiols on the triamine. Then, we immobilized the triamine on a Tentagel S resin and built a DCL out of this resin, EvaCA, and thiol-containing peptides. This DCL was tested for its templating effect against BSA as a test analyte. Preliminary HPLC analysis of the library after removal of the peptides from the resin by DTT suggests that there is a template effect from the protein.

5.7 FUTURE WORK

Further work in the area of protein labeling could include taking the thiol MEEEEE and transforming the alcohol to an azide, as we did with 2-mercaptoethanol to make 2-azidoethanethiol. Then, this thiol could be tested for addition to LysEva and for its reactivity toward CuAAC conditions with an alkyne. If this process proved to be an efficient way to conjugate a protein to recognition units, then proteins labeled in this way could be used for DCC.

For the resin-bound triamine portion of the work, the DCLs should be analyzed with a more sophisticated LC-MS instrument to improve the resolution and allow for conclusive identification of the peptides. If the apparent template effect is confirmed, then more DCLs could be set up from these peptides to template against different proteins. The data for the DCLs for each protein could be used in conjunction with pattern recognition algorithms like LDA and PCA to differentiate the proteins.

5.8 CONTRIBUTIONS

J. Logan Bachman assisted in the synthesis of the EvaCA derivatives. Ye (Eva) Zhong assisted in the EvaCA reactivity studies. Dr. Jared Shaw and Scott Robotham performed the protein mass spectral analysis.

5.9 EXPERIMENTAL METHODS

5.9.1 General

Unless otherwise indicated, chemicals and reagents were obtained from Sigma Aldrich and used without further purification. The resins, amino acids, and coupling reagents were obtained from NovaBioChem. *N*-succinimidyl *S*-acetylthioacetate (SATA) and 2-anthracene carboxylic acid were obtained from TCI America. NBD-FA was synthesized according to the literature procedure described in Chapter 2.

HPLC-grade solvents were prepared with 0.1% TFA (v/v) and filtered through a 0.2 micron filter. HPLC was performed on a Shimadzu instrument with a preparative C-18 column using a water and acetonitrile mobile phase. LC-MS analysis was performed on an Agilent 1200 Series HPLC with an Agilent 6130 single quadrupole mass spectrometer (ESI and APCI ionization). ¹H NMR spectra were obtained on a 400 MHz Varian.

The fluorescence titration experiments were performed with a PTI fluorimeter using an 814 photomultiplier detection system and a 75W xenon short arc lamp. Absorbance measurements were taken using a Beckmann Coulter DU-800 UV/VIS Spectrophotometer.

5.9.2 Synthesis

5.9.2.1 *N*-succinimidyl *S*-tritylthioacetate (STTA)

S-tritylthioglycolic acid³⁶

Thioglycolic acid (3 mL, 43 mmol, 1 eq) and triphenylmethanol (11.2 g, 43 mmol, 1 eq) were placed in a flask with 30 mL DCM under argon. Acetic acid (23 mL) and boron trifluoride diethyl etherate (7.4 mL, 60 mmol, 1.4 eq) were added to the reaction. The reaction was stirred at room temperature for 1.5 hr. The solvent was removed under reduced pressure, and the product was precipitated with water. The solid was washed with water, acetonitrile, and cold diethyl ether. The product required no further purification (9.9 g, 69% yield). ¹H-NMR (400 MHz, CDCl₃, ppm): δ 3.05 (s, 2H, CH₂), 7.30 (m, 9H, Ar-H), 7.42 (m, 6H, Ar-H). LRMS (ESI): 333.2 (M-1).

N-succinimidyl *S*-tritylthioacetate

S-tritylglycolic acid (6.7 g, 20 mmol, 1 eq) and *N*-hydroxysuccinimide (2.3 g, 20 mmol, 1 eq) were dissolved in 35 mL dry DCM. Dicyclohexylcarbodiimide (4.2 g, 20 mmol, 1 eq) was added, and the reaction was stirred for 18 hr under argon. The reaction mixture was filtered to remove solids, and the solvent was removed under reduced pressure. The crude product was recrystallized from isopropanol. The product was a white crystalline solid (3.0 g, 35% yield). ¹H-NMR (400 MHz, CDCl₃, ppm): δ 2.82 (s, 4H, CH₂), 3.18 (s, 2H, CH₂), 7.23 (m, 3H, Ar-H), 7.31 (m, 6H, Ar-H), 7.42 (m, 6H, Ar-H).

¹³C-NMR (400 MHz, CDCl₃, ppm): δ 25.53 (CH₂), 31.40 (CH₂), 68.01 (C), 127.13 (Ar-C), 128.23 (Ar-C), 129.48 (Ar-C), 143.52 (Ar-C), 165.05 (C=O), 168.56 (C=O).

5.9.2.2 1-(4-azidophenyl)prop-2-en-1-one (attempted)

To 4-azidobenzoic acid (1 g, 6 mmol, 1 eq) was added 27 mL of thionyl chloride. The reaction was refluxed for 2.5 hr, and the solvent was removed under reduced pressure. To this residue was added 10 mL of dry THF, and the flask was placed under argon. Tributyl vinyl tin (1.75 mL, 6.0 mmol, 1 eq) and a catalytic amount of tetrakis(triphenylphosphine)palladium were added to the reaction. The reaction was heated to reflux for 3 hr. Then ether was added to the reaction, and solids were removed by filtration. The solution was washed repeatedly with a saturated KF solution (aq) to remove tin species. The solution was washed with brine, dried with MgSO₄, and the solvent was removed under reduced pressure. The crude product was analyzed by LC-MS but did not contain the desired product.

5.9.2.3 PyrCA via 2-(bromomethyl)pyridine (attempted)

Meldrum's acid (200 mg, 1.4 mmol, 1 eq), carbon disulfide (0.13 mL, 2.2 mmol, 1.2 eq), and trimethylamine (0.4 mL, 2.8 mmol, 2 eq) were placed in a flask with 5 mL acetonitrile. The reaction was stirred for 1 hr. Meanwhile the 2-(bromomethyl)pyridine hydrobromide was dissolved in 10 mL of a saturated sodium bicarbonate solution (aq), which was stirred briefly. The solution was extracted three times with ether, and the solvent was removed under reduced pressure. The above reaction was cooled in an ice bath and the 2-(bromomethyl)pyridine was added. The reaction was stirred overnight at room temperature. The solvent was removed under reduced pressure, yielding a black residue. Water was added, and the solution was extracted with chloroform three times. The organic layer was washed with brine and dried with MgSO₄. The solvent was removed under

reduced pressure to yield a reddish-black residue. ¹H-NMR and LC-MS analysis did not indicate the presence of any product. Attempts to separate the species present on alumina were unsuccessful.

5.9.2.4 *EvaCA*⁴⁰

To a solution of Meldrum's acid (10 g, 0.07 mol) in DMSO (30 mL) was added triethylamine (14.2 g, 0.14 mol) and carbon disulfide (5.3 g, 0.07 mol). The reaction was stirred for 1 hr at room temperature under argon. Then methyl iodide (19.9 g, 0.14 mol) was added dropwise. The reaction was stirred for 18 hr. Then ice was added to precipitate the product. The crude product was recrystallized from methanol (9.5 g, 55% yield). ¹H NMR (400 MHz, CD₃CN, ppm): δ 1.71 (s, 3H, CH₃), 2.62 (s, 3H CH₃).

5.9.2.5 *BenzCA (attempted)*

Meldrum's acid (0.5 g, 3.5 mmol, 1 eq), trimethylamine (1 mL, 7.3 mmol, 2.1 eq), and carbon disulfide (0.34 mL, 5.6 mmol, 1.6 eq) were placed in a flask under argon with 10 mL of acetonitrile. The reaction was stirred for 1 hr, and the benzylbromide (0.86 mL, 6.9 mmol, 2.0 eq) was added. The reaction was stirred for 18 hr. A white precipitate formed and was collected by filtration. The crude product was analyzed by NMR but not purified. ¹H NMR (400 MHz, CDCl₃, ppm): δ 0.63 (s, 3H, CH₃), 3.45 (s, 2H, CH₂), 7.21 (m, 5H, Ar-H). LRMS (ESI): 308.8 (M-1).

5.9.2.6 *QuinCA (attempted)*

*2-(bromomethyl)quinolone*⁴¹

To 100 mL of water was added 2-(chloromethyl)quinoline hydrochloride (10 g, 46.7 mmol, 1 eq). To this solution was added 50 mL 1 M NaOH (aq). A white precipitate formed and was collected. The solid was recrystallized from hexanes. The purified solid was dissolved in 200 mL THF, and LiBr (40 g, 46.7 mmol, 10 eq) was added. The reaction

was heated to reflux for 18 hr. The reaction was concentrated under reduced pressure. Water was added, and the solution was extracted twice with ether. The ether was washed with brine and removed under reduced pressure. The produce was an orange solid (4.0 g, 39% yield). ¹H NMR (400 MHz, CDCl₃, ppm): δ 4.65 (s, 2H, CH₂), 7.52 (m, 2H, Ar-H), 7.66 (t, 1H, Ar-H), 7.76 (d, 1H, Ar-H), 7.99 (d, 1H, Ar-H), 8.10 (d, 1H, Ar-H).

QuinCA (attempted)

Meldrum's acid (650 mg, 4.5 mmol, 1 eq), carbon disulfide (0.33 mL, 5.6 mmol, 1.6 eq), and trimethylamine (1.25 mL, 9.0 mmol, 2.1 eq) were placed in a flask with 10 mL acetonitrile. The reaction was stirred for 1 hr. The above reaction was cooled in an ice bath and the 2-(bromomethyl)quinoline was added. The reaction was stirred overnight at room temperature. A yellow precipitate had formed and was collected by filtration. ¹H-NMR and LC-MS analysis did not indicate the presence of any product. Attempts to separate the components of the reaction mixture on alumina were unsuccessful.

5.9.2.7 PyrCA via 2-pyridinemethanethiol (attempted)

*2-pyridinemethanethiol*⁴¹

Thiourea (1.28 g, 15.2 mmol, 1.1 eq) and 2-(chloromethyl)pyridine hydrochloride (2.5 g, 15.0 mol, 1 eq) were dissolved in 20 mL of water. The reaction was heated to 85 °C for 1 hr. The reaction mixture was cooled in an ice bath and sparged with nitrogen for 2 hr. With the flow of nitrogen continuing, NaOH was added (1.8 g, 45.0 mmol, 3 eq), and the reaction was stirred overnight at room temperature. The reaction mixture was washed twice with *t*-butylmethylether. The aqueous layer was cooled in an ice bath, and the pH was adjusted to 7 with 1 M HCl. The solution was extracted three times with DCM. The organic layers were dried with MgSO₄, and the solvent was removed under reduced pressure. The product was a white solid (119 mg, 6.3% yield). ¹H NMR (400 MHz, CDCl₃,

ppm): δ 2.00 (s, 1H, SH), 3.84 (s, 2H, CH₂), 7.16 (m, 1H, Ar-H), 7.34 (m, 1H, Ar-H), 7.65 (m, 1H, Ar-H), 8.53 (m, 1H, Ar-H).

PyrCA via 2-pyridinemethanethiol (attempted)

EvaCA (118 mg, 0.48 mmol, 1 eq) and 2-pyridinemethanethiol (119 mg, 0.95 mmol, 2 eq) were dissolved in acetonitrile. DIPEA (165 μ L, 0.95 mmol, 2 eq) was added, and the reaction was stirred under nitrogen for 24 hr. Water was added to the reaction, and the solution was extracted twice with DCM. The DCM layer was washed with brine and dried with MgSO₄. The solvent was removed under reduced pressure. The crude material was analyzed by LC-MS and ¹H NMR. The desired product was not observed.

5.9.2.8 2-Azidoethanethiol

*Bis-(2-bromoethyldisulfide)*⁴⁵

Concentrated H₂SO₄ (10 mL) was added slowly to a solution of 48% HBr in an ice bath with stirring. Then *bis*-(2-hydroxyethyl)disulfide (0.3 mL, 2.45 mmol) was added slowly to the acid mixture. The reaction stirred overnight at room temperature. DCM was added to the reaction and washed with water and 10% Na₂CO₃ (aq). The DCM layer was dried with MgSO₄ and removed under reduced pressure. The product was a liquid (0.68 g, 99% yield). ¹H NMR (400 MHz, CDCl₃, ppm): δ 3.08 (t, 4H, CH₂), 3.59 (t, 4H, CH₂).

*Bis-(2-azidoethyl)disulfide*⁴⁵

The *bis*-(2-bromoethyl)disulfide (0.23 g, 0.82 mmol, 1 eq) and sodium azide (0.26 g, 4.0 mmol, 4.9 eq) were dissolved in 10 mL DMF. The reaction was heated to 80 °C for 18 hr. Then water was added, and the product was extracted three times with ether. The organic layer was dried with MgSO₄ and removed under reduced pressure (0.14 g, 82% yield). ¹H NMR (400 MHz, CDCl₃, ppm): δ 2.87 (t, 4H, CH₂), 3.60 (t, 4H, CH₂).

*2-azidoethanethiol*³⁸

Bis-(2-azidoethyl)disulfide (0.15 g, 0.75 mmol, 1 eq), dithiothreitol (0.34 g, 2.2 mmol, 3 eq), and trimethylamine (0.6 mL, 4.3 mmol, 6 eq) were dissolved in THF under argon. The reaction was stirred for 2 hr at room temperature and was monitored by TLC with an iodine stain. Then 1 M HCl (aq) was added to acidify the solution. The resultant solution was extracted twice with DCM. The DCM was washed with 1 M HCl and water until the DTT spot on the TLC disappeared. The DCM was allowed to evaporate. The product was a volatile yellow liquid (74 mg, 49% yield). ¹H NMR (400 MHz, CDCl₃, ppm): δ 1.58 (t, 1H, SH), 2.70 (q, 2H, CH₂), 3.46 (t, 2H, CH₂).

5.9.2.9 Triamine compound

*1,3,5-(azidomethyl)benzene*⁴⁸

To a flask was added 1,3,5-(bromomethyl)benzene (1 g, 2.8 mmol, 1 eq) in 5 mL DMF. Sodium azide (1.1 g, 16.8 mmol, 6 eq) was added in portions over several minutes, and then the reaction was stirred at room temperature for 24 hr. Water was added to the reaction, and the solution was extracted twice with DCM. The DCM was washed with brine, dried with MgSO₄, and removed under reduced pressure. The product was a brown oil (459 mg, 67% yield). ¹H NMR (400 MHz, CDCl₃, ppm): δ 4.38 (s, 6H, CH₂), 7.23 (s, 3H, Ar-H).

1,3,5-(aminomethyl)benzene^{47,48}

To a flask was added 1,3,5-(azidomethyl)benzene (0.59 g, 2.3 mmol, 1 eq) and triphenylphosphine (3.8 g, 14.5 mmol, 6.3 eq) in 10:1 THF/water. The reaction was stirred at room temperature for 48 hr. The solvent was removed under reduced pressure. The residue was dissolved in a small amount of water and filtered to remove solids. The water was removed by lyophilization to yield a yellow hygroscopic solid (0.25 g, 63% yield). ¹H NMR (400 MHz, d-DMSO, ppm): δ 3.63 (s, 6H, CH₂), 7.06 (s, 3H, Ar-H).

5.9.3 Peptide Synthesis

The peptides were synthesized on an automated peptide synthesizer (Protein Technologies) using Fmoc chemistry and the Wang resin pre-loaded with Fmoc-12-aminododecanoic acid. For a 200 μmol -scale reaction, the Wang resin was swelled in DMF and then deprotected with 20% piperidine in DMF twice. In between the resin was washed with DMF and DCM. Each amino acid was dissolved in DMF (100 mM) and introduced to the resin after the deprotection step. PyBOP (300 mM) and DIPEA (1.2 M) in DMF were added along with the amino acid to perform each coupling (twice). The final residue was deprotected at the end. The resin was then removed from the synthesizer.

The resin was washed with DMF, DCM, MeOH, and glacial acetic acid and then dried overnight on the high vacuum line. The peptides were cleaved from the resin with 95% TFA, 2.5% TIPS, and 1% H_2O for 4 hr. The TFA solution was filtered from the resin, and the TFA was removed. To the residue was added diethyl ether to precipitate the peptides.

5.9.3.1 Model Peptide

Before IAHRFKDLGE was cleaved from the resin, the N-terminus was acetylated by placing the resin in 5 mL of 50% acetic anhydride in DMF and 100 μL pyridine. The reaction was allowed to proceed for 3 hr with shaking to mix. Then the resin was washed and cleaved from the resin. The crude peptide (100 μmol) was dissolved in 3 mL acetonitrile and 2 mL water. DIPEA (120 μL , 600 μmol , 6 eq) and EvaCA (50 mg, 200 μL , 2 eq) were added, and the reaction was stirred in an open vial for 24hr. The presence of the product was confirmed by LC-MS. The solvent was removed under reduced pressure.

5.9.3.2 Purification

The crude peptides were dissolved in water/DMSO (no more than 50% DMSO) and purified by HPLC. Fractions containing the desired peptide were verified by LC-MS, combined, and evaporated:

- acIAHRFK(Eva)DLGE and acAHRFK(Eva)DLGE: 34.8 mg (25% yield); LRMS (ESI): 657.6, 714.1 ((M+1)/2), 1313.8, 1426.8 (M+1); HRMS (ESI): expected 1314.58, found 1314.58, expected 1427.66, found 1427.66 (M+1)
- VVKLKC: 59.3 mg (58% yield). LRMS (ESI): 689.4 (M+1); HRMS (ESI): expected 689.44, found 689.44 (M+1)
- KGGRKC: 61.9 mg (56% yield); LRMS (ESI): 648.4 (M+1); HRMS (ESI): expected 648.36, found 648.36 (M+1)
- KRGSKC: 48.9 mg (43% yield); LRMS (ESI): 678.4 (M+1); HRMS (ESI): expected 678.37, found 678.37 (M+1)
- KRSSKC: 64.3 mg (55% yield); LRMS (ESI): 708.4 (M+1); HRMS (ESI): expected 708.38, found 708.38 (M+1)
- KAGLKC: 52.9 mg (55% yield); LRMS (ESI): 619.4 (M+1); HRMS (ESI): expected 310.18, found 310.18 ((M+1)/2)
- KYPYKC: 56.6 mg (50% yield); LRMS (ESI): 801.4 (M+1); HRMS (ESI): expected 401.20, found 401.20 ((M+1)/2)
- PRGDKC: 61.4 mg (60% yield); LRMS (ESI): 675.3 (M+1); HRMS (ESI): expected 338.17, found 338.17 ((M+1)/2)
- KDGRKC: 65.0 mg (56% yield); LRMS (ESI): 706.3 (M+1); HRMS (ESI): expected 706.37, found 706.37 (M+1)
- KGARCC: 40.7 mg (42% yield); LRMS (ESI): 637.3 (M+1); HRMS (ESI): expected 637.29, found 637.29 (M+1)

5.9.4 Small Molecule Test Reactions

5.9.4.1 *LysEvaCys*

EvaCA (5 mg, 20 μ mol, 1 eq) was dissolved in 100 μ L acetonitrile. The α -Boc-lysine (5 mg, 20 μ mol, 1 eq) was dissolved in 1 mL of 50 mM phosphate buffer, H₂O, 10 mM EDTA, pH 7.5. The EvaCA solution was added dropwise to the lysine solution. The reaction was stirred for 24 hr at room temperature. LC-MS analysis indicated that the reaction had gone to completion (**Figure 5.20** and **5.21**). LRMS (ESI): 445.2 (M-1). Then NAC (3.6 mg, 22 μ mol, 1.1 eq) was added to the solution. The reaction was gently sparged with nitrogen and monitored by LC-MS over eight days, at which point the reaction had mostly completely (**Figure 5.22**). LRMS (ESI): 560.2 (M-1).

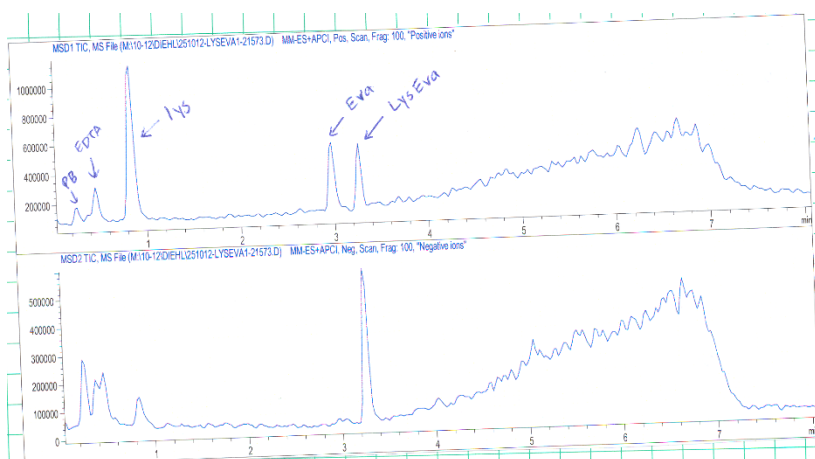


Figure 5.20 LC-MS analysis of *LysEva* reaction after 3 hr. Top: TIC, positive mode, bottom: TIC, negative mode.

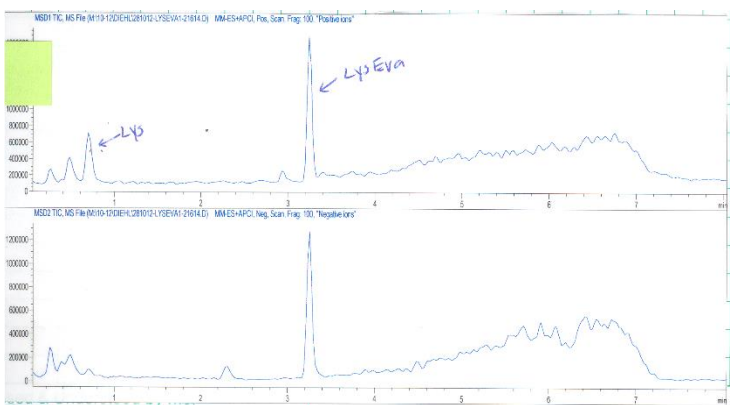


Figure 5.21 LC-MS analysis of *LysEva* reaction after 72 hr. Top: TIC, positive mode, bottom: TIC, negative mode.

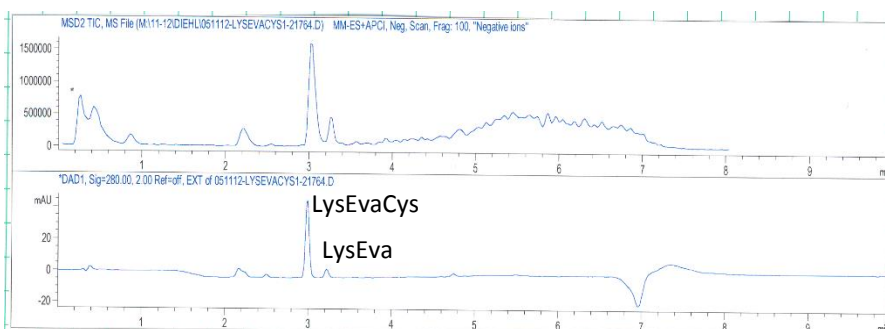


Figure 5.22 LC-MS analysis of *LysEvaCys* reaction after 8 days. Top: TIC, negative mode, bottom: UV trace, $\lambda = 280$ nm.

5.9.4.2 *LysEvaCys* + DTT

LysEvaCys was prepared as described. Then three equivalents of dithiothreitol was added to the solution, and the reaction was heated to 40 °C for 18 hr. LC-MS analysis showed no *LysEvaCys* was present.

5.9.4.3 *CysEvaCys*

EvaCA (6.0 mg, 24.2 μ mol, 1 eq) was dissolved in 1 mL acetonitrile. *NAC* (8.3 mg, 50.7 μ mol, 2.1 eq) was dissolved in 5 mL of 50 mM phosphate buffer, H₂O, 10 mM EDTA, pH 7.5. The two solutions were combined. The reaction was gently sparged with nitrogen and reacted overnight. LC-MS analysis showed *CysEvaCys* (**Figure 5.23**).

LRMS (ESI): 477.2 (M-1). Then α -Boc-lysine (6.0 mg, 24.2 μ mol, 1 eq) was added, and the reaction was stirred for 3 days. It was monitored by LC-MS (**Figure 5.24**). LRMS (ESI): 560.2 (M-1).

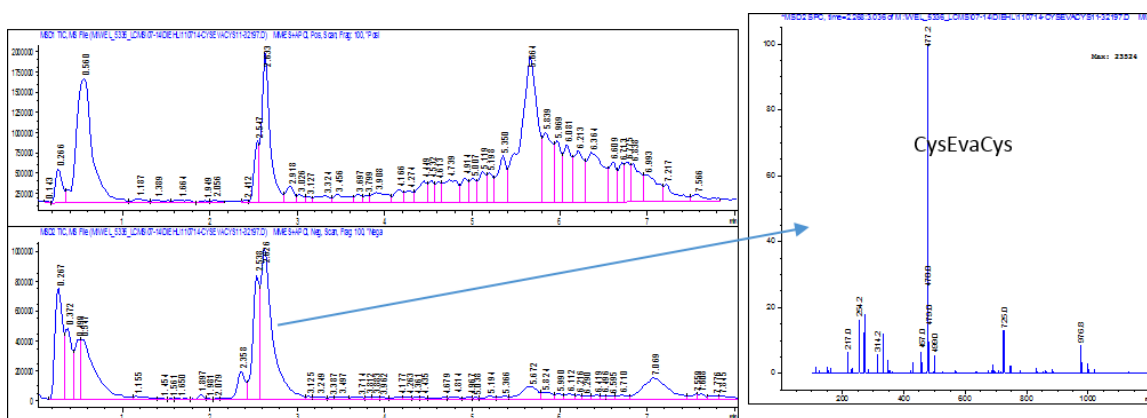


Figure 5.23 LC-MS analysis of *CysEvaCys* after 18 hrs. Top: TIC, positive mode, bottom: TIC, negative mode.

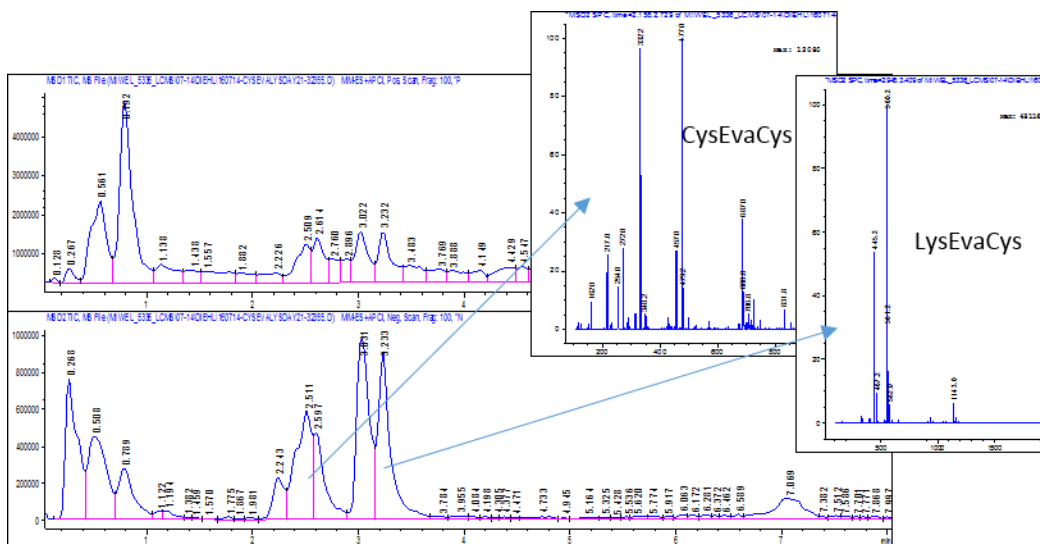


Figure 5.24 LC-MS analysis of *LysEvaCys* after 3 days. Top: TIC, positive mode, bottom: TIC, negative mode.

5.9.4.4 *LysEvaBME*

LysEva (19 μmol) was prepared as described. Then 2-mercaptoethanol (BME, 1900 μmol , 100 eq) was added, and the reaction was stirred for 5 hr. LRMS (ESI): 475.2 (M-1).

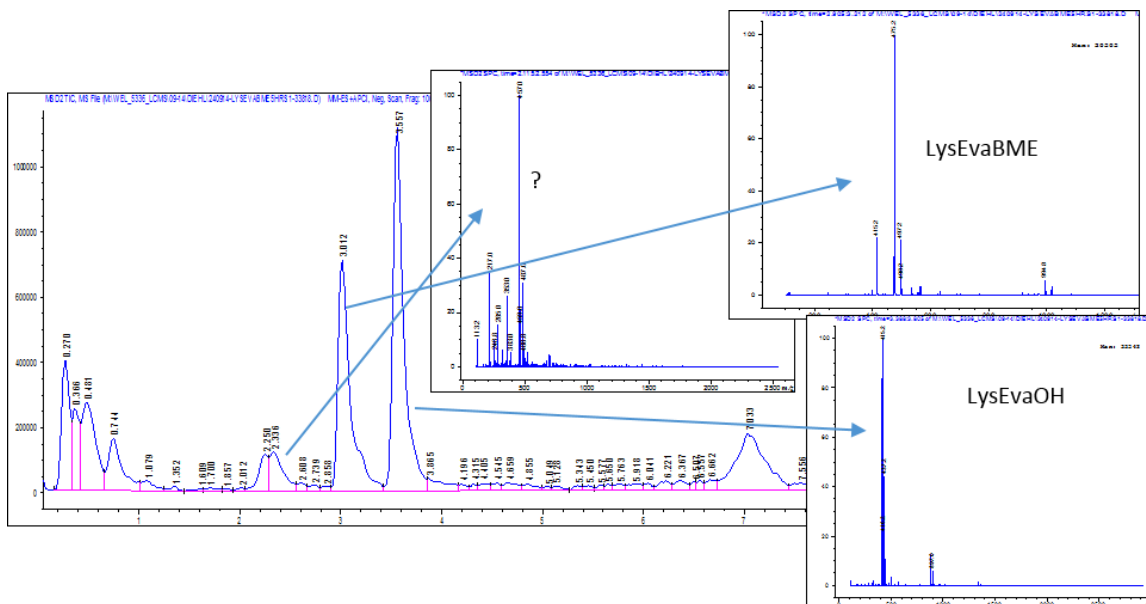


Figure 5.25 LC-MS analysis of *LysEvaBME* after 5 hr. TIC, negative mode.

5.9.4.5 *MPEvaVVKLKC*

The peptides acIAHRFK(Eva)DLGE/acAHRFK(Eva)DLGE (1.9 mg, 1.15 μmol , 1 eq) and VVKLKC (2.7 mg, 2.62 μmol , 2.3 eq) were dissolved in 10 mM carbonate buffer, pH 9.2. The buffer capacity was exceeded, so 1 M NaOH (aq) was added to adjust the pH to 7.5. The total reaction volume at the end of the adjustment was 1 mL. The reaction was allowed to stir at room temperature for two weeks in an open Eppendorf tube. The reaction was periodically replenished with H_2O as it evaporated. After this time, the solution was filtered and analyzed by LC-MS. The desired products were present. A sample was

submitted for MS analysis by our collaborators. LRMS (ESI): 1034.7 ((M+1)/2), 978.2 ((M+1)/2), 690.3 ((M+1)/3), 652.6 ((M+1)/3), 518.0 ((M+1)/4), 489.7 ((M+1)/4).

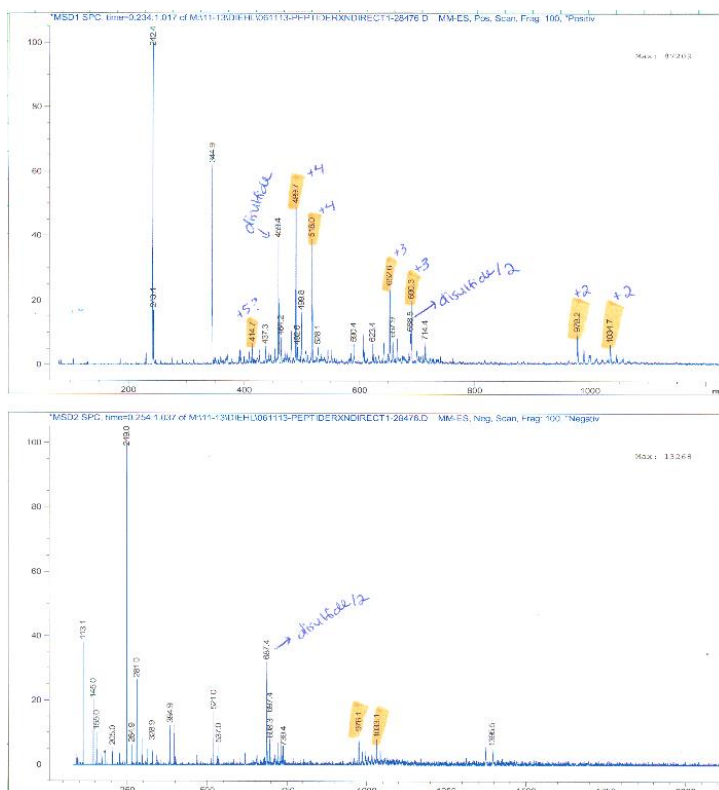


Figure 5.26 LC-MS analysis of MPEvaVVKLKC after 2 weeks. Top: TIC, positive mode, bottom: TIC, negative mode.

5.9.4.6 MPEvaVVKLKC + DTT

To the solution of MPEvaVVKLKC (1.15 μmol , 1 eq) was added DTT (1 mg, 6.5 μmol , 5.7 eq). The reaction was stirred at room temperature in an open container for three days. LC-MS analysis indicated that the reaction had occurred (**Figure 5.27**). LRMS (ESI): VVKLKC- 230.7 ((M+1)/3), 345.2 ((M+1)/2), 689.4 (M+1); MP- 372.3 ((M+1)/3), 410.0 ((M+1)/3).

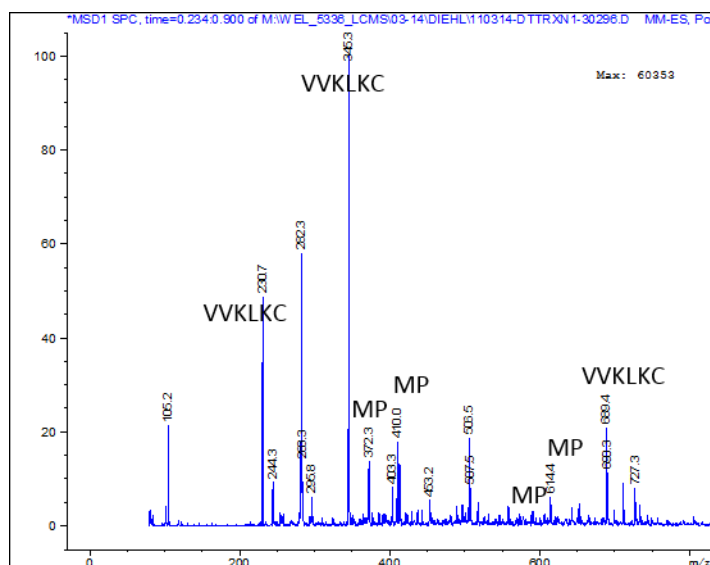


Figure 5.27 MS analysis of MPEvaVVKLKC + DTT after 3 days.

5.9.4.7 LysEvaLys and LysEvaLys + DTT

EvaCA (5.8 mg, 23.5 μmol , 1 eq) was dissolved in 300 μL acetonitrile, and α -Boc-lysine (33.7 mg, 137 μmol , 5.8 eq) was dissolved in 2.5 mL of 50 mM phosphate buffer, H_2O , 10 mM EDTA, pH 7.2. The EvaCA solution was added to the lysine solution. The reaction was heated to 45 $^\circ\text{C}$ for three days. LC-MS analysis indicated some of the LysEvaLys product had formed (**Figure 5.28**). LRMS (ESI): 643.2 (M+1). Then dithiothreitol (42 mg, 272 μmol , 11.6 eq) was added to the reaction, and the heating was continued for 24 hr. LC-MS analysis indicated that the LysEva reacted with DTT, but the LysEvaLys did not react with DTT (**Figure 5.29**).

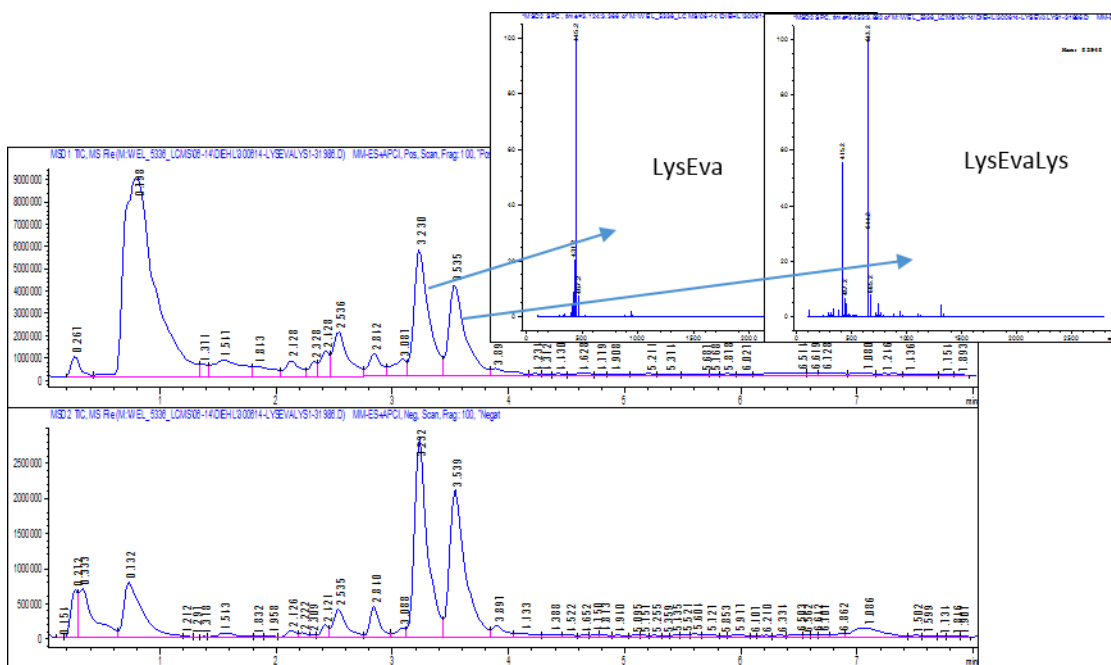


Figure 5.28 LC-MS analysis of *LysEvaLys* after 3 days. Top: TIC, positive mode, bottom: TIC, negative mode.

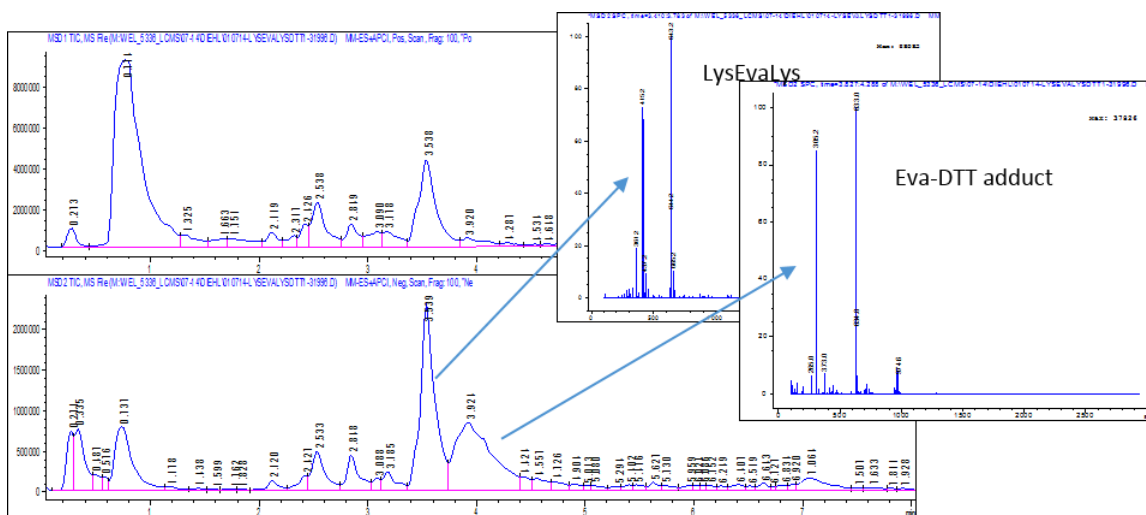


Figure 5.29 LC-MS analysis of *LysEvaLys* + DTT after 24 hr. Top: TIC, positive mode, bottom: TIC, negative mode.

5.9.4.8 LysEvaAz and CuAAC reaction

EvaCA (4.7 mg, 19 μmol , 1 eq) was dissolved in 0.4 mL acetonitrile and added to a 2 mL solution of α -Boc-lysine (6.4 mg, 26 μmol , 1.4 eq) in 50 mM phosphate buffer, H_2O , 10 mM EDTA, pH 7.5. The reaction was gently sparged with nitrogen for 48 hr. Then 2-azidoethanethiol (5.9 mg, 57 μmol , 3 eq) was dissolved in 1 mL methanol and added to the reaction. The reaction was stirred for 18 hr in an open vial, and then analyzed by LC-MS to confirm LysEvaAz (**Figure 5.30**). LRMS (ESI): 500.2 (M-1). Then, 3-butyn-1-ol (73 μL , 970 μmol), $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ (36 mg, 144 μmol , 15%), and sodium ascorbate (38 mg, 192 μmol , 20%) were added. The reaction as stirred for 24 hr, and then a sample was taken and analyzed by LC-MS (**Figure 5.31**). The reaction did not yield the desired product.

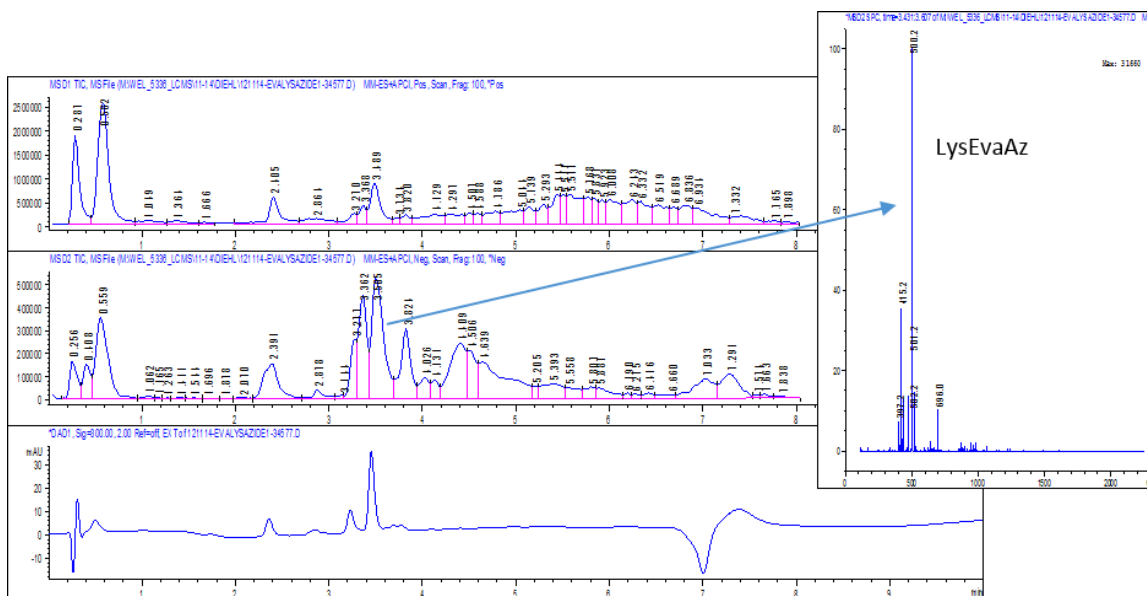


Figure 5.30 LC-MS analysis of LysEvaAz after 18 hr. Top: TIC, positive mode, middle: TIC, negative mode, bottom: UV trace $\lambda = 300$ nm.

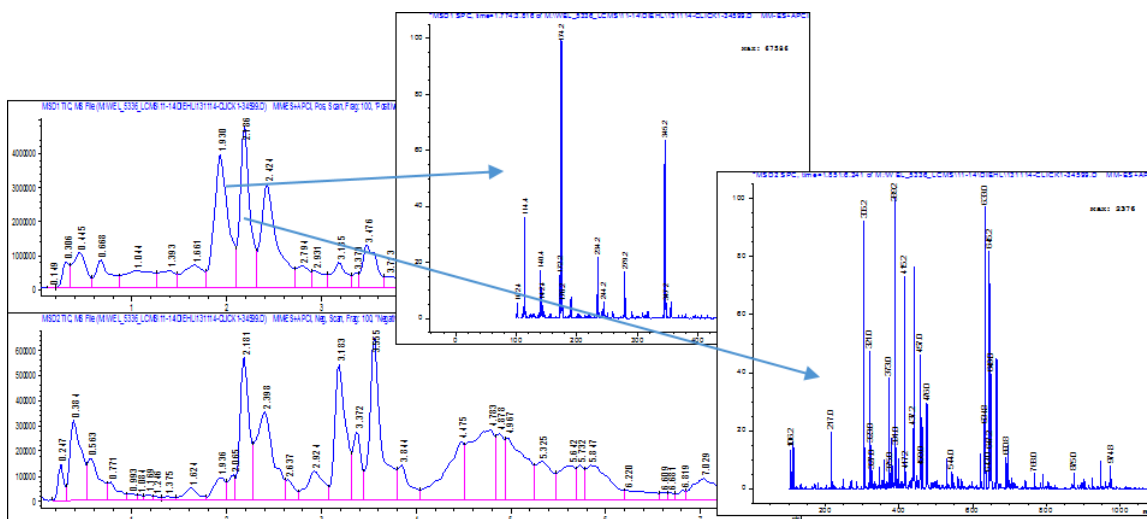


Figure 5.31 LC-MS analysis of Click reaction with *LysEvaAz* and 3-butyn-1-ol after 24 hr. Top: TIC, positive mode, bottom: TIC, negative mode.

5.9.4.9 *LysEvaMEEEE* and *LysEvaMEEEE* + DTT

LysEva was prepared as described (38 μmol). To *LysEva* was added MEEEE (22 μL , 115 μmol , 3 eq), and the reaction was stirred in an open container for 18 hr. The reaction was analyzed by LC-MS to confirm *LysEvaMEEEE* (**Figure 5.32**). LRMS (ESI): 607.2 (M-1). Then, dithiothreitol (53 mg, 344 μmol , 9 eq) was added, and the reaction was stirred for 18 hr. LC-MS analysis showed MEEEE and DTT-Eva adduct (**Figure 5.33**). LRMS (ESI): 305.2 (M-1).

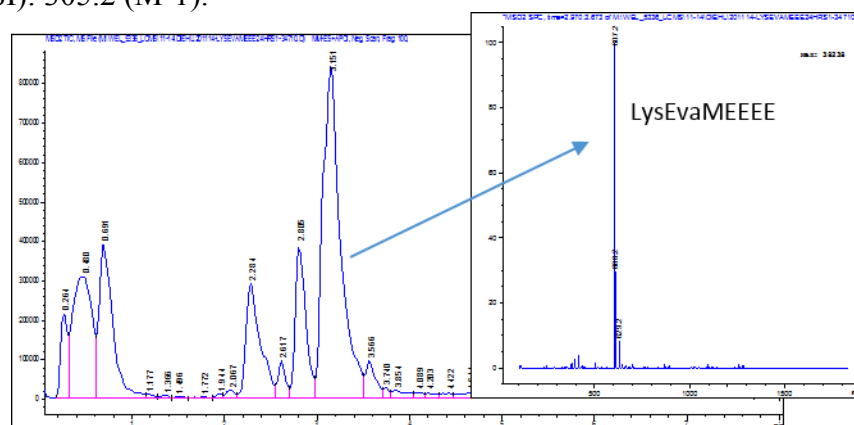


Figure 5.32 LC-MS analysis of *LysEvaMEEEE* reaction after 18 hr. TIC negative mode.

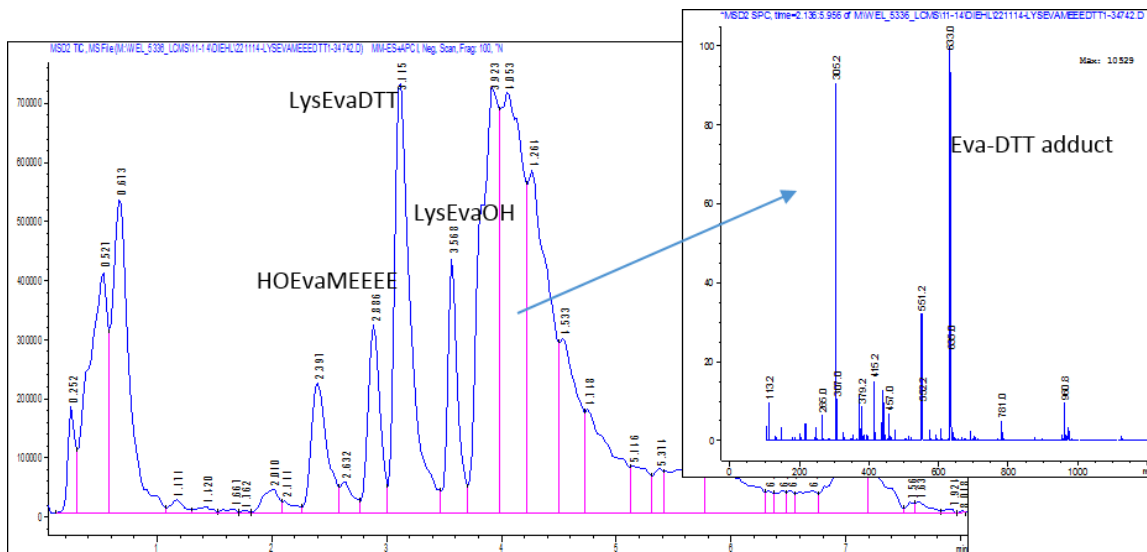


Figure 5.33 LC-MS analysis of LysEvaMEEEE + DTT after 18 hr. TIC negative mode.

5.9.4.10 MPEvaMEEEE

The peptides acIAHRFK(Eva)DLGE/acAHRFK(Eva)DLGE (1.29 mg, 0.78 μ mol, 1 eq) were dissolved in 250 μ L of water. The pH of the solution was adjusted to 7 with 10 mM NaOH (aq), resulting in a total volume of 1 mL. MEEEE (2 μ L, 10.4 μ M, 13.3 eq) was added to the solution, and the reaction was stirred for 20 hr in an open vial. LC-MS analysis showed the desired product (**Figure 5.34**). LRMS (ESI): 738.9 and 795.4 ((M+1)/2).

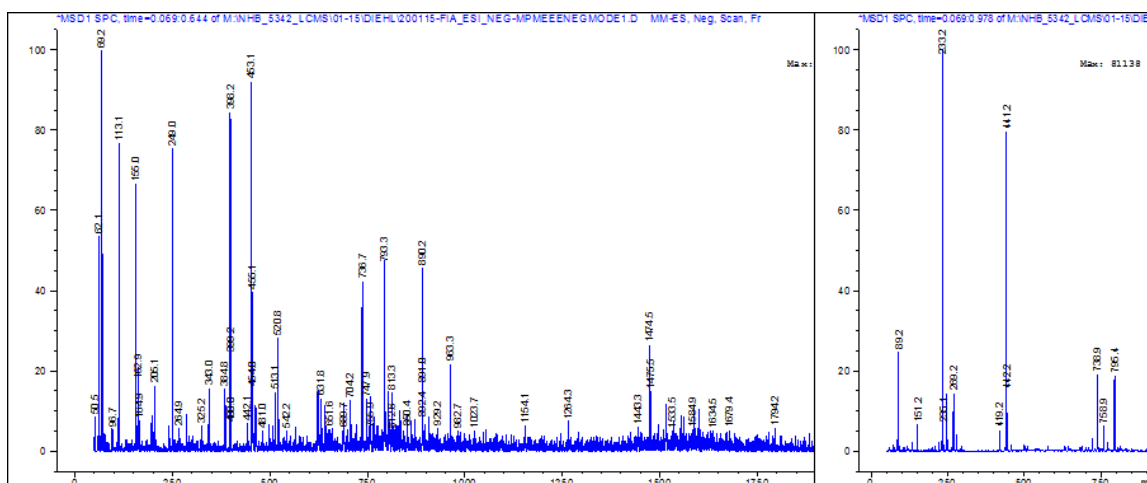


Figure 5.34 MS analysis of MPEvaMEEEE after 20 hr. Left: negative mode, Right: positive mode.

5.9.4.11 DTT-Eva Adduct

EvaCA (3.37 mg, 1.5 μmol , 1 eq) and dithiothreitol (2.27 mg, 1.5 μmol , 1 eq) were dissolved in 1.5 mL CD_3CN . The reaction was stirred in an open vial for 24 hr. ^1H NMR (400 MHz, CD_3CN , ppm): δ 1.66 (s, 6H, CH_3), 2.82 (m, 2H, CH_2), 3.41 (m, 2H, CH_2), 3.59 (m, 1H, CH), 3.81 (s, 1H, OH), 3.89 (s, 1H, OH), 5.36 (m, 1H, CH). ^{13}C NMR (400 MHz, CD_3CN , ppm): δ 26.01 (CH_3), 27.33 ($\text{CH}_2\text{-SH}$), 30.69 ($\text{CH}_2\text{-SH}$), 72.45 ($\text{CH}_2\text{-OH}$), 91.29 (CH-OH), 103.45 (C-O, $-\text{CH}_3$, $-\text{CH}_3$), 103.45 (C=C, $-\text{C}$, $-\text{C}$), 158.13 (C=O), 163.95 (C=O), not observed (C=C, $-\text{S}$, $-\text{O}$). LRMS (ESI): 305.2 (M-1).

5.9.4.12 EDT-Eva Adduct

EvaCA (6.7 mg, 27 μmol) and ethanedithiol (2 μL , 24 μmol) were dissolved in 0.75 mL CD_3CN . The reaction was stirred in an open vial overnight. ^{13}C NMR (400 MHz, CD_3CN , ppm): 20.73 ($\text{CH}_3\text{-C}$), 25.93 ($\text{CH}_3\text{-S}$), 26.12 ($\text{CH}_2\text{-SH}$), 38.38 ($\text{CH}_2\text{-S}$), 102.98 (C=C, $-\text{C}$, $-\text{C}$), 103.67 (C-O, $-\text{CH}_3$, $-\text{CH}_3$), 159.74, 161.30 (C=O), not observed (C=C, $-\text{S}$, $-\text{S}$).

5.9.4.13 PE_3C_3

The triamine (6.35 mg, 38 μ mol, 1 eq) was dissolved in 4 mL of 0.1 M phosphate buffer, 0.15 M NaCl, H₂O, pH 7.2. EvaCA (29.5 mg, 119 μ mol, 3 eq) was dissolved in 2 mL acetonitrile. The solutions were combined and stirred for 36 hr. The solution became cloudy, so MeCN was added to dissolve the components. NAC (39 mg, 239 μ mol, 6.3 eq) was added, and the reaction was monitored over four days with LC-MS (**Figure 5.35** and **5.36**). LRMS (ESI): PE_3 764.0 (M-1), PE_3C 879.0 (M-1), PE_3C_2 994.0 (M-1), PE_3C_3 1109.0 (M-1).

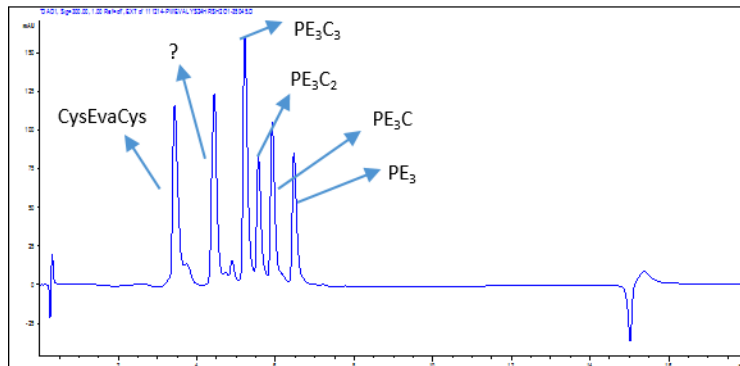


Figure 5.35 LC-MS analysis of PE_3 reaction with NAC after 48 hr. UV trace $\lambda = 300$ nm.

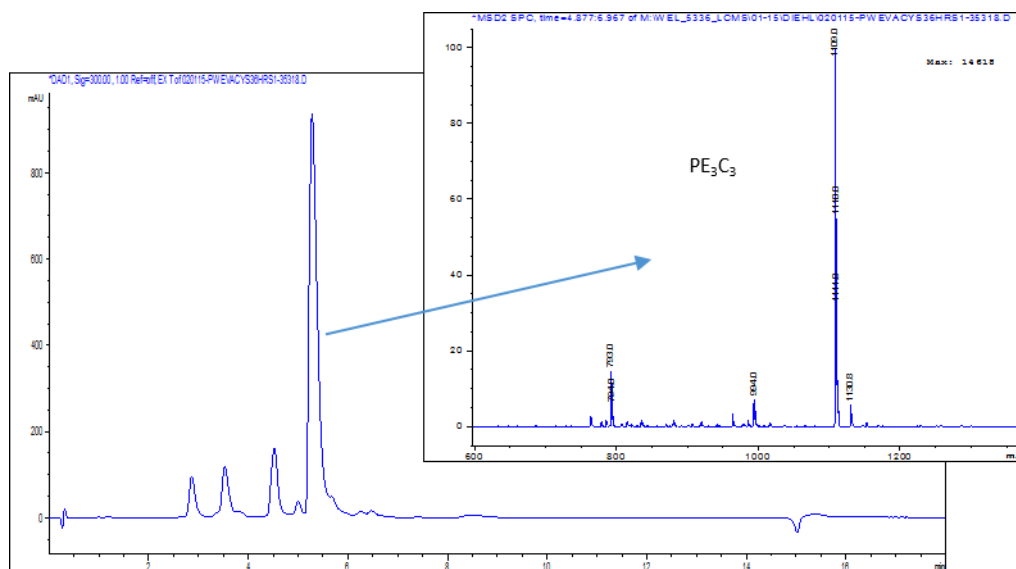


Figure 5.36 LC-MS analysis of PE_3 reaction with NAC after 4 days. UV trace $\lambda = 300$ nm.

5.9.4.14 $PE_3C_3 + DTT$

PE_3C_3 was prepared as described (41 μ mol) in 36 hr by adding more MeCN to dissolved PE_3 and gently sparging with nitrogen. Then, DTT (63 mg, 408 μ mol, 10 eq) was added and reacted for three days (**Figure 5.37**). LRMS (ESI): 305.2 (M-1).

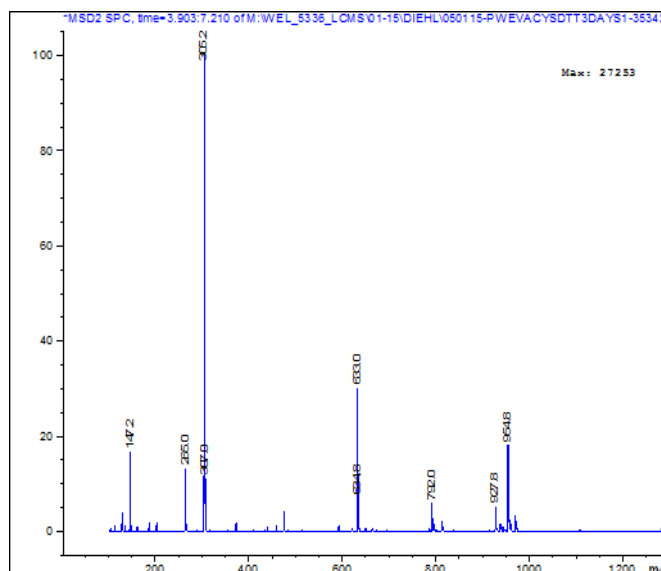


Figure 5.37 MS analysis of PE_3C_3 reaction with DTT after 3 days. Negative mode.

5.9.4.15 PE_3C_3 + 3-Mercaptopropionic Acid

PE_3C_3 was prepared as described (16.3 μmol). Then, 3-mercaptopropionic acid (8.7 μL , 100 μmol , 6 eq) was added, and the reaction was monitored by LC-MS for ten days (**Figure 5.38**). LRMS (ESI): PE_3 764.0 (M-1), PE_3C 879.0 (M-1), PE_3C_2 994.0 (M-1), PE_3C_3 1109.0 (M-1), PE_3M_3 938.0 (M-1), PE_3M_2C 994.8 (M-1), PE_3MC_2 1051.8 (M-1), PE_3M_2 879.8 (M-1).

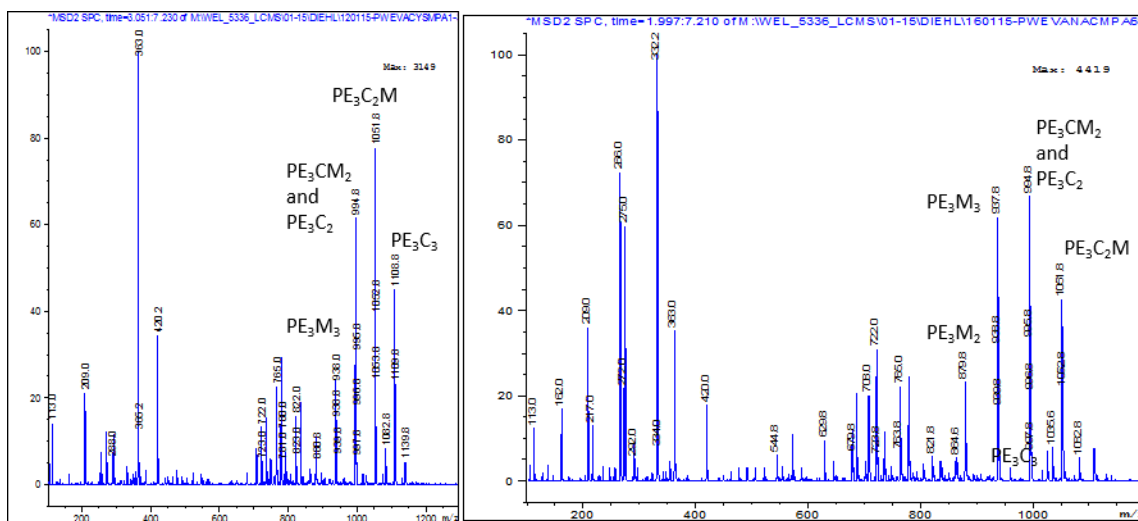


Figure 5.38 MS analysis of reaction of PE_3C_3 with 3-mercaptopropionic acid. Left: after 48 hr, negative mode, right: after 8 days, negative mode.

5.9.4.16 Coupling of Triamine to Tentagel S Br (Resin-P)

Tentagel S Br (0.26 mmol/g, 107 mg, 27.8 μ mol, 1 eq) was placed in 5 mL of 1:1 water/MeCN and stirred for 15 min to swell the resin. The triamine (14.1 mg, 85.5 μ mol, 3.1 eq), proton sponge (18.5 mg, 86.3 μ mol, 3.1 eq), and NaI (16 mg, 107 μ mol, 3.8 eq) were added to the solution. The reaction was heated to 80 $^{\circ}$ C for 20 hr. The beads were washed with water and methanol and subjected to a Kaiser test. The test was positive for amine. The resin was dried thoroughly under vacuum.

5.9.4.17 Kaiser Test

Solutions: KCN (16.5 mg KCN in 25 mL H_2O , then 1 mL of this solution was diluted to 50 mL total in pyridine). Ninhydrin (1 g ninhydrin dissolved in 10 mL ethanol). Phenol (20 g phenol in 10 mL ethanol). To a few beads of resin was added three drops of each of the three solutions. The solution was heated to 100 $^{\circ}$ C for 10 minutes to develop the color. Blue/purple = amine.

5.9.4.18 Resin-PE₂

Resin-P (107 mg, 27.8 μmol , 1 eq) was placed in 3 mL of 0.1 M phosphate buffer, 0.15 M NaCl, H₂O, pH 7. EvaCA (28 mg, 113 μmol , 2 eq) was dissolved in 2 mL MeCN and added to the resin. The reaction was gently sparged with nitrogen for 24 hr. Then the resin was washed with water and methanol and subjected to a Kaiser test (negative for amine).

5.9.4.19 Resin-PE₂C₂

Resin-PE₂ (11.67 mg) was placed in 1 mL of 0.1 M phosphate buffer, 0.15 M NaCl, H₂O, pH 7.2, and NAC (2 mg, 12.14 μmol) was added. The reaction was gently sparged with nitrogen for 2.5 days. The resin was washed with water and methanol. Then, a solution of DTT (6.15 mg, 40 μmol) was added to the resin. The reaction was stirred for 36 hr. Then the solution was collected and analyzed by LC-MS (**Figure 5.39** and **5.40**). The resin was washed and subjected to a Kaiser test (positive for amine).

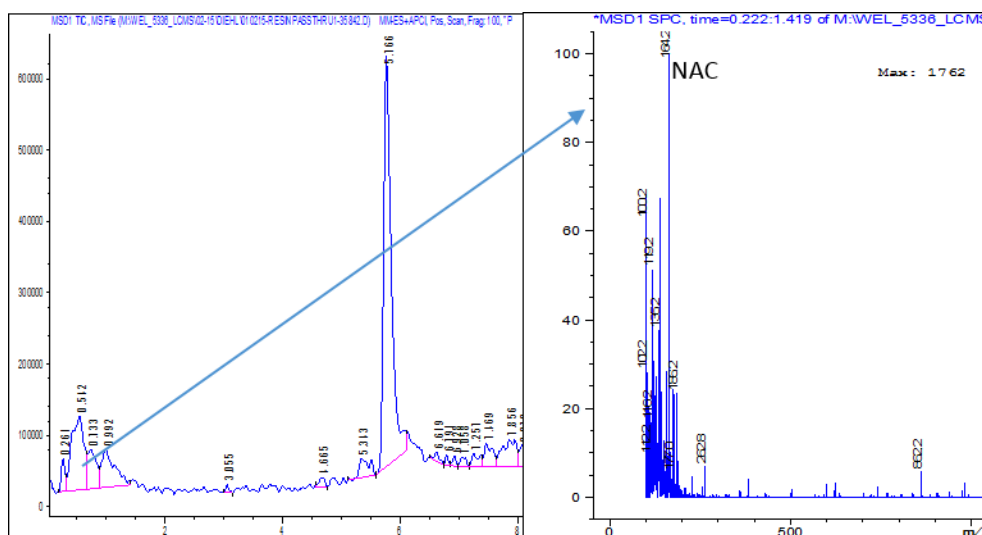


Figure 5.39 LC-MS analysis of the flow-through from the DTT reaction with resin-PE₂C₂. TIC positive mode.

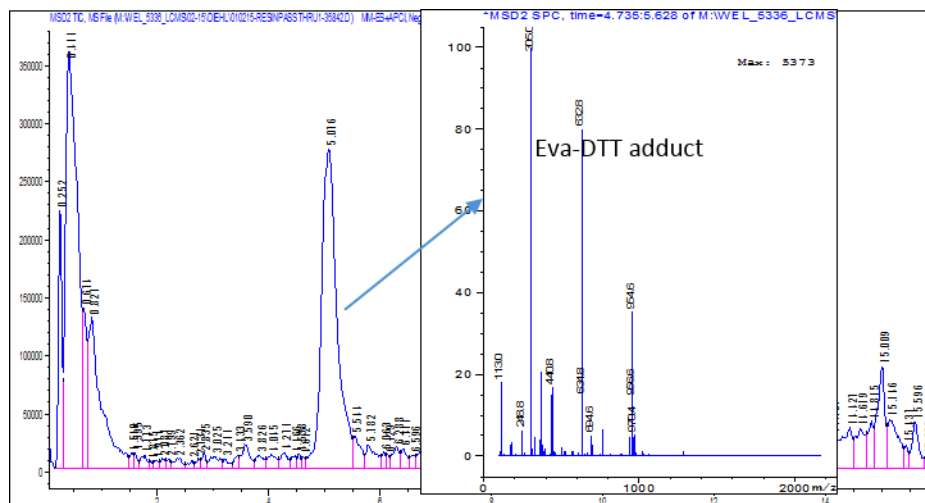


Table 5.2 Summary of peptide solutions and DCL setup.

Peptide	MW (g/mol)	mg	vol (mL)	concentration (mM)	vol added to library (μ L)	concentration in library (mM)
KGGRKC	1103.89	61.9	5	11.2	44.6	0.8
KRGSKC	1133.92	48.9	5	8.6	58.0	0.8
KRSSKC	1163.94	64.3	5	11.0	45.3	0.8
KAGLKC	960.86	52.9	5	11.0	45.4	0.8
KYPYKC	1143.04	56.6	5	9.9	50.5	0.8
PRGDKC	1016.84	61.4	5	12.1	41.4	0.8
KDGRKC	1161.93	65.0	5	11.2	44.7	0.8
KGARCC	978.86	40.7	5	8.32	60.1	0.8

5.9.5 Protein Modification Reactions

5.9.5.1 BSA Reduction by DTT²¹

A 1.5-g portion of BSA was added to an argon-flushed flask. To this flask 40 mL of deoxygenated buffer (20 mM PBS, 45 mM NaCl, 1.1 μ M EDTA) was added. A 0.55 M solution of DTT in buffer was prepared and deoxygenated and then added to the BSA solution. The reaction was stirred for 45 min. The reaction mixture was transferred to dialysis tubing under nitrogen and dialyzed against 6 L of 0.1 mM EDTA, 100 mM HCl, 50 mM NaCl with two changes over 18 hr. The solution was removed from the tubing under N₂ and placed in a flask under Ar. It was subjected to the Ellman test as described below. Ellman test results varied between samples with an average of 20 mol SH/mol BSA (s = 5) based on seven reaction products.

5.9.5.2 BSA Thiolation with AHTL²¹

In 10 mL of water 0.5 g BSA and 0.06 g *N*-acetylhomocysteine thiolactone (AHTL, 50 eq) were dissolved. A 100 mM silver nitrate solution was prepared in water (0.085 g in 5 mL) and added dropwise to the BSA/AHTL solution over 15 min. In order to maintain a pH of about 7 during the course of this addition, 1 M NaOH was added concurrently over

the 15 min. The pH was monitored by narrow-range pH paper. As the silver nitrate was added, a white precipitate formed in the solution. To the reaction was added 3.25 mL of a 500 mM thiourea solution in water, and then the reaction was acidified with 1 M HCl to pH 2. During the acidification the solution first cleared and turned yellow and then turned a milky brown at pH 2. The solution was filtered and passed through an ion exchange column to remove silver (Amberlite IR-120 H⁺ ion exchange resin) with a water mobile phase. This process was repeated three times. The solution was dialyzed against 24 L of 5 mM HCl over 18 hr with two changes. The product was subjected to the Ellman test as described below. Ellman test registered an average of 17 mol SH per mol BSA (s = 2) based on five reaction products.

5.9.5.3 BSA Thiolation with SATA and Deprotection¹⁵

BSA (1.0 g, 15.04 μ mol) was dissolved in 40 mL of 0.1M PBS buffer (0.15M NaCl, pH 7.2). *N*-succinimidyl *S*-acetylthioacetate (172 mg, 752 μ mol, 50 eq) was dissolved in 4 mL DMSO and added to the BSA solution. The reaction was stirred overnight (18 hr). The reaction mixture was dialyzed against 9 L total of 0.1M PBS, 50 mM EDTA, 0.2 w/v % NaN₃, pH 7.5 with two changes for several days until the SATA smell was absent. The protein solution was removed from the dialysis tubing and filtered. The protein concentration was measured as described below to be 253 μ M (n = 4, s = 5) with a total volume of about 33 mL. A 20-mL portion of this solution was taken and to it was added 2.53 mL of a 1 M solution of hydroxylamine in buffer (0.1M PBS, pH 7.2). To adjust the pH back to 7.5, 1 M NaOH was used. The reaction was stirred for 2.5 hr. Next, the reaction was subjected to the Ellman test and then dialyzed against 0.1 M PBS, 50 mM EDTA, 0.2 w/v % NaN₃, pH 7.5 (5 L total with two changes over 24 hr). The Ellman test registered an average of 11 mol SH per mol BSA (s = 4) from four reactions.

5.9.5.4 Reaction of t-BSA with Methyl Vinyl Ketone

A portion of t-BSA solution in buffer (0.1 M PBS, 50 mM EDTA, 0.2 w/v % NaN₃, pH 7.5) was made 5 mM in DTT. The reaction was placed under nitrogen and half of the DTT was added. After stirring for 30 min, the other half of the DTT was quickly added. To the protein solution was added methyl vinyl ketone. The amount of MVK added was calculated by the following equation so that an appropriate excess would be achieved based on the approximate number of thiols on the protein and the thiols from unreacted DTT in the solution:

$$\text{mol MVK} = [(\text{mol BSA} * 20) + (\text{mol DTT} * 2)] * 50$$

The reaction stirred overnight at which point the reaction had turned orange. The reaction mixture was dialyzed against 6 L of water overnight (one change) until the smell of MVK was absent. The solution was concentrated by about one third with Aquacide II. Then it was dialyzed against 4 L of 10 mM PBS, 0.2 w/v % NaN₃, pH 7 overnight (one change). Finally, the solution was filtered and the concentration measured.

5.9.5.5 BSAEvaVVKLKC

EvaCA (17.4 mg, 70 μmol, 115 eq) was dissolved in 0.7 mL MeCN. BSA (40.5 mg, 0.6 μmol, 1 eq) was dissolved in 4.5 mL of 50 mM phosphate buffer, H₂O, 10 mM EDTA, pH 7.5. The EvaCA solution was added to the BSA solution. The reaction was stirred in an open vial for 18 hr. The reaction was filter and purified by 10 kDa MWCO centrifugal filter unit against 50 mM phosphate buffer, H₂O, 10 mM EDTA, pH 7.5. The concentration of the protein solution was found to be 107 μM by Bradford assay (0.11 μmol, 1 mL). To the BSAEva was added VVKLKC (2.18 mg, 2.11 μmol, 20 eq). The reaction was stirred in an open vial for two weeks. The reaction was filtered and purified by 10 kDa MWCO centrifugal filter unit against 50 mM phosphate buffer, H₂O, 10 mM EDTA, pH 7.5. The sample was digested and analyzed by LC-ESI-MS³ (see 5.8.6).

5.9.5.6 BSAEvaVVKLKC + DTT

BSAEvaVVKLKC was prepared as described (0.04 μmol). To the protein solution was added dithiothreitol (4.5 mg, 29 μmol , 725 eq). The reaction was heated to 40 °C for 18 hr. The protein solution was purified by 10 kDa MWCO centrifugal filter unit against 50 mM phosphate buffer, H₂O, 10 mM EDTA, pH 7.5. The flowthrough was collected and subjected to LC-MS analysis, which showed VVKLKC and DTT-Eva adduct.

5.9.5.7 BSAEvaCys

EvaCA (12 mg, 48 μmol , 1 eq) was dissolved in 0.5 mL acetonitrile and added to a 5 mL solution of NAC (16.6 mg, 102 μmol , 2.1 eq) in 0.1 M phosphate buffer, 0.15 M NaCl, H₂O, pH 7.2. The reaction was gently sparged with nitrogen for 24 hr. Then BSA (64 mg, 0.96 μmol , 0.02 eq) was added, and the reaction was stirred for three days. The reaction was filtered and purified by 10 kDa MWCO centrifugal filter unit against 0.1 M phosphate buffer, 0.15 M NaCl, H₂O, pH 7.2. The sample was digested and analyzed by LC-ESI-MS³ (see 5.8.6).

Eva- 100, 156, 198, 204, 264, 297, 336, 340, 399, 402, 580

EvaCys- 580, a few on 204, 336, 340

5.9.5.8 MyoEvaMEEEE

Equine skeletal muscle myoglobin (14 mg, 0.82 μmol , 1 eq) was dissolved in 4 mL of 0.1 M phosphate buffer, 0.15 M NaCl, H₂O, pH 7.2. EvaCA (8.2 mg, 33 μmol , 40 eq) was dissolved in 0.5 mL acetonitrile and was added to the myoglobin solution. The reaction was stirred in an open container for 24 hr. The protein was filtered and purified against 0.1 M phosphate buffer, 0.15 M NaCl, H₂O, pH 7.2 with 10 kDa MWCO centrifugal filter units. Then the MEEEE (16 μL , 84 μmol , 102 eq) was added, and the reaction was stirred for another 24 hrs. The protein was filtered and purified against 0.1 M

phosphate buffer, 0.15 M NaCl, H₂O, pH 7.2 with 10 kDa MWCO centrifugal filter units. A sample was taken of MyoEvaMEEEE and submitted for LC-ESI-MS³ analysis (see 5.8.6). To the remainder was added dithiothreitol (12.8 mg, 83 μmol, 101 eq), and the reaction was stirred for 36 hr. The protein was filtered and purified against 0.1 M phosphate buffer, 0.15 M NaCl, H₂O, pH 7.2 with 10 kDa MWCO centrifugal filter units. The protein sample was submitted for intact ESI-MS, while the flow-through was collected and analyzed by LC-MS (**Figure 5.41**). MEEEE and DTT-Eva adduct was observed.

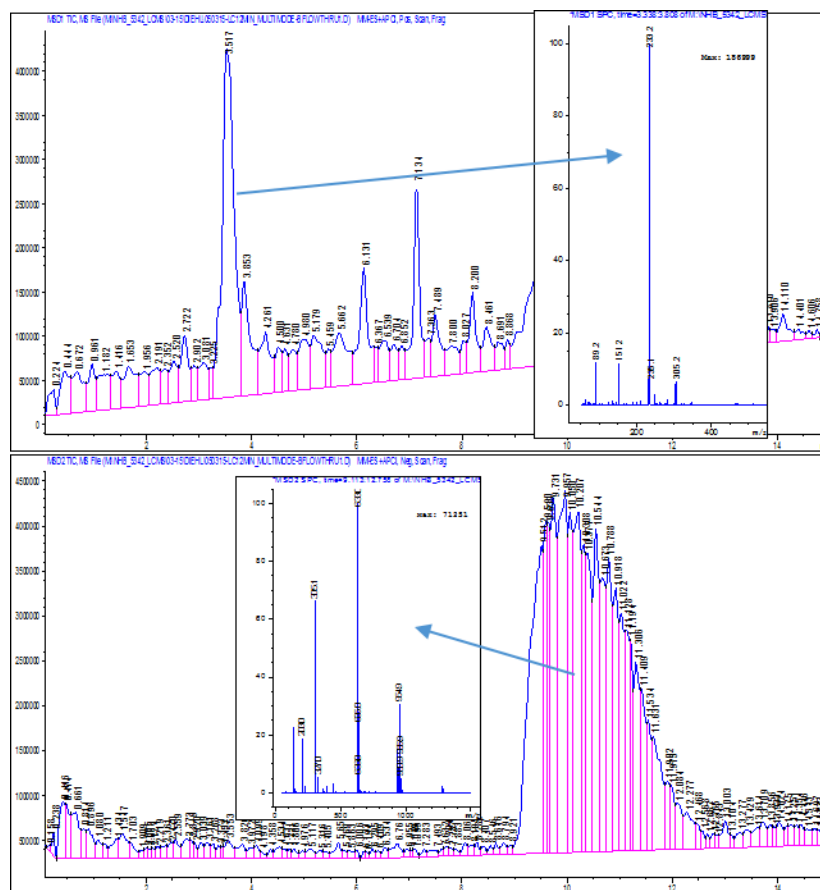


Figure 5.41 LC-MS analysis of flow-through from reaction of MyoEvaMEEEE and DTT. Top: TIC positive mode, bottom: TIC negative mode.

5.9.6 Protein Characterization

5.9.6.1 BSA Concentration Measurement

*Lowry Method*²¹

To a quartz UV-VIS cuvette was added 2 mL of water, which was used to blank the spectrophotometer over 250-350 nm. Protein solution was added to the water, and the absorbance was measured over that range. The reported extinction coefficient of 43,824 $\text{cm}^{-1}\text{M}^{-1}$ at 280 nm for BSA was used to calculate the concentration of the protein solution in the cuvette. The concentration of the original protein solution was calculated from the dilution. Several trials were performed and averaged to obtain the protein concentration.

*Bradford Assay*⁵⁵

A solution of Coomassie Blue G dye was prepared by dissolving 100 mg of dye in 50 mL of methanol. This solution was added to 100 mL of 85% H_3PO_4 , and water was added to dilute to a total volume of 200 mL to yield a dark red solution. Standard solutions of BSA were prepared in 10 mM phosphate buffer, H_2O , pH 7.0, 0.02% NaN_3 (2, 4, 6, 8, 10, 12 μM). Then, 1 part dye stock was diluted with 4 parts water (e.g. 10 mL dye solution + 40 mL water). The protein samples to be measured were generally diluted by 1/10 to bring them within the concentration range of the BSA standards. The diluted dye solution (2 mL) was mixed with 30 μL of each BSA standard and each diluted sample to be measured. The absorbance at 595 nm of each solution was measured to generate a calibration curve from the BSA standards. From the calibration curve, the concentration of the unknown protein samples was determined based on their absorbance at 595 nm.

5.9.6.2 *Ellman test*^{17,18}

A solution of DTNB was prepared by dissolving 20 mg in 5 mL of 0.1 M PBS, pH 8. The protein solution was diluted to a concentration of about 0.2 μM in 0.1 M PBS, pH

8. To 3 mL of this protein solution was added 0.1 mL of DTNB solution. It was allowed to react for 15 minutes. A “blank” was also prepared by adding 0.1 mL of DTNB solution to 3 mL of 0.1 M PBS, pH 8. The absorbance for each sample was read at 410 nm. The “blank” absorbance was subtracted from the absorbance of each sample. The concentration of thiol was calculated from the extinction coefficient for 2-nitro-5-mercaptobenzoic acid, $13,560 \text{ cm}^{-1}\text{M}^{-1}$. Several trials were performed and averaged to obtain the concentration of thiol.

5.9.6.3 Quantitative Ninhydrin Test^{19,56}

The protein samples to be tested were diluted to approximately 10 μM in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3 . A native BSA sample was used as a control for the total measurable amine content. Amine standards were prepared from glycine (10 μM - 500 μM in 10 mM phosphate buffer, H_2O , pH 7.00, 0.02% NaN_3). The dye solution was prepared from ninhydrin (2 g) and hydrindantin (0.3 g) dissolved in 75 mL DMSO and 25 mL 4 M acetate buffer, H_2O , pH 5.5. Each of the sample solutions (1 mL of protein or glycine solution) were mixed with dye solution (1 mL). The solutions were heated to 100 $^\circ\text{C}$ for 15 min. The solutions were allowed to cool, and then 5 mL of a 1:1 water/ethanol solution were added. The absorbance of the resultant solutions was measured at 570 nm. A calibration curve was constructed from the glycine standards absorbance values, and the amine content of the protein samples was determined from this plot.

5.9.6.4 Mass Spectrometric Analysis

Protein digestion

A 200- μg portion of protein was reduced with a 100-fold excess of dithiothreitol (DTT) with respect to protein for 60 minutes at 60 $^\circ\text{C}$ then incubated with a 400-fold excess of iodoacetamide with respect to protein at 37 $^\circ\text{C}$ for 30 minutes in the dark. Excess

iodoacetamide was quenched with additional DTT. Then a 1 $\mu\text{g}/\mu\text{L}$ solution of GluC in 1 mM HCl was added giving a protein-to-enzyme ratio of 20:1. The digestion solution was incubated over night at 37 °C. The digest was desalted using C18 cartridges then dried and reconstituted in water to a concentration of 1 $\mu\text{g}/\mu\text{L}$.

MS analysis

The LC instrument used was a Dionex Ultimate 3000. Mobile phase A was water with 0.1% formic acid, and mobile phase B was acetonitrile with 0.1% formic acid. The gradient was 5% to 40% mobile phase B over 60 min. The ion source was electrospray ionization. The mass analyzer used was a Bruker Daltronics HCT quadrupole ion trap. Collision-induced dissociation with helium gas was used for ion fragmentation. The algorithm used was MassMatrix (MassMatrix.net).

5.9.7 Titrations

The experimental details for the titrations can be found in the corresponding figure captions.

5.10 REFERENCES

- (1) Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J.-L.; Sanders, J. K. M.; Otto, S. *Chem. Rev.* **2006**, *106*, 3652–3711.
- (2) Lehn, J.-M. *Science* **2001**, *291*, 2331–2332.
- (3) Otto, S.; Furlan, R. L. .; Sanders, J. K. . *Drug Discov. Today* **2002**, *7*, 117–125.
- (4) Ingerman, L. A.; Cuellar, M. E.; Waters, M. L. *Chem. Commun. (Camb)*. **2010**, *46*, 1839–1841.
- (5) Ingerman, L. A.; Waters, M. L. *J. Org. Chem.* **2009**, *74*, 111–117.
- (6) Shi, B.; Stevenson, R.; Campopiano, D. J.; Greaney, M. F. *J. Am. Chem. Soc.* **2006**, *128*, 8459–8467.

- (7) Sakai, S.; Shigemasa, Y.; Sasaki, T. *Tetrahedron Lett.* **1997**, *38*, 8145–8148.
- (8) Klekota, B.; Miller, B. L. *Trends Biotechnol.* **1999**, *17*, 205–209.
- (9) Huc, I.; Lehn, J.-M. *Proc. Natl. Acad. Sci.* **1997**, *94*, 2106–2110.
- (10) Otto, S.; Furlan, R. L. E.; Sanders, J. K. M. *Science* **2002**, *297*, 590–593.
- (11) Shi, B.; Greaney, M. F. *Chem. Commun.* **2005**, 886.
- (12) Anslyn, E. V.; Dougherty, D. A. *Modern Physical Organic Chemistry*; University Science Books: Sausalito, CA, 2006.
- (13) Joshi, G.; Anslyn, E. V. *Org. Lett.* **2012**, *14*, 4714–4717.
- (14) Zhong, Y.; Xu, Y.; Anslyn, E. V. *European J. Org. Chem.* **2013**, *2013*, 5017–5021.
- (15) Hermanson, G. T. *Bioconjugate Techniques*; 2nd ed.; Academic Press: London, UK, 2008.
- (16) Riddles, P. W.; Blakeley, R. L.; Zerner, B. *Methods Enzymol.* **1983**, *91*, 49–60.
- (17) Sedlak, J.; Lindsay, R. H. *Anal. Biochem.* **1968**, *25*, 192–205.
- (18) Collier, H. B. *Anal. Biochem.* **1973**, *56*, 310–311.
- (19) Prochazkova, S.; Vårum, K. M.; Ostgaard, K. *Carbohydr. Polym.* **1999**, *38*, 115–122.
- (20) Verheyen, E.; Delain-Bioton, L.; Van Der Wal, S.; El Morabit, N.; Barendregt, A.; Hennink, W. E.; Van Nostrum, C. F. *Macromol. Biosci.* **2010**, *10*, 1517–1526.
- (21) Ewing, J. F.; Young, D. V.; Janero, D. R.; Garvey, D. S.; Grinnell, T. A. *J. Pharmacol. Exp. Ther.* **1997**, *283*, 947–954.
- (22) Biemann, K.; Scoble, H. A. *Science* **1987**, *237*, 992–998.
- (23) Domon, B.; Aebersold, R. *Science* **2006**, *312*, 212–217.
- (24) Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 806–812.

- (25) Siuti, N.; Kelleher, N. L. *Nat. Methods* **2007**, *4*, 817–821.
- (26) Papayannopoulos, I. A. *Mass Spectrom. Rev.* **1995**, *14*, 49–73.
- (27) Wells, J. M.; McLuckey, S. A. *Methods Enzymol.* **2005**, *402*, 148–185.
- (28) Peters, T. *All About Albumin: Biochemistry, Genetics, and Medical Applications*; Academic Press: San Diego, CA, 1996.
- (29) Kolthoff, I. M.; Anastasi, A.; Tan, B. H. *J. Am. Chem. Soc.* **1958**, *80*, 3235–3240.
- (30) Prem P., B.; Katsushi, S.; Takuya, U.; Kunio, T. *Int. J. Biochem.* **1989**, *21*, 857–862.
- (31) Benesch, R.; Benesch, R. E. *J. Am. Chem. Soc.* **1956**, *78*, 1597–1599.
- (32) Benesch, R.; Benesch, R. E. *Proc. Natl. Acad. Sci. U. S. A.* **1958**, *44*, 848–853.
- (33) Weber, G.; Daniel, E. *Biochemistry* **1966**, *5*, 1900–1907.
- (34) Prabakaran, S.; Tepp, W.; DasGupta, B. . *Toxicon* **2001**, *39*, 1515–1531.
- (35) Duncan, R. J.; Weston, P. D.; Wrigglesworth, R. *Anal. Biochem.* **1983**, *132*, 68–73.
- (36) Aufort, M.; Gonera, M.; Le Gal, J.; Czarny, B.; Le Clainche, L.; Thai, R.; Dugave, C. *Chembiochem* **2011**, *12*, 583–592.
- (37) Huisgen, R. *Angew. Chemie Int. Ed. English* **1963**, *2*, 565–598.
- (38) Lin, Z.; Lu, P.; Yu, X.; Zhang, W.-B.; Huang, M.; Wu, K.; Guo, K.; Wesdemiotis, C.; Zhu, X.; Zhang, Z.; Yue, K.; Cheng, S. Z. D. *Macromolecules* **2014**, *47*, 4160–4168.
- (39) Wentrup, C.; Lorencak, P. *J. Am. Chem. Soc.* **1988**, *110*, 1880–1883.
- (40) Ben Cheikh, A.; Chuche, J.; Manisse, N.; Pommelet, J. C.; Netsch, K. P.; Lorencak, P.; Wentrup, C. *J. Org. Chem.* **1991**, *56*, 970–975.
- (41) Chu, D. T. W.; Chen, J.; Zhang, W.; Li, X.; Song, J.; Wang, B.; Cong, Q.; James, D. R. Spiro derivatives as lipoxygenase inhibitors. 20060128790, 2006.

- (42) Remuzon, P.; Bouzard, D.; DiCesare, P.; Essiz, M.; Jacquet, J.-P.; Nicolau, A.; Martel, A.; Menard, M.; Bachand, C. *Tetrahedron* **1995**, *51*, 9657–9670.
- (43) Gade, A. M. *Sensing Approaches for the Discrimination of Small Molecules and Multivalent Analytes*, 2014.
- (44) Bujacz, A. *Acta Crystallogr. D. Biol. Crystallogr.* **2012**, *68*, 1278–1289.
- (45) Sharma, M.; Tiwari, M.; Chandra, R. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5347–5350.
- (46) Nagao, S.; Osuka, H.; Yamada, T.; Uni, T.; Shomura, Y.; Imai, K.; Higuchi, Y.; Hirota, S. *Dalton Trans.* **2012**, *41*, 11378–11385.
- (47) Wallace, K. J.; Hanes, R.; Anslyn, E.; Morey, J.; Kilway, K. V.; Siegel, J. *Synthesis (Stuttg.)* **2005**, *2005*, 2080–2083.
- (48) Gunasekara, R. W.; Zhao, Y. *J. Am. Chem. Soc.* **2015**, *137*, 843–849.
- (49) McNaughton, B. R.; Miller, B. L. *Org. Lett.* **2006**, *8*, 1803–1806.
- (50) Quarrell, R.; Claridge, T. D. W.; Weaver, G. W.; Lowe, G. *Mol. Divers.* **1996**, *1*, 223–232.
- (51) Raju, B.; Kogan, T. P. *Tetrahedron Lett.* **1997**, *38*, 4965–4968.
- (52) Pucci, M.; Alfei, S.; Castellaro, S.; Lucchesini, F.; Milanese, M.; Bertini, V. *Polym. J.* **2013**, *45*, 1146–1152.
- (53) Naffin, J. L.; Han, Y.; Olivos, H. J.; Reddy, M. M.; Sun, T.; Kodadek, T. *Chem. Biol.* **2003**, *10*, 251–259.
- (54) Liang, R.; Loebach, J.; Horan, N.; Ge, M.; Thompson, C.; Yan, L.; Kahne, D. *Proc. Natl. Acad. Sci.* **1997**, *94*, 10554–10559.
- (55) Kruger, N. J. *Methods Mol. Biol.* **1994**, *32*, 9–15.
- (56) Lamothe, P. J.; McCormick, P. G. *Anal. Chem.* **1973**, *45*, 1906–1911.

Bibliography

- Adams, M. M.; Anslyn, E. V. *J. Am. Chem. Soc.* **2009**, *131*, 17068–17069
- Adlof, R. *J. Chromatogr. A* **2007**, *1148*, 256–259.
- Aerneck, M. J.; Walt, D. R. *Sensors Actuators B Chem.* **2009**, *142*, 464–469.
- Agbaria, R. A.; Oldham, P. B.; McCarroll, M.; McGown, L. B.; Warner, I. M. *Anal. Chem.* **2002**, *74*, 3952–3962.
- Ajayaghosh, A. *Acc. Chem. Res.* **2005**, *38*, 449–459.
- Aleman, L. B. *Chem. Phys. Lipids* **2002**, *120*, 33–44.
- Aluyor, O., E.; Ozigagu, E., C.; Oboh, I., O.; *P. Sci. Res. Essays* **2009**, *4*, 191–197.
- An, H. J.; Kronewitter, S. R.; de Leoz, M. L. A.; Lebrilla, C. B. *Curr. Opin. Chem. Biol.* **2009**, *13*, 601–607.
- Andrikopoulos, N. K. *Crit. Rev. Food Sci. Nutr.* **2002**, *42*, 473–505.
- Anslyn, E. V.; Dougherty, D. A. *Modern Physical Organic Chemistry*; University Science Books: Sausalito, CA, 2006.
- Arun, K. T.; Jayaram, D. T.; Avirah, R. R.; Ramaiah, D. *J. Phys. Chem. B* **2011**, *115*, 7122–7128.
- Arunkumar, E.; Chithra, P.; Ajayaghosh, A. *J. Am. Chem. Soc.* **2004**, *126*, 6590–6598.
- Ashbrook, J. D.; Spector, A. A.; Fletcher, J. E. *J. Biol. Chem.* **1972**, *247*, 7038–7042.
- Aufort, M.; Gonera, M.; Le Gal, J.; Czarny, B.; Le Clainche, L.; Thai, R.; Dugave, C. *Chembiochem* **2011**, *12*, 583–592.
- Axel, R. *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 6110–6127.
- Bajaj, A.; Miranda, O. R.; Phillips, R.; Kim, I. B.; Jerry, D. J.; Bunz, U. H. F.; Rotello, V. *M. J. Am. Chem. Soc.* **2010**, *132*, 1018–1022.
- Baldini, L.; Wilson, A. J.; Hong, J.; Hamilton, A. D. *J. Am. Chem. Soc.* **2004**, *126*, 5656–5657.

- Basheer, M. C.; Alex, S.; George Thomas, K.; Suresh, C. H.; Das, S. *Tetrahedron* **2006**, *62*, 605–610.
- Ben Cheikh, A.; Chucho, J.; Manisse, N.; Pommelet, J. C.; Netsch, K. P.; Lorencak, P.; Wentrup, C. *J. Org. Chem.* **1991**, *56*, 970–975.
- Benesch, R.; Benesch, R. E. *J. Am. Chem. Soc.* **1956**, *78*, 1597–1599.
- Benesch, R.; Benesch, R. E. *Proc. Natl. Acad. Sci. U. S. A.* **1958**, *44*, 848–853.
- Benitez-Sanchez, P. L.; Leon-Camacho, M.; Aparicio, R. *Eur. Food Res. Technol.* **2003**, *218*, 13–19.
- Biemann, K.; Scoble, H. A. *Science* **1987**, *237*, 992–998.
- Biermann, U.; Meier, M. A. R.; Butte, W.; Metzger, J. O. *Eur. J. Lipid Sci. Technol.* **2011**, *113*, 39–45.
- Bigelow, R. W.; Freund, H.-J. *Chem. Phys.* **1986**, *107*, 159–174.
- Bonizzoni, M.; Long, S. R.; Rainwater, C.; Anslyn, E. V. *J. Org. Chem.* **2012**, *77*, 1258–1266.
- Brandão, L. F. P.; Braga, J. W. B.; Suarez, P. A. Z. *J. Chromatogr. A* **2012**, *1225*, 150–157.
- Brasaemle, D. L. *J. Lipid Res.* **2007**, *48*, 2547–2559.
- Brereton, R. G. *Chemometrics Data Analysis for the Laboratory and Chemical Plant*; Wiley: West Sussex, UK, 2003.
- Brodersen, R. *CRC Crit. Rev. Clin. Lab. Sci.* **1980**, *11*, 305–399.
- Buck, L. B. *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 6128–6140.
- Bujacz, A. *Acta Crystallogr. D. Biol. Crystallogr.* **2012**, *68*, 1278–1289.
- Burns, J. A.; Whitesides, G. M. *Chem. Rev.* **1993**, *93*, 2583–2601.
- Buryak, A.; Severin, K. *J. Am. Chem. Soc.* **2005**, *127*, 3700–3701.
- Carey, J. R.; Suslick, K. S.; Hulkower, K. I.; Imlay, J. A.; Imlay, K. R. C.; Ingison, C. K.; Ponder, J. B.; Sen, A.; Wittrig, A. E. *J. Am. Chem. Soc.* **2011**, *133*, 7571–7576.

- Carter, D. C.; Ho, J. X. *Adv. Protein Chem.* **1994**, *45*, 153–203.
- Cermenati, G.; Abbiati, F.; Cermenati, S.; Brioschi, E.; Volonterio, A.; Cavaletti, G.; Saez, E.; De Fabiani, E.; Crestani, M.; Garcia-Segura, L. M.; Melcangi, R. C.; Caruso, D.; Mitro, N. *J. Lipid Res.* **2012**, *53*, 300–310.
- Chen, H.; Farahat, M. S.; Law, K.-Y.; Whitten, D. G. *J. Am. Chem. Soc.* **1996**, *118*, 2584–2594.
- Cole, E. L.; Arunkumar, E.; Xiao, S.; Smith, B. A.; Smith, B. D. *Org. Biomol. Chem.* **2012**, *10*, 5769–5773.
- Collier, H. B. *Anal. Biochem.* **1973**, *56*, 310–311.
- Collins, B. E.; Wright, A. T.; V., A. E. *Top. Curr. Chem.* **2007**, *277*, 181–218.
- Collins, C. G.; Peck, E. M.; Kramer, P. J.; Smith, B. D. *Chem. Sci.* **2013**, *4*, 2557.
- Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J.-L.; Sanders, J. K. M.; Otto, S. *Chem. Rev.* **2006**, *106*, 3652–3711.
- Corrie, J. E. T.; Trentham, D. R. *J. Chem. Soc. Perkin Trans. 1* **1995**, 1993.
- Chu, D. T. W.; Chen, J.; Zhang, W.; Li, X.; Song, J.; Wang, B.; Cong, Q.; James, D. R. Spiro derivatives as lipoxygenase inhibitors. 20060128790, 2006.
- Das, S.; Kamat, P. V.; De la Barre, B.; Thomas, K. G.; Ajayaghosh, A.; George, M. V. *J. Phys. Chem.* **1992**, *96*, 10327–10330.
- Das, S.; Thanulingam, T. L.; Thomas, K. G.; Kamat, P. V.; George, M. V. *J. Phys. Chem.* **1993**, *97*, 13620–13624.
- De, M.; Rana, S.; Akpınar, H.; Miranda, O. R.; Arvizo, R. R.; Bunz, U. H. F.; Rotello, V. *M. Nat. Chem.* **2009**, *1*, 461–465.
- De Bruin, B.; Hauwert, P.; Reek, J. N. H. *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 2660–2663.
- Diehl, K. L.; Anslyn, E. V. *Chem. Soc. Rev.* **2013**, *42*, 8596–8611.
- Dirk, C. W.; Herndon, W. C.; Cervantes-Lee, F.; Selnau, H.; Martinez, S.; Kalamegham, P.; Tan, A.; Campos, G.; Velez, M. *J. Am. Chem. Soc.* **1995**, *117*, 2214–2225.

- Dobson, G.; Christie, W. W.; Nikolova-Damyanova, B. *J. Chromatogr. B Biomed. Sci. Appl.* **1995**, *671*, 197–222.
- Domon, B.; Aebersold, R. *Science* **2006**, *312*, 212–217.
- Doty, R. L. *Handbook of Olfaction and Gustation*; 2nd ed.; Marcel Dekker: New York, 2003.
- Duncan, R. J.; Weston, P. D.; Wrigglesworth, R. *Anal. Biochem.* **1983**, *132*, 68–73.
- Edwards, N. Y.; Sager, T. W.; McDevitt, J. T.; Anslyn, E. V. *J. Am. Chem. Soc.* **2007**, *129*, 13575–13583.
- Eichmann, T. O.; Kumari, M.; Haas, J. T.; Farese, R. V.; Zimmermann, R.; Lass, A.; Zechner, R. *J. Biol. Chem.* **2012**, *287*, 41446–41457.
- Er, J. C.; Vendrell, M.; Tang, M. K.; Zhai, D.; Chang, Y.-T. *ACS Comb. Sci.* **2013**, *15*, 452–457.
- Ewing, J. F.; Young, D. V.; Janero, D. R.; Garvey, D. S.; Grinnell, T. A. *J. Pharmacol. Exp. Ther.* **1997**, *283*, 947–954.
- Fairclough, R. H.; Cantor, C. R. *Methods Enzymol.* **1978**, *48*, 347–379.
- Fan, J.; Chen, C.; Lin, Q.; Fu, N. *Sensors Actuators B Chem.* **2012**, *173*, 874–881.
- Fan, J.; Wang, Z.; Zhu, H.; Fu, N. *Sensors Actuators B Chem.* **2013**, *188*, 886–893.
- Folmer-Andersen, J. F.; Kitamura, M.; Anslyn, E. V. *J. Am. Chem. Soc.* **2006**, *128*, 5652–5653.
- Gade, A. M. *Sensing Approaches for the Discrimination of Small Molecules and Multivalent Analytes*, 2014.
- Germain, M. E.; Knapp, M. J. *Inorg. Chem.* **2008**, *47*, 9748–9750.
- Ghazarossian, V.; Pease, J. S.; Hu, M. W.; Laney, M.; Tarnowski, T. L. *Fluorescent Dyes*. EP89308412, 1989.
- Girousse, A.; Langin, D. *Int. J. Obes. (Lond)*. **2012**, *36*, 581–594.
- Gude, M.; Ryf, J.; White, P. D. *Lett. Pept. Sci.* **2002**, *9*, 203–206.

- Guilherme, A.; Virbasius, J. V.; Puri, V.; Czech, M. P. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 367–377.
- Gunasekara, R. W.; Zhao, Y. *J. Am. Chem. Soc.* **2015**, *137*, 843–849.
- Gunstone, F. D. *The Chemistry of Oils and Fats: Sources, Compositions, Properties, and Uses*; CRC Press: Boca Raton, FL, 2004.
- Hazel, J. R.; Sidell, B. D. *J. Exp. Biol.* **2004**, *207*, 897–903.
- Hemmateenejad, B.; Yousefinejad, S. *Anal. Bioanal. Chem.* **2009**, *394*, 1965–1975.
- Hermanson, G. T. *Bioconjugate Techniques*; 2nd ed.; Academic Press: London, UK, 2008.
- Herrero, A.; Cruz Ortiz, M.; Arcos, J.; López-Palacios, J.; Sarabia, L. *Anal. Chim. Acta* **1994**, *293*, 277–293.
- Hewage, H. S.; Anslyn, E. V. *J. Am. Chem. Soc.* **2009**, *131*, 13099–13106.
- Hirsch, T.; Kettenberger, H.; Wolfbeis, O. S.; Mirsky, V. M. *Chem. Commun.* **2003**, 432–433.
- Hong, Y.; Feng, C.; Yu, Y.; Liu, J.; Lam, J. W. Y.; Luo, K. Q.; Tang, B. Z. *Anal. Chem.* **2010**, *82*, 7035–7043.
- Huc, I.; Lehn, J.-M. *Proc. Natl. Acad. Sci.* **1997**, *94*, 2106–2110.
- Huisgen, R. *Angew. Chemie Int. Ed. English* **1963**, *2*, 565–598.
- Hurd, C. D.; Schmerling, L. *J. Am. Chem. Soc.* **1937**, *59*, 112–117.
- Ingerman, L. A.; Cuellar, M. E.; Waters, M. L. *Chem. Commun. (Camb)*. **2010**, *46*, 1839–1841.
- Ingerman, L. A.; Waters, M. L. *J. Org. Chem.* **2009**, *74*, 111–117.
- Isgor, Y. G.; Akkaya, E. U. *Tetrahedron Lett.* **1997**, *38*, 7417–7420.
- Ivy, M. A. Differential Sensing of Hydrophobic Analytes with Serum Albumin, 2012.
- Ivy, M. A.; Gallagher, L. T.; Ellington, A. D.; Anslyn, E. V. *Chem. Sci.* **2012**, *3*, 1773.

- J-Aggregates, Volume 2*; Kobayashi, T., Ed.; World Scientific Publishing Co. Pte. Ltd.: Singapore, 2012.
- James, T. D.; Sandanayake, K. R. A. S.; Shinkai, S. *Angew. Chemie Int. Ed. English* **1996**, *35*, 1910–1922.
- Jisha, V. S.; Arun, K. T.; Hariharan, M.; Ramaiah, D. *J. Am. Chem. Soc.* **2006**, *128*, 6024–6025.
- Jisha, V. S.; Arun, K. T.; Hariharan, M.; Ramaiah, D. *J. Phys. Chem. B* **2010**, *114*, 5912–5919.
- Jolliffe, I. T. *Principal Component Analysis*; 2nd ed.; Springer: New York, 2002.
- Jose, J.; Ueno, Y.; Burgess, K. *Chemistry* **2009**, *15*, 418–423.
- Joshi, G.; Anslyn, E. V. *Org. Lett.* **2012**, *14*, 4714–4717.
- Jurs, P. C.; Bakken, G. A.; McClelland, H. E. *Chem. Rev.* **2000**, *100*, 2649–2678.
- Karimi, M. A.; Mazloun Ardakani, M.; Behjatmanesh Ardakani, R.; Mashhadizadeh, M. H.; Zadeh, N. Z. *Anal. Bioanal. Electrochem.* **2009**, *1*, 142–158.
- Kasai, S.; Horie, T.; Mizuma, T.; Awazu, S. *J. Pharm. Sci.* **2006**, *76*, 387–392.
- Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 806–812.
- Kirkland, T. A.; Lynn, D. M.; Grubbs, R. H. *J. Org. Chem.* **1998**, *63*, 9904–9909.
- Klekota, B.; Miller, B. L. *Trends Biotechnol.* **1999**, *17*, 205–209.
- Kolthoff, I. M.; Anastasi, A.; Tan, B. H. *J. Am. Chem. Soc.* **1958**, *80*, 3235–3240.
- Kovac, A.; Scheib, H.; Pleiss, J.; Schmid, R. D.; Paltauf, F. *Eur. J. Lipid Sci. Technol.* **2000**, *102*, 61–77.
- Kraemer, F. B. *J. Lipid Res.* **2002**, *43*, 1585–1594.
- Kritchevsky, D. *J. Nutr. Biochem.* **1995**, *6*, 172–178.
- Kruger, N. J. *Methods Mol. Biol.* **1994**, *32*, 9–15.

- Kubarych, C. J.; Adams, M. M.; Anslyn, E. V. *Org. Lett.* **2010**, *12*, 4780–4783.
- Lacowicz, J. R. *Principles of Fluorescence Spectroscopy*; Plenum Press: New York, 1986.
- Lamothe, P. J.; McCormick, P. G. *Anal. Chem.* **1973**, *45*, 1906–1911.
- Large, V.; Arner, P. *Diabetes Metab.* **1998**, *24*, 409–418.
- Larsson, K. In *The Lipid Handbook*; Gunstone, F. D.; Hawood, J. L.; Padley, F. B., Eds.; Chapman and Hall: London, 1986; pp. 321–384.
- Lavigne, J. J.; Anslyn, E. V. *Angew. Chemie Int. Ed.* **2001**, *40*, 3118–3130.
- Law, K. Y. *Chem. Rev.* **1993**, *93*, 449–486.
- Law, K.-Y.; Bailey, F. C. *Dye. Pigment.* **1988**, *9*, 85–107.
- Law, K.-Y.; Bailey, F. C.; Bluett, L. J. *Can. J. Chem.* **1986**, *64*, 1607–1619.
- Lee, J. W.; Lee, J.-S.; Chang, Y.-T. *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 6485–6487.
- Lehn, J.-M. *Science.* **2001**, *291*, 2331–2332.
- Lévêque, N. L.; Héron, S.; Tchaplal, A. *J. Mass Spectrom.* **2010**, *45*, 284–296.
- Li, X.; Evans, J. J. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 2528–2538.
- Liang, R.; Loebach, J.; Horan, N.; Ge, M.; Thompson, C.; Yan, L.; Kahne, D. *Proc. Natl. Acad. Sci.* **1997**, *94*, 10554–10559.
- Liao, S.; Han, W.; Ding, H.; Xie, D.; Tan, H.; Yang, S.; Wu, Z.; Shen, G.; Yu, R. *Anal. Chem.* **2013**, *85*, 4968–4973.
- Lim, S. H.; Musto, C. J.; Park, E.; Zhong, W.; Suslick, K. S. *Org. Lett.* **2008**, *10*, 4405–4408.
- Lin, Z.; Lu, P.; Yu, X.; Zhang, W.-B.; Huang, M.; Wu, K.; Guo, K.; Wesdemiotis, C.; Zhu, X.; Zhang, Z.; Yue, K.; Cheng, S. Z. D. *Macromolecules* **2014**, *47*, 4160–4168.
- Liu, X.-D.; Sun, R.; Ge, J.-F.; Xu, Y.-J.; Xu, Y.; Lu, J.-M. *Org. Biomol. Chem.* **2013**, *11*, 4258–4264.

- Lorber, A. *Anal. Chem.* **1986**, *58*, 1167–1172.
- Lorber, A.; Faber, K.; Kowalski, B. R. *Anal. Chem.* **1997**, *69*, 1620–1626.
- Lozano, V. A.; Ibañez, G. A.; Olivieri, A. C. *Anal. Chim. Acta* **2009**, *651*, 165–172.
- Luo, C.; Zhou, Q.; Zhang, B.; Wang, X. *New J. Chem.* **2011**, *35*, 45.
- Malnic, B.; Hirono, J.; Sato, T.; Buck, L. B. *Cell* **1999**, *96*, 713–723.
- Marchesini, S.; Preti, A.; Aleo, M. F.; Casella, A.; Dagan, A.; Gatt, S. *Chem. Phys. Lipids* **1990**, *53*, 165–175.
- Marth, J. D.; Grewal, P. K. *Nat. Rev. Immunol.* **2008**, *8*, 874–887.
- Martinez, A. W.; Phillips, S. T.; Whitesides, G. M.; Carrilho, E. *Anal. Chem.* **2010**, *82*, 3–10.
- McNaughton, B. R.; Miller, B. L. *Org. Lett.* **2006**, *8*, 1803–1806.
- Meisner, H.; Neet, K. *Mol. Pharmacol.* **1978**, *14*, 337–346.
- Miranda, O. R.; Li, X.; Garcia-Gonzalez, L.; Zhu, Z. J.; Yan, B.; Bunz, U. H. F.; Rotello, V. M. *J. Am. Chem. Soc.* **2011**, *133*, 9650–9653.
- Miranda, O. R.; You, C.-C.; Phillips, R.; Kim, I.-B.; Ghosh, P. S.; Bunz, U. H. F.; Rotello, V. M. *J. Am. Chem. Soc.* **2007**, *129*, 9856–9857.
- Mohseni, N.; Bahram, M.; Olivieri, A. C. *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* **2014**, *122*, 721–730.
- Mol, J. C. *Green Chem.* **2002**, *4*, 5–13.
- Murphy, R. C.; Fiedler, J.; Hevko, J. *Chem. Rev.* **2001**, *101*, 479–526.
- Murphy, R. C.; Gaskell, S. J. *J. Biol. Chem.* **2011**, *286*, 25427–25433.
- Murphy, R. C.; Leiker, T. J.; Barkley, R. M. *Biochim. Biophys. Acta* **2011**, *1811*, 776–783.
- Naffin, J. L.; Han, Y.; Olivos, H. J.; Reddy, M. M.; Sun, T.; Kodadek, T. *Chem. Biol.* **2003**, *10*, 251–259.

- Nagao, S.; Osuka, H.; Yamada, T.; Uni, T.; Shomura, Y.; Imai, K.; Higuchi, Y.; Hirota, S. *Dalton Trans.* **2012**, *41*, 11378–11385.
- Nakamoto, T. *Handbook of Machine Olfaction: Electronic Nose Technology*; Pearce, T. C.; Schiffman, S. S.; Nagle, H. T.; Gardner, J. W., Eds.; Wiley-VCH: Weinheim, Germany, 2006.
- Nguyen, B. T.; Anslyn, E. V. *Coord. Chem. Rev.* **2006**, *250*, 3118–3127.
- Nishijima, M.; Kato, H.; Fukuhara, G.; Yang, C.; Mori, T.; Maruyama, T.; Otagiri, M.; Inoue, Y. *Chem. Commun.* **2013**, *49*, 7433.
- Oils and Fats Authentication*; Jee, M., Ed.; CRC Press: Boca Raton, FL, 2002.
- Organic Synthesis and Molecular Engineering*; Nielsen, M. B., Ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2013.
- Otto, S.; Furlan, R. L. .; Sanders, J. K. . *Drug Discov. Today* **2002**, *7*, 117–125.
- Otto, S.; Furlan, R. L. E.; Sanders, J. K. M. *Science* **2002**, *297*, 590–593.
- Ouadahi, K.; Sbagoud, K.; Allard, E.; Larpent, C. *Nanoscale* **2012**, *4*, 727–732.
- Papayannopoulos, I. A. *Mass Spectrom. Rev.* **1995**, *14*, 49–73.
- Park, K. K.; Park, J. W.; Hamilton, A. D. *Org. Biomol. Chem.* **2009**, *7*, 4225–4232.
- Pearlman, W. H.; Crepy, O. *J. Biol. Chem.* **1967**, *242*, 182–189.
- Peters, T. *All About Albumin: Biochemistry, Genetics, and Medical Applications*; Academic Press: San Diego, CA, 1996.
- Petry, S.; Ben Ali, Y.; Chahinian, H.; Jordan, H.; Kleine, H.; Müller, G.; Carrière, F.; Abousalham, A. *J. Lipid Res.* **2005**, *46*, 603–614.
- Pezzato, C.; Lee, B.; Severin, K.; Prins, L. J. *Chem. Commun. (Camb)*. **2013**, *49*, 469–471.
- Pocci, M.; Alfei, S.; Castellaro, S.; Lucchesini, F.; Milanese, M.; Bertini, V. *Polym. J.* **2013**, *45*, 1146–1152.
- Podlaha, O.; Töregård, B. *J. Chromatogr. A* **1989**, *482*, 215–226.

- Prabakaran, S.; Tepp, W.; DasGupta, B. . *Toxicon* **2001**, *39*, 1515–1531.
- Prem P., B.; Katsushi, S.; Takuya, U.; Kunio, T. *Int. J. Biochem.* **1989**, *21*, 857–862.
- Prochazkova, S.; Vårum, K. M.; Ostgaard, K. *Carbohydr. Polym.* **1999**, *38*, 115–122.
- Quarrell, R.; Claridge, T. D. W.; Weaver, G. W.; Lowe, G. *Mol. Divers.* **1996**, *1*, 223–232.
- Raclot, T.; Holm, C.; Langin, D. *J. Lipid Res.* **2001**, *42*, 2049–2057.
- Rainwater, J. C.; Anslyn, E. V. *Chem. Commun. (Camb)*. **2010**, *46*, 2904–2906.
- Raju, B.; Kogan, T. P. *Tetrahedron Lett.* **1997**, *38*, 4965–4968.
- Ramaiah, D.; Eckert, I.; Arun, K. T.; Weidenfeller, L.; Epe, B. *Photochem. Photobiol.* **2004**, *79*, 99–104.
- Ramaiah, D.; Joy, A.; Chandrasekhar, N.; Eldho, N. V.; Das, S.; George, M. V. *Photochem. Photobiol.* **1997**, *65*, 783–790.
- Ramsey, B. L.; Westphal, U. *Biochim. Biophys. Acta* **1978**, *529*, 115–122.
- Rangin, M.; Basu, A. *J. Am. Chem. Soc.* **2004**, *126*, 5038–5039.
- Reed, R. G.; Feldhoff, R. C.; Peters, T. *Biochemistry* **1976**, *15*, 5394–5398.
- Remuzon, P.; Bouzard, D.; DiCesare, P.; Essiz, M.; Jacquet, J.-P.; Nicolau, A.; Martel, A.; Menard, M.; Bachand, C. *Tetrahedron* **1995**, *51*, 9657–9670.
- Rhee, E. P.; Cheng, S.; Larson, M. G.; Walford, G. A.; Lewis, G. D.; McCabe, E.; Yang, E.; Farrell, L.; Fox, C. S.; O'Donnell, C. J.; Carr, S. A.; Vasan, R. S.; Florez, J. C.; Clish, C. B.; Wang, T. J.; Gerszten, R. E. *J. Clin. Invest.* **2011**, *121*, 1402–1411.
- Richieri, G. V.; Anel, A.; Kleinfeld, A. M. *Biochemistry* **1993**, *32*, 7574–7580.
- Riddles, P. W.; Blakeley, R. L.; Zerner, B. *Methods Enzymol.* **1983**, *91*, 49–60.
- Ritchie, C. D.; Sager, W. F. *Progress in Physical Organic Chemistry*; Cohen, S. G.; Streitwieser, A.; Taft, R. W., Eds.; Progress in Physical Organic Chemistry; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1964; Vol. 2.

- Roda, A.; Cappelleri, G.; Aldini, R.; Roda, E.; Barbara, L. *J. Lipid Res.* **1982**, *23*, 490–495.
- Rodriguez, J. A.; Ben Ali, Y.; Abdelkafi, S.; Mendoza, L. D.; Leclaire, J.; Fotiadu, F.; Buono, G.; Carrière, F.; Abousalham, A. *Biochim. Biophys. Acta* **2010**, *1801*, 77–83.
- Ros-Lis, J. V.; García, B.; Jiménez, D.; Martínez-Máñez, R.; Sancenón, F.; Soto, J.; Gonzalvo, F.; Valdecabres, M. C. *J. Am. Chem. Soc.* **2004**, *126*, 4064–4065.
- Ros-Lis, J. V.; Martínez-Máñez, R.; Soto, J. *Chem. Commun.* **2002**, 2248–2249.
- Ros-Lis, J. V.; Martínez-Máñez, R.; Soto, J.; Villaescusa, L. A.; Rurack, K. *J. Mater. Chem.* **2011**, *21*, 5004.
- Sadao, A. Production of 3-(N-alkylamino)-acylanilide. JPH08143523, 1996.
- Sakai, S.; Shigemasa, Y.; Sasaki, T. *Tetrahedron Lett.* **1997**, *38*, 8145–8148.
- Sauer, M.; Hofkens, J.; Enderlein, J. In *Handbook of Fluorescence Spectroscopy and Imaging*; Wiley-VCH: Weinheim, Germany, 2011; pp. 1–30.
- Saxberg, B. E. H.; Kowalski, B. R. *Anal. Chem.* **1979**, *51*, 1031–1038.
- Schiano, V.; Laurenzano, E.; Brevetti, G.; De Maio, J. I.; Lanero, S.; Scopacasa, F.; Chiariello, M. *Clin. Nutr.* **2008**, *27*, 241–247.
- Schiller, A.; Wessling, R. A.; Singaram, B. *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 6457–6459.
- Sedlak, J.; Lindsay, R. H. *Anal. Biochem.* **1968**, *25*, 192–205.
- Sena, M. M.; Trevisan, M. G.; Poppi, R. J. *Talanta* **2006**, *68*, 1707–1712.
- Shafeekh, K. M.; Rahim, M. K. A.; Basheer, M. C.; Suresh, C. H.; Das, S. *Dye. Pigment.* **2013**, *96*, 714–721.
- Sharma, M.; Tiwari, M.; Chandra, R. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5347–5350.
- Shi, B.; Greaney, M. F. *Chem. Commun.* **2005**, 886.
- Shi, B.; Stevenson, R.; Campopiano, D. J.; Greaney, M. F. *J. Am. Chem. Soc.* **2006**, *128*, 8459–8467.

- Sieber, P. *Tetrahedron Lett.* **1987**, 28, 6147–6150.
- Silvestri, F.; Irwin, M. D.; Beverina, L.; Facchetti, A.; Pagani, G. A.; Marks, T. J. *J. Am. Chem. Soc.* **2008**, 130, 17640–17641.
- Simard, J. R.; Zunszain, P. A.; Hamilton, J. A.; Curry, S. *J. Mol. Biol.* **2006**, 361, 336–351.
- Siuti, N.; Kelleher, N. L. *Nat. Methods* **2007**, 4, 817–821.
- Skoog, D. A.; Holler, J. F.; Crouch, S. R. *Principles of Instrumental Analysis*; 6th ed.; Thomson Brooks/Cole: Belmont, CA, 2007.
- Small, D. M. *Annu. Rev. Nutr.* **1991**, 11, 413–434.
- Springsteen, G.; Wang, B. *Tetrahedron* **2002**, 58, 5291–5300.
- Squire, P. G.; Moser, P.; O’Konski, C. T. *Biochemistry* **1968**, 7, 4261–4272.
- Sreejith, S.; Carol, P.; Chithra, P.; Ajayaghosh, A. *J. Mater. Chem.* **2008**, 18, 264.
- Stitzel, S. E.; Aernecke, M. J.; Walt, D. R. *Annu. Rev. Biomed. Eng.* **2011**, 13, 1–25.
- Sudlow, G.; Birkett, D. J.; Wade, D. N. *Mol. Pharmacol.* **1975**, 11, 824–832.
- Sudlow, G.; Birkett, D. J.; Wade, D. N. *Mol. Pharmacol.* **1976**, 12, 1052–1061.
- Sułkowska, A. *J. Mol. Struct.* **2002**, 614, 227–232.
- Sun, Q.; Qian, J.; Tian, H.; Duan, L.; Zhang, W. *Chem. Commun. (Camb)*. **2014**, 50, 8518–8521.
- Takehara, K.; Yuki, K.; Shirasawa, M.; Yamasaki, S.; Yamada, S. *Anal. Sci.* **2009**, 25, 115–120.
- Tamaoku, K.; Murao, Y.; Akiura, K.; Ohkura, Y. *Anal. Chim. Acta* **1982**, 136, 121–127.
- Terpetsching, E.; Szmecinski, H.; Lakowicz, J. R. *Anal. Chim. Acta* **1993**, 282, 633–641.
- Thomas, M. C.; Mitchell, T. W.; Harman, D. G.; Deeley, J. M.; Murphy, R. C.; Blanksby, S. J. *Anal. Chem.* **2007**, 79, 5013–5022.
- Thumser, A. E. A.; Buckland, A. G.; Wilton, D. C. *J. Lipid Res.* **1998**, 39, 1033–1038.

- Treibs, A.; Jacob, K. *Angew. Chemie Int. Ed. English* **1965**, *4*, 694–694.
- Varmuza, K.; Filzmoser, P. *Introduction to Multivariate Statistical Analysis in Chemometrics*; CRC Press: Boca Raton, FL, 2009.
- Verheyen, E.; Delain-Bioton, L.; Van Der Wal, S.; El Morabit, N.; Barendregt, A.; Hennink, W. E.; Van Nostrum, C. F. *Macromol. Biosci.* **2010**, *10*, 1517–1526.
- Voet, D.; Voet, J. G.; Pratt, C. W. *Fundamentals of Biochemistry*; 3rd ed.; Wiley & Sons: New York, 2008; p. Chapter 9.
- Voigtritter, K.; Ghorai, S.; Lipshutz, B. H. *J. Org. Chem.* **2011**, *76*, 4697–4702.
- Volkova, K. D.; Kovalska, V. B.; Losytskyy, M. Y.; Bento, A.; Reis, L. V.; Santos, P. F.; Almeida, P.; Yarmoluk, S. M. *J. Fluoresc.* **2008**, *18*, 877–882.
- Volkova, K. D.; Kovalska, V. B.; Tatarets, A. L.; Patsenker, L. D.; Kryvorotenko, D. V.; Yarmoluk, S. M. *Dye. Pigment.* **2007**, *72*, 285–292.
- Wada, T.; Nishijima, M.; Fujisawa, T.; Sugahara, N.; Mori, T.; Nakamura, A.; Inoue, Y. *J. Am. Chem. Soc.* **2003**, *125*, 7492–7493.
- Wallace, K. J.; Hanes, R.; Anslyn, E.; Morey, J.; Kilway, K. V.; Siegel, J. *Synthesis (Stuttg.)* **2005**, *2005*, 2080–2083.
- Warwel, S.; Borgdorf, R.; Brühl, L. *Biotechnol. Lett.* **1999**, *21*, 431–436.
- Weber, G.; Daniel, E. *Biochemistry* **1966**, *5*, 1900–1907.
- Wei, G.; Wang, S.; Renshaw, K.; Thompson, M. E.; Forrest, S. R. *ACS Nano* **2010**, *4*, 1927–1934.
- Wells, J. M.; McLuckey, S. A. *Methods Enzymol.* **2005**, *402*, 148–185.
- Wentrup, C.; Lorencak, P. *J. Am. Chem. Soc.* **1988**, *110*, 1880–1883.
- West, R.; Powell, D. L. *J. Am. Chem. Soc.* **1963**, *85*, 2577–2579.
- Wright, A. T.; Griffin, M. J.; Zhong, Z.; McCleskey, S. C.; Anslyn, E. V.; McDevitt, J. T. *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 6375–6378.
- Xiao, X.; Wei, G.; Wang, S.; Zimmerman, J. D.; Renshaw, C. K.; Thompson, M. E.; Forrest, S. R. *Adv. Mater.* **2012**, *24*, 1956–1960.

- Yagi, S.; Hyodo, Y.; Hirose, M.; Nakazumi, H.; Sakurai, Y.; Ajayaghosh, A. *Org. Lett.* **2007**, *9*, 1999–2002.
- Yan, Z.; Guang, S.; Xu, H.; Liu, X. *Analyst* **2011**, *136*, 1916–1921.
- Yates, F. E.; Urquhart, J. *Physiol. Rev.* **1962**, *42*, 359–433.
- Yousefinejad, S.; Hemmateenejad, B. *Drug Test. Anal.* **2012**, *4*, 507–514.
- Zamora-Olivares, D.; Kaoud, T. S.; Jose, J.; Ellington, A.; Dalby, K. N.; Anslyn, E. V. *Angew. Chemie* **2014**, *126*, 14288–14292.
- Zhang, Y.; Askim, J. R.; Zhong, W.; Orlean, P.; Suslick, K. S. *Analyst* **2014**, *139*, 1922–1928.
- Zhang, Y.; Yue, X.; Kim, B.; Yao, S.; Bondar, M. V.; Belfield, K. D. *ACS Appl. Mater. Interfaces* **2013**, *5*, 8710–8717.
- Zhang, Y.-Z.; Zhou, B.; Liu, Y.-X.; Zhou, C.-X.; Ding, X.-L.; Liu, Y. *J. Fluoresc.* **2007**, *18*, 109–118.
- Zhong, Y.; Xu, Y.; Anslyn, E. V. *European J. Org. Chem.* **2013**, *2013*, 5017–5021.
- Zhou, H.; Baldini, L.; Hong, J.; Wilson, A. J.; Hamilton, A. D. *J. Am. Chem. Soc.* **2006**, *128*, 2421–2425.
- Zhou, H.; Jiao, P.; Yang, L.; Li, X.; Yan, B. *J. Am. Chem. Soc.* **2011**, *133*, 680–682.

Vita

Katharine Louise Diehl was born in Augusta, Georgia. She completed her undergraduate degree at the University of North Carolina at Chapel Hill in May 2010, receiving a Bachelor of Science degree in chemistry and a minor in German. As an undergraduate, Katharine received several awards from the university, including the Johnston Scholarship, the James M. Maguire Memorial Award, and the David L. Stern Scholar Award. She was also inducted into Phi Beta Kappa. Outside of class, Katharine worked in the lab of Dr. Marcey Waters for two years and also studied abroad in Berlin for a summer. In August 2010, Katharine entered the chemistry Ph.D. program at the University of Texas at Austin. She joined the research group of Dr. Eric Anslyn. As a graduate student, Katharine received the Henze Teaching Excellence Award and the Dorothy B. Banks Fellowship from the chemistry department. She also served on the American Chemical Society's 2014 Graduate Student Symposium Planning Committee. She will receive her doctorate in organic chemistry in May 2015. After graduation, Katharine will begin a post-doctoral research position at Princeton University in the research group of Dr. Tom Muir.

Permanent email address: katharineldiehl@gmail.com

This dissertation was typed by Katharine Louise Diehl.