The Dissertation Committee for Alexandra Moore Gade Certifies that this is the approved version of the following dissertation:

Sensing Approaches for the Discrimination of Small Molecules and Multivalent Analytes

Committee:

Eric V. Anslyn, Supervisor

Jonathan Sessler

Andrew Ellington

David Hoffman

Richard Aldrich

Sensing Approaches for the Discrimination of Small Molecules and Multivalent Analytes

by

Alexandra Moore Gade, BS

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

December 2014

Dedication

This dissertation is dedicated to the memory of my grandfather, Dr. Carl E. Moore, the first chemist I ever loved.

Acknowledgements

This work would not have been possible without the support of my advisor, Dr. Eric Anslyn. I am grateful to have been given the opportunity to do research in his group and to have observed his great passion for his job. Under his advisement, I was able to assume a leadership role and develop my independent research skills. I am indebted to him for believing in me, especially at times when I did not believe in myself. Without his guidance and support throughout the difficult times of my graduate career, this dissertation would not have come to fruition.

This research was also made possible by the generous donation of resources and training provided by the Ellington lab.

Sensing Approaches for the Discrimination of Small Molecules and Multivalent Analytes

Alexandra Moore Gade, PhD

The University of Texas at Austin, 2014

Supervisor: Eric V. Anslyn

Differential arrays, composed of receptors that are capable of generating unique patterns of responses, have been shown to be useful for discrimination of molecular analytes. Herein, differential arrays have been developed and utilized for the discrimination of small molecule and multivalent biological analytes using cross-reactive receptors. A variety of carboxylate-binding guanidinium-based receptors were tested for their ability to discriminate carboxylate enantiomers. Lanthanide complexes showed the most promising enantiodifferentiation. A dynamic receptor for multivalent biological analytes was developed using self-assembling components designed to target cancer cell lines in a cross-reactive manner. Using this differential array, cancer cell lines of different tissue origin were classified using principal component analysis. The receptors in the array responded to targets as hypothesized but also behaved in a cross-reactive manner that allowed for analyte differentiation. The classification response of the array was reproducible. Boronic acid receptors and receptor arrays were also developed for discrimination of cell surface glycans. In this work, the success of cross-reactive receptors with designed components in differential sensing for small molecules as well as complex multivalent analytes is demonstrated.

Table of Contents

List of Tables

List of Figures

List of Schemes

Chapter 1: Small Molecule Sensing

1.1 Introduction

In an effort to understand the natural world from the perspective of atoms and molecules, scientists have developed a variety of techniques to separate, identify, and quantify chemicals. Wet laboratory and instrumental techniques exist for both qualitative analysis to identify chemical components and quantitative analysis to ascertain the amount of a chemical component. These techniques either use direct determination of a chemical *via* its physical properties, such as its absorbance of light using UV/visible spectroscopy, or transformation of a chemical to indicate its presence and/or amount, such as oxidation using Tollens' reagent in the silver mirror test for aldehydes.

As indicated above, the interaction of molecules with one another is important to qualifying and quantifying a particular chemical. Beyond such analysis, chemists also strive to understand the manner in which molecules interact with one another in order to manipulate those interactions to some end, such as synthesis of new molecules *via* the creation of covalent bonds. Non-covalent interactions such as ion pairing, dipole-dipole, dipole-induced dipole, hydrogen bonding (H-bonding), van der Waals and London dispersion forces, pi-interactions, and hydrophobic interactions, as well as reversible covalent bonds are important to molecular recognition.[1] Designed receptors often combine several of these forces for the detection of molecular analytes.

1.1.1 Designed Receptors for Small Molecules

Inspired by nature, researchers have designed synthetic receptors based on molecular recognition in biological systems. In 1894, Emil Fischer introduced the "Lock-and-Key" model to describe enzyme-ligand interactions[2]:

> I would like to say that enzyme and glucoside have to fit together like lock and key in order to exert a chemical effect on each other.

He later expounded:

> The geometrical structure exerts such a profound influence on the playing of the chemical affinities, that it appeared legitimate to me to compare the interacting molecules with key and lock.
>
> This had led to the hypothesis, that there must be a similarity in the molecular configuration between the enzymes and their object of attack, if reaction is to take place. To make this thought more perspicuous, I have used the picture of lock and key.

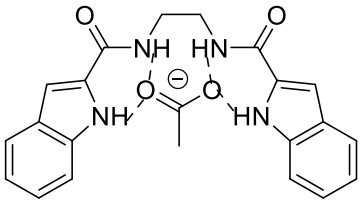
Enzymes interact with substrates with high specificity and selectivity; in Fischer's analogy, an enzyme behaves as a lock that will only fit a particular ligand key that is complementary to the lock's structure.

Artificial receptors for molecular recognition, also known as hosts, can be tailored to particular target analytes, also known as guests, using the lock-and-key principle. Such hosts are designed and synthesized to possess complementary structures to a particular target guest in order to bind with high specificity and selectivity. Complementarity in "geometrical structure" and similarity in "molecular configuration" are the molecular features of the host that allow for attractive, reversible interactions with the guest, and a researcher can manipulate such interactions through host structure design.

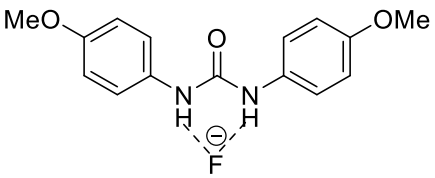
Using molecular recognition as an analytical technique to identify or quantify an analyte requires that the host/guest interaction results in a signal to indicate that binding between the species has occurred. When binding results in a measurable change to the physical properties of one or both binding partners, transduction of the binding event can be achieved using spectroscopic, electrochemical, or calorimetric techniques. Signal transduction is achieved by observing changes to the portions of the host or guest directly

involved in recognition or by synthetically appending the host structure with a component that can generate an optical or electrochemical signal.

For example, shifts in the ppm values of nuclear magnetic resonance (NMR) signals due to changes in the electronic environment upon molecular interaction can be used as a direct signal of a binding event. Furthermore, identification of the atoms involved in a binding interaction can be determined from the structural information provided by the NMR spectrum. A recent report by RoyChowdhury *et al*. explored the anion binding properties of indole- and urea-based receptors of different cleft size (Figure 1.1).[3] Titration of selected monoanions with the indole receptor monitored by ^{1}H-NMR led to broadening and downfield shifts of the aromatic and aliphatic -NH protons. This was explained by anion recognition through H-bonding, resulting in a change of electron distribution of the bound species and the observed proton shifts. In contrast to the indole receptor, the urea-based receptor showed no signal response to acetate by ^{1}H-NMR. Selectivity of the receptors for different anions can be explained by differences in receptor design that affect receptor shape and size, leading to different H-bonding interactions.



1.1



1.2

Figure 1.1 Two anion binding receptors of different cleft sizes from RoyChowdhury *et al*. and their binding of different anions through H-bonding.[3]

Isothermal titration calorimetry (ITC) is another technique for direct observation of a binding event, but it does not give information about atom-level interactions. Instead, the thermodynamic driving forces for binding can be determined by measurement of heat changes, expressed as ΔH enthalpy values, upon titration of binding partners. These data are iteratively fit to give a binding stoichiometry and equilibrium association constant, K_{ass}, from which the entropy of the interaction, ΔS, can be calculated using the Gibb's free energy equation (Equation 1.1):

$$\Delta G = \Delta H - T\Delta S \qquad \text{(Equation 1.1)}$$

When a binding interaction cannot be directly transduced into an observable signal, it is necessary to append a molecule to the host that is capable of generating a response upon the binding event. The addition of chromophores, fluorophores, and electrochemically sensitive groups to hosts give signals that can be modulated upon the binding interaction, resulting in transduction of the binding event. Fluorogenic tripodal anion receptors based on 1,3,5-triethylbenzene substituted with pyrrole and thiourea anion-binding groups have been reported by Anzenbacher *et al*. (Figure 1.2).[4] These receptors are modified with fluorophores, including fluorescein (**1.2c**) and naphthalimide (**1.2b and 1.3c**), and exhibit strong fluorescence in the presence of anions. The fluorescence turn-on occurs in response to the binding of anions in the cavity formed by the alternately arranged hydrogen-bond donor arms and serves as a reporter of the binding event, although the fluorophore is not directly involved in anion recognition. The higher affinity for C_3-symmetrical anions, such as phosphonate and phosphate, is likely due to the bowl-shaped design of the receptor cavity and makes these fluorophore-appended receptors useful for optical signaling of the hydrolysis products of sarin nerve gas.

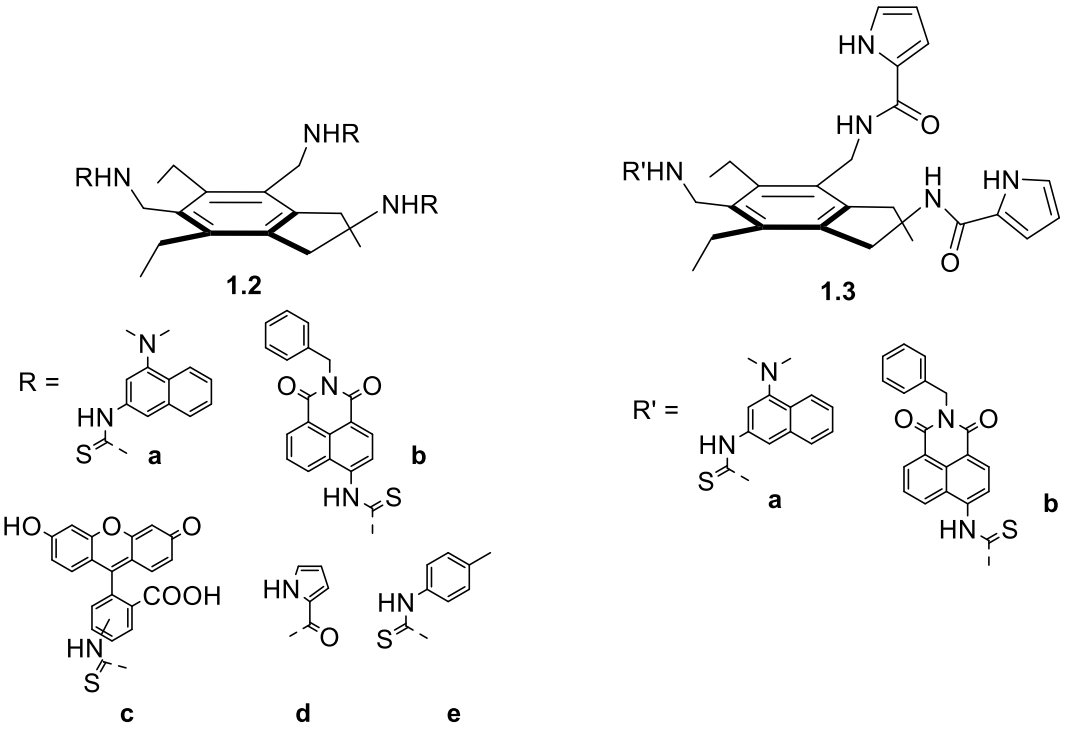


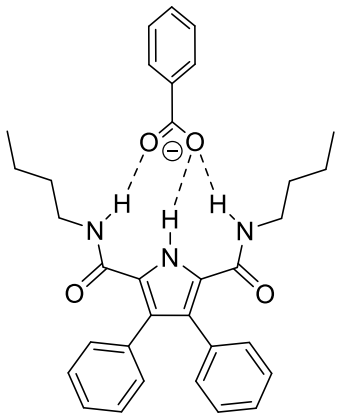
Figure 1.2 Tripodal receptors derived from 1,3,5-triethylbenzene and appended with fluorogenic groups for anion detection, prepared by Anzenbacher *et al*.[4]

1.1.1.1 Synthetic Receptors for Anion Sensing

Anions are important targets for molecular sensing owing to their prominence in biological functions that depend on, for example, anion transport;[5] they are also found in industrial processes and are present in the environment as pollutants.[6] While abiotic sensors for cationic analytes have been well developed in the past four decades, the development of analogous anion sensors was slow after first being introduced in the late 1960s.[7] This disparity is due to the greater difficulty of molecular recognition of anions, compared to cations and neutral molecules, due to their lower charge density, diverse geometries and charge delocalized structures, and large negative free energies of solvation.[8,9] Challenged by such difficulty, research in anion sensing has grown, and a

wide variety of receptor functionalities, for example pyrrole-,[8] urea-,[10] and guanidinium-based[11] ligands, have been published.

Pyrrole contains a single H-bond donor, but incorporation of this motif into a macrocycle, such as calix[n]pyrrole or porphyrin, allows for arrangement of multiple H-bond donors in a three-dimensional structure that is favorable for anion binding.[12] Acyclic receptors using pyrrole have been synthesized that show anion binding selectivity. Gale *et al*. prepared 2,5-bis-amidopyrroles that show selectivity for oxoanions and that have higher affinity than the monoamido analogs. All of the hydrogen bond donor groups in the diamidopyrroles are involved in oxoanion binding (Figure 1.3).[13]



1.4

Figure 1.3 A complex of a 2,5-bis-amidopyrrole with benzoate adapted from the crystal structure published by Gale *et al*.[13] This structure illustrates the contribution of all three H-bond donors in the receptor in complex formation.

Urea groups can form strong hydrogen-bond complexes with oxoanions. Amides in proteins assist in formation of tertiary structures by interactions between N-H and C=O groups. Urea, which contains a second N-H group bonded to an amide carbonyl, acts as a H-bond donor when neutral, and the adjacent N-H functionality provides directional and

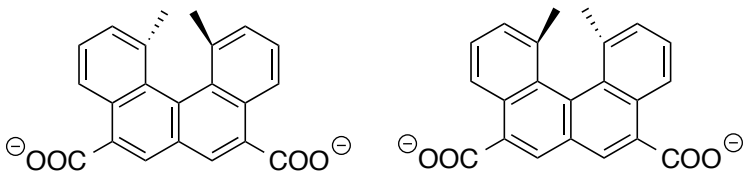
chelating properties for anion coordination.[14] Substitution of the carbonyl oxygen with sulfur weakens the N-H bond and increases its H-bond donating ability. Martinez-Máñez *et al*. synthesized colorimetric sensors containing H-bond donating groups appended to an azodye.[14-15] Of the amide, carbamate, urea, and thiourea derivatives, only the latter two were able to bind oxoanions significantly, presumably due to the better anion coordinating properties of urea.

1.1.1.2 Receptors for Enantiomeric Discrimination

The ability to discriminate enantiomers is an important area of research, especially in drug development where one enantiomer of a drug can have a positive effect while the opposite enantiomer can be detrimental (e.g. Thalidomide). Enantiomers share the same properties unless they are subjected to an asymmetric environment; as a result, synthetic receptors capable of differentiating chiral molecules are themselves chiral. Enantiodifferentiation with chiral receptors occurs when diastereomeric complexes of chiral receptor and chiral analyte have different stabilities.[16]

Chiral receptors have been developed with structures that span a variety of functional groups and target different types of chiral anions. Chromatography and NMR spectroscopy have been widely adopted to evaluate the ability of a particular receptor to resolve enantiomers. Axially chiral enantiomers (*M*)-**1.5** and (*P*)-**1.5** can be discriminated by ^{1}H-NMR spectroscopy in D_2O using β- or γ-cyclodextrin.[17] (Figure 1.4) The enantiomers of **1.5** can be distinguished by the splitting of the proton signals upon the addition of cyclodextrin, and the signals that correspond to the (*M*)-enantiomer are shifted farther downfield. The association constants for (*M*)-**1.5** with both β- and γ-cyclodextrin are higher than those of (*P*)-**1.5** because complexation with (*P*)-**1.5** results in an enthalpically unfavorable interaction of one carboxylate with the cavity of the

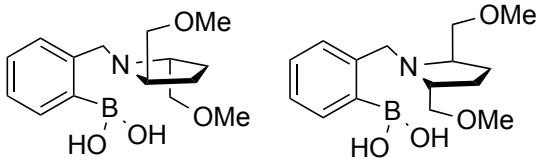
cyclodextrin, while complexation of (*M*)-**1.5** leaves both carboxylates solvent-exposed. For β-cyclodextrin, the differences in the free energy changes of the diastereomeric complexes (ΔΔG) is 5.2 kJ/mol.



(*M*)-**1.5** (*P*)-**1.5**

Figure 1.4 Helical dicarboxylate enantiomers that can be discriminated with cyclodextrins through the use of NMR, from reference [17].

Colorimetric techniques for differentiation of chiral analytes are gaining popularity. For example, Zhu *et al*. prepared a chiral boronic acid receptor for the differentiation of enantiomers of α-hydroxy acids.[18] Using absorbance spectroscopy, host (*R,R*)-**1.6** displayed a preference for complexation with D-phenyllactic acid via reversible covalent boronate ester formation, while the (*S,S*)-enantiomer had the opposite selectivity (Figure 1.5).



(*R,R*)-**1.6** (*S,S*)-**1.6**

Figure 1.5 Boronic acid-based receptors for enantiodifferentiation of α-hydroxy acids, prepared by Zhu *et al*.[18]

1.1.2 Array Sensing for Small Molecules

1.1.2.1 Differential Arrays

Although biological systems use highly specific interactions for recognition, differential sensing is inspired by another type of sensing motif found in nature: biological recognition of analytes by receptors in the nose and tongue. The mammalian senses of taste and smell are produced by receptors that are sensitive to classes of analytes. Differentials arrays utilize cross-reactive receptors, often biased towards a class of analytes, that do not necessarily require specificity or high affinity.[19] Instead, analytes can be discriminated by their diagnostic patterns of interaction with an array of receptors of moderate affinity, where each receptor responds to each analyte to a different degree.[20] Therefore, the fingerprint that results is not necessarily a response of specific recognition interactions between host and guest.[19, 21]

Array sensing is a powerful tool for analyte discrimination, and as such, a variety of chemical noses and tongues have been developed using electrochemical and optical responses for vapor- and solution-based analytes.[21]

1.1.2.2 Indicator Displacement Assays

Popularized by Anslyn *et al*. in 2001,[22] indicator displacement assays (IDAs) are used for colorimetric and/or fluorometric sensing by binding of a pH or solvatochromatic indicator to a receptor through reversible interactions, causing the indicator to change color.[23] The indicator is displaced upon addition of an analyte, which competes for the binding site on the receptor, leading to a UV/Vis absorption or fluorescence signal change. Because synthesis is not required to attach the chromophore to the receptor, indicators can easily be screened and stoichiometrically adjusted to improve the optical

signal, giving IDAs an advantage over sensors that use a non-reversibly attached chromophore.

Through the use of an IDA, Metzger and Anslyn developed a chemosensor specific for citrate in highly competitive media.[24] A binding constant for the tripodal guanidinium-based receptor with citrate was determined to be 3 x 10^5 M^{-1} by indicator displacement in aqueous methanol using a UV/Visible titration. The absorbance intensity from 5-carboxyfluorescein, which was electrostatically bound to the receptor, decreased upon addition of citrate, During both titration of host into indicator and displacement of indicator with citrate an isosbestic point was observed, which indicates that two absorbing species were present. This receptor-indicator pair was used to determine citrate concentrations in complex mixtures like sports drinks, sodas, and juices.

1.1.2.3 Pattern-based enantiodifferentiation

Mammalian sensory systems for taste and smell are based on proteins composed from chiral amino acids, specifically the L-enantiomers. These receptors are thus capable of enantiodifferentiation, which is in fact an important aspect of olfaction and gustation. For example, enantiomers of carvone exhibit different odor qualities, where (+)-carvone smells of caraway, and (-)-carvone smells of spearmint.[25] However, human olfaction exhibits low chiral specificity.[26] Many D-amino acids are reported to taste sweet, and in 2002 Zuker *et al*. showed that a mammalian taste-specific receptor, T1R2+3, was responsive to D-amino acids but did not respond to the standard L-enantiomers.[27]

Differential arrays inspired by mammalian gustation and olfaction have been combined with IDAs for pattern-based chiral differentiation[28] and enantiomeric excess determination.[29] Shabbir *et al*. developed an array of three chiral boronic acids and three pH indicators that was able to generate patterns to discriminate identity, enantiomeric

excess (*ee*), and concentration of structurally similar chiral threo diols.[30] The use of IDAs in enantiodifferentiation is based on the displacement of the indicator being enantioselective.[18] Because the indicator and analyte compete for a binding site on the receptor, differences in affinities of chiral analytes for chiral receptors lead to different extents of indicator displacement, which can manifest as distinct optical signals for the enantiomers. For threo diol enantiomers, Shabbir *et al*. used Principal Component Analysis (PCA) and an Artificial Neural Network (ANN) as the pattern recognition protocols to convert the absorbance values from the array to diagnostic patterns for identity, *ee*, and concentration discrimination.[30]

1.2 Chiral Carboxylate Sensing

1.2.1 Chiral Carboxylate Sensors

Many chiral carboxylate sensors can be found in the literature; such hosts utilize ureas,[31] diamines,[32] amino acids.[32c, 33] and transition metals,[31, 34] incorporated into chiral molecules as recognition units for chiral amino acids and chiral aromatic carboxylates. However, differentiation of other types of chiral carboxylates is still underdeveloped.[35]

Sessler *et al*. used ITC and ^{1}H-NMR to show differential binding of enantiomers of 2-phenylbutyrate to enantiomerically pure calix[4]pyrrole appended with an optically active BINOL group. The (*S*)-receptor exhibited a preference for the (*S*)-carboxylate manifested as a ~10 times higher association constant in MeCN.[36]

Joyce *et al*. reported the formation of enantiomeric complexes with an achiral copper(II) host upon chiral carboxylate coordination.[37] The induced helicity from carboxylate binding generated an exciton signal that was observable by circular

dichroism spectroscopy (CD). The resulting enantiospecific CD spectra were used to identify and discriminate between six pairs of enantiomeric α-chiral carboxylates.

1.2.2 Guanidinium Groups for Carboxylate Sensing

In particular, the guanidinium group has emerged as a successful abiotic sensor for oxoanions, inspired by biogenic examples of carboxylate, phosphate, and sulfate binding by the amino acid arginine, which contains a guanidinium group in its side chain (Figure 1.6).[15]

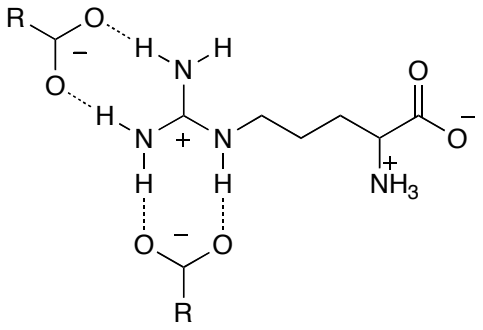


Figure 1.6 The amino acid arginine and its possible interactions with oxoanions

The side chain of arginine is used for oxoanion recognition in proteins, enzymes, and antibodies, and contributes to protein tertiary structure by formation of salt bridges with carboxylates.[35] The guanidinium group binds oxoanions through an ionic hydrogen bond, which is facilitated by three planar nitrogens that share a positive charge and remain protonated over a large pH range (pK_a 12-13). Arginine is able to complement the bidentate structure of oxoanions through two possible binding modes involving four hydrogen bonds (Figure 1.6). This interaction is strong in hydrophobic pockets where recognition commonly occurs in biological systems; however, most synthetic receptors are designed to work in polar solvents where complexation competes with solvation. Therefore, receptors must be designed to provide favorable interactions that outweigh the

free energies of solvation of both binding partners. Such affinity has been achieved through structural changes to the receptor that increase non-covalent interactions like coulombic attraction and H-bonding. Carsten Schmuck designed a receptor for carboxylates in aqueous solvents by adding pyrrole and amide functionality to the guanidinium group for additional H-bond donation, improving carboxylate binding over 100-fold compared to guanidinium chloride.[38]

Incorporation of the guanidinium functionality into a bicyclo[4.4.0]decaline structure increases its solubility in apolar solvents, where higher association constants are observed due to greater H-bonding.[35] The bicyclic structure forces the *syn* conformation, allowing a robust donor/donor-acceptor/acceptor (DD-AA) interaction (Figure 1.7). Schmidtchen *et al*. reported the first bicyclic guanidinium receptor in 1988, observing K_{diss}=7×10^{-6} M for allyl substituted guanidinium and the tetrabutylammonium (TBA) salt of *p*-nitrobenzoate in $CDCl_3$.[39] Modifications to the guanidinium anchor groups that introduced charge neutrality[40,41] and additional hydrogen bond donors[42] were hypothesized to improve oxoanion binding through increased enthalpic attraction. Surprisingly, ITC studies showed that more favorable association constants for oxoanions were due to increased entropies of association, arising from a variety of distinct anion binding modes.[41-42]

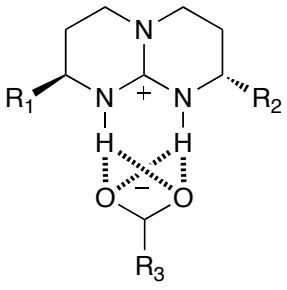


Figure 1.7 The DD-AA arrangement in the *syn*-conformation of bicycloguanidinium

The introduction of stereogenic centers to the bicyclic structure can impart the molecule with C_2 symmetry and allows for chiral oxoanion recognition through binding of guests in an asymmetric environment.[35] Enantiorecognition studies using this structure have focused on chiral carboxylate anions; the first of such receptors extracted *N*-protected sodium tryptophan salts from water by binding the C-terminus through ionic hydrogen bonding at the guanidinium group while naphthoyl side arms engaged in π-stacking.[43] Addition of a crown ether as an anchor group for ammonium binding expanded the receptor's utility to selective extraction of D- or L-enantiomers of phenylalanine, tryptophan, and valine using the (*S*,*S*)- and (*R*,*R*)-receptors respectively[44]. Although Schmidtchen was able to observe diastereoisomeric complexes of racemic carboxylates and a bicyclic guanidinium by ^{1}H-NMR,[45] ITC studies showed no difference in association constants for α-chiral carboxylate enantiomers until the host was incorporated into a macrocycle, where restricting the conformational space of the receptor leads to geometrically unique binding modes for enantiomers and enantiodifferentiation is entropy driven (Figure 1.8).[46,47]

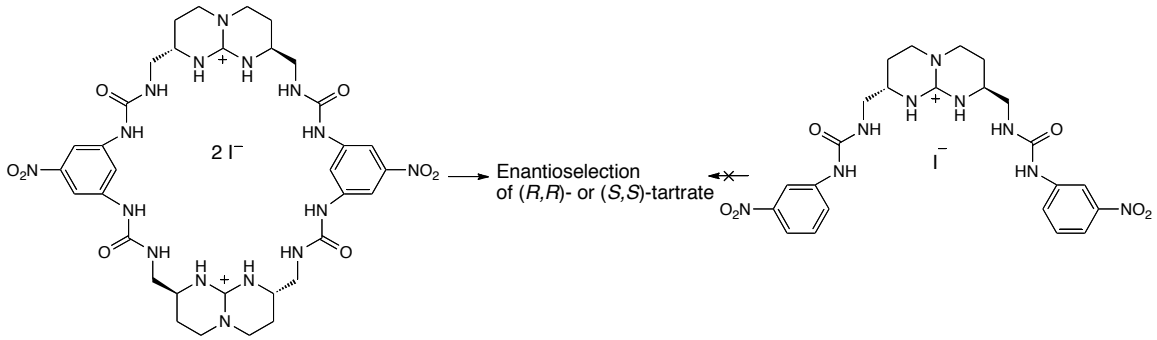
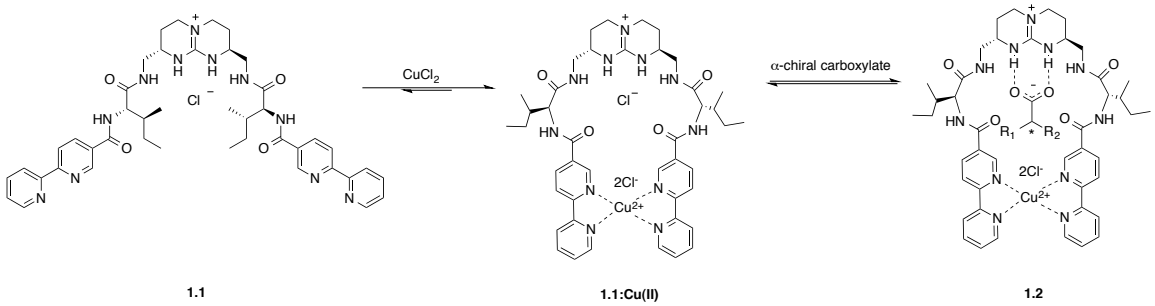


Figure 1.8 Adapted from Jadhav and Schmidtchen: enantioselection occurs only for the macrocycle.[48]

1.3 BIPYRIDYL-FUNCTIONALIZED BICYCLOGUANIDINIUM COMPOUNDS FOR CHIRAL CARBOXYLATE SENSING

This research aimed to develop a colorimetric assay for the detection of chiral carboxylate anions through the use of an IDA. The guanidinium receptor used in this research for α-chiral carboxylate sensing has the bicyclic guanidinium core functionalized with isoleucine arms capped with bipyridyl (BIPY) units (Figure 1.9).



Scheme 1.1 The bicycloguanidinium group containing amino acid arms capped with BIPY units (**1.1**); the receptor complexed with copper (**1.1:Cu(II)**); binding of a chiral α-carboxylate by **1.1:Cu(II)** (**1.2**).

Collins[49] synthesized **1.1** as the chloride salt by first attaching the isoleucine residues to the bicyclic core and then adding the BIPY groups. The isoleucine groups introduce additional hydrogen-bond donor functionality and act as a chiral spacer, while the BIPY groups will chelate copper(II) in a 2:1 BIPY:Cu(II) complex.[50] Complexation by copper ensures that the amino acid arms are conformationally locked around the binding site, creating a sufficiently restricted environment for enantioselection. The IDA used in this research thus consists of a bicyclic guanidinium receptor, Cu(II) chloride, and a UV/Vis indicator Alizarin Red S or Pyrocatechol Violet.

1.3.1 Receptor Assembly with Copper (II)

In order to investigate the complexation of Cu(II) with **1.1**, a titration of **1.1** with $CuCl_2$ was undertaken. An increase in the d-to-d copper transitions was observed, with λ_{max} shifting from 715 nm to 775 nm over the course of the titration (Figure 1.9).

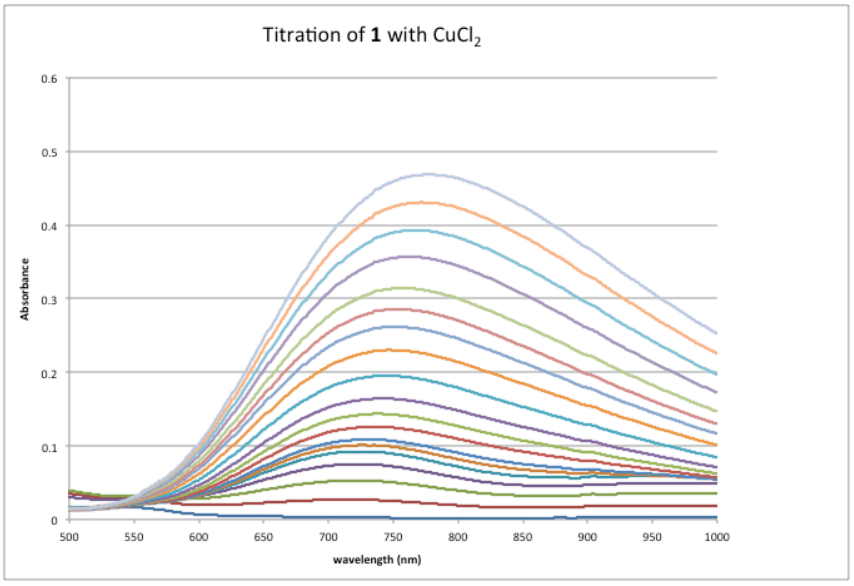


Figure 1.9 Titration of 100mM $CuCl_2$ in 90%DMSO/H_2O (v/v) with 10mM HEPES into 1mM **1.1** in the same buffer.

A control titration of $CuCl_2$ into buffer showed an increase of the d-to-d copper transitions at 800 nm (Figure 1.10).

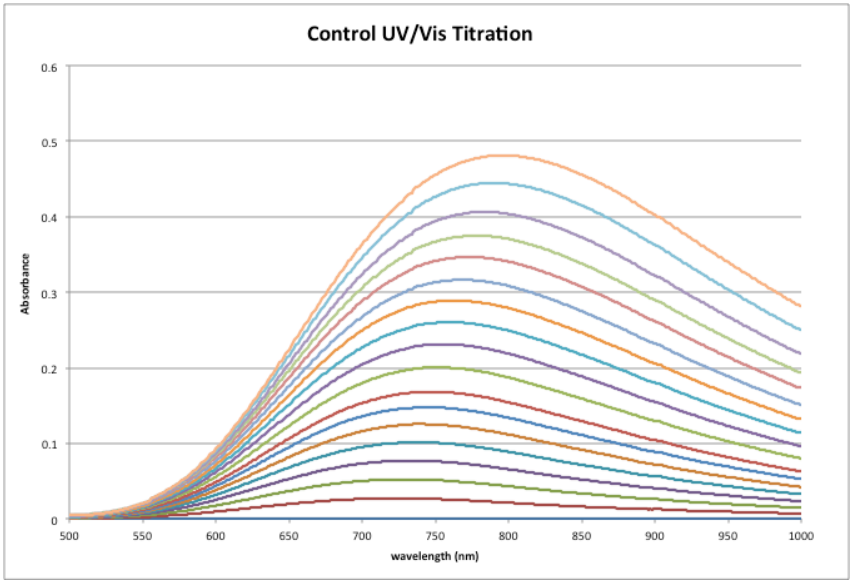


Figure 1.10 Control Titration of 100 mM $CuCl_2$ in 90%DMSO/H_2O (v/v) with 10mM HEPES into the same buffer

The control titration data at 775 nm was fit to a polynomial and then the calculated value at the corresponding equivalents was subtracted from the host titration data at 775 nm. Although the resulting absorbance intensity was very weak, a binding isotherm was obtained, and repetition of the experiment gave consistent results (Figure 1.11).

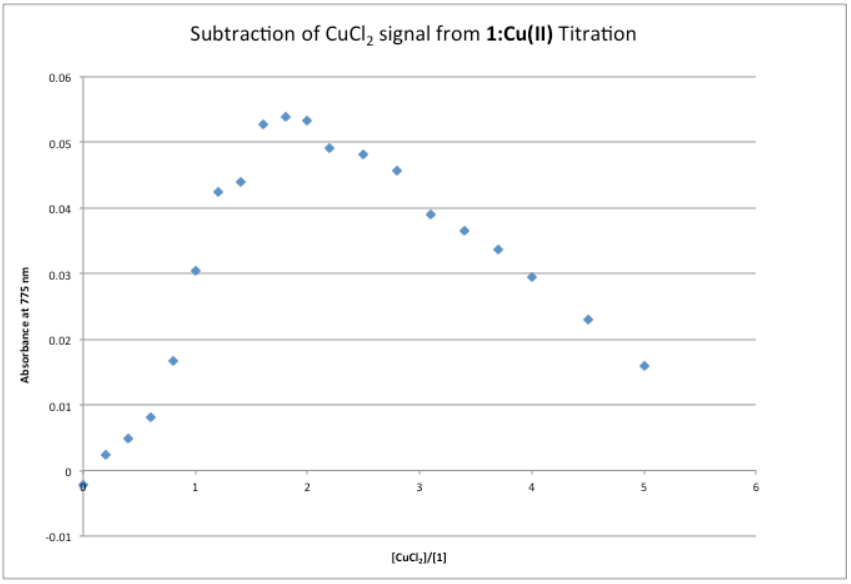


Figure 1.11 Titration of 100 mM $CuCl_2$ in 90% DMSO/H_2O (v/v) with 10 mM HEPES into 1mM **1.1** in the same buffer, subtracted from control data at 775 nm.

The data indicate that the copper sequentially formed 1:1 and 2:1 complexes with the host. Considering the requirement of host design that copper complexation would pre-organize the host into a macrocycle, one equivalent of $CuCl_2$ was used in further experiments with indicator and carboxylate ions.

1.3.2 Receptor Association with Indicators and Chiral Carboxylates

1.3.2.1 Studies with Pyrocatechol Violet

As another control titration, aliquots of $CuCl_2$ were added to a solution of pyrocatechol violet (PV) to examine the interaction of indicator alone with Cu(II). Initially, the free indicator peak at 435 nm decreased, while an increase at 280 nm was observed (Figure 1.12).

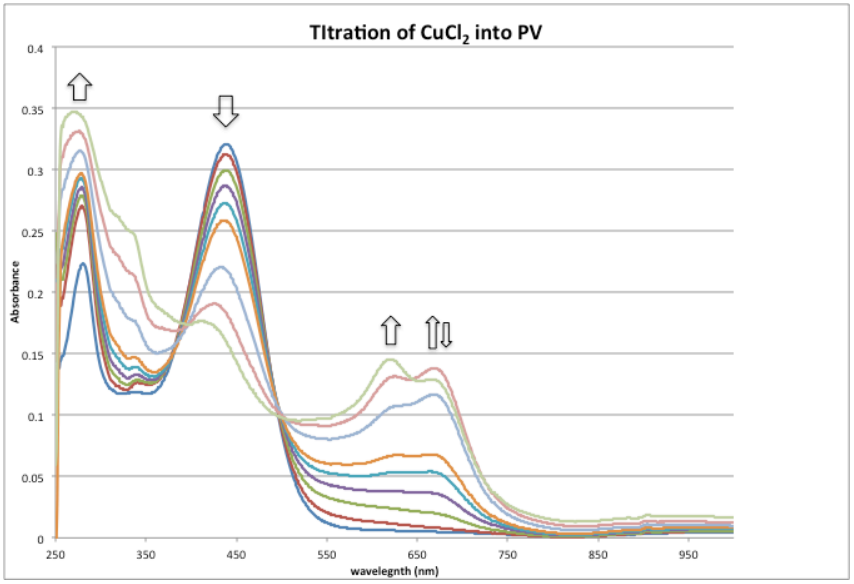


Figure 1.12 Titration of 1.5 mM $CuCl_2$ in 90%DMSO/H_2O (v/v) with 10 mM HEPES into 20 μM Pyrocatechol Violet in the same buffer.

After addition of 0.6 equivalents of $CuCl_2$, a peak at 675 nm appeared, and the solution changed colors from yellow to green. Over the course of the titration, the peak at 675 nm began to decrease, and a peak at 610 nm appeared. Repeated measurements after addition of 3.75 equivalents of $CuCl_2$ indicated that the system continued changing over time. The absorbance values at 675 nm and 440 nm were therefore monitored at one-minute intervals over a period of 15 minutes and 30 minutes, respectively, after addition of one equivalent of $CuCl_2$ to a solution of PV. After these time intervals, the absorbance intensities at both wavelengths continued to decrease and did not stop changing over a period of several days. From these studies it was determined that PV is not an adequate indicator for the conditions used in this system.

An indicator uptake assay was initially employed *in lieu* of an IDA to overcome the kinetic instability of PV in the presence of Cu(II). Addition of excess PV to an equimolar solution of **1.1:Cu(II)** and an achiral oxoanion, phosphate or acetate, resulted

in the appearance of two peaks: the free indicator at 435 nm and the indicator-**1.1:Cu(II)** complex at 675 nm (Figure 1.13).

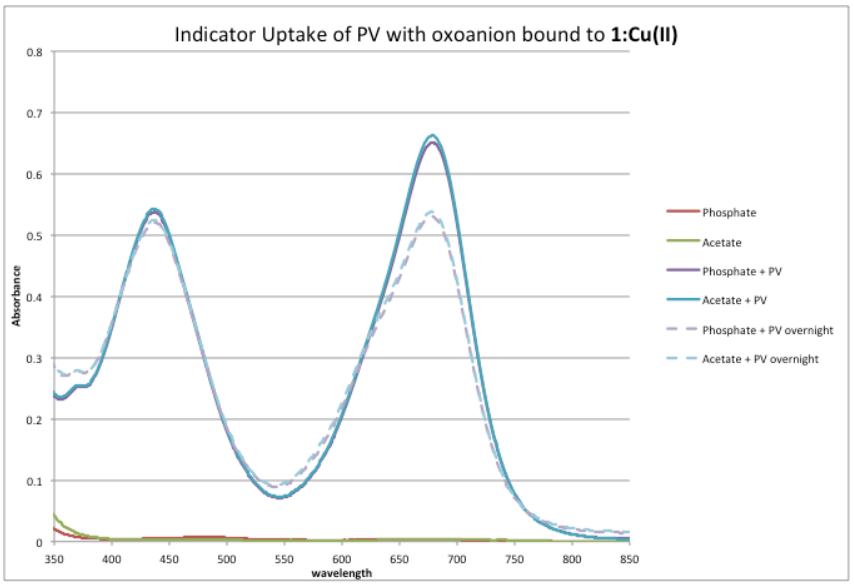


Figure 1.13 Addition of Excess PV (20 μM) in 90% DMSO/H_2O (v/v) with 10 mM HEPES to 75 μM oxoanion (phosphate or acetate) pre-complexed with 75 μM **1.1:Cu(II)** in the same buffer.

Over 12 hours, both peaks diminished in intensity, with the most dramatic decrease for the peak at 675 nm. There was no appreciable difference in the spectral behavior of acetate versus phosphate upon addition of PV, indicating the receptor cannot differentiate between different oxoanion types with the same charge.

Following the same approach, an α-chiral carboxylate, (*R*)-or (*S*)-2-phenylbutyrate, was complexed to **1.1:Cu(II)** and an excess of PV was then added. Again, two peaks at 435 nm and 675 nm appeared, corresponding to the free indicator and the indicator-**1.1:Cu(II)** complex respectively (Figure 1.14).

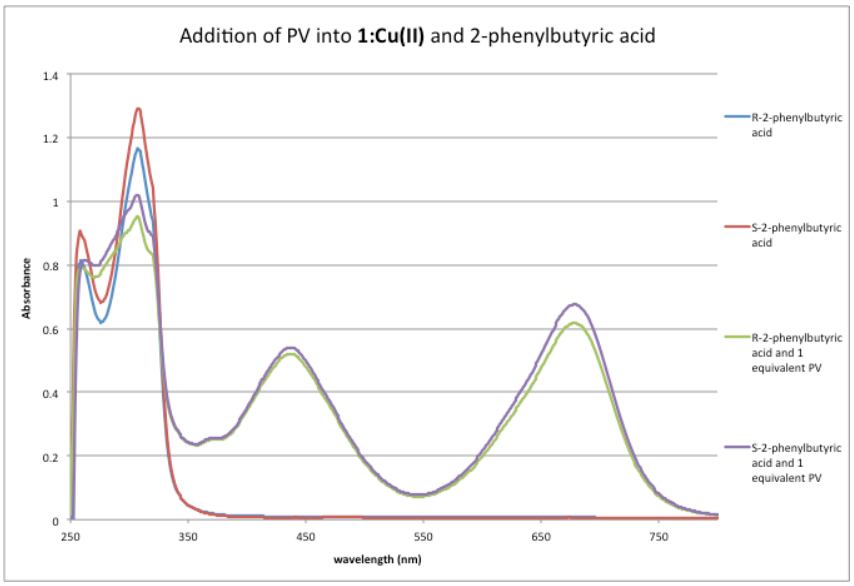


Figure 1.14 Addition of excess PV in 90% DMSO/H_2O (v/v) with 10 mM HEPES to 1.5 mM (*R*)- or (*S*)-2-phenylbutyrate pre-complexed with 1.5 μM **1.1:Cu(II)** in the same buffer.

Initial results showed that **1.1:Cu(II)** gave a larger signal change for the (*S*)-2-phenylbutyrate enantiomer. Because the receptor was only synthesized as the (*S*,*S*)-enantiomer, cross reactivity of the receptor was not studied. An IDA was therefore undertaken to confirm the selectivity of the receptor for (*S*)-2-phenylbutyrate.

1.3.2.2 Studies with Alizarin Red S

Due to the kinetic instability of PV in this system an indicator screening was undertaken; Alizarin Red S (ARS) gave a signal change when bound to the receptor that took only one hour to reach equilibrium and was chosen for further studies. Titration of **1.1:Cu(II)** into a constant concentration of ARS showed a decrease of the free indicator at 435 nm and an increase in the host-bound indicator at 530 nm, with an isosbestic point at 460 nm (Figure 1.15 inset).

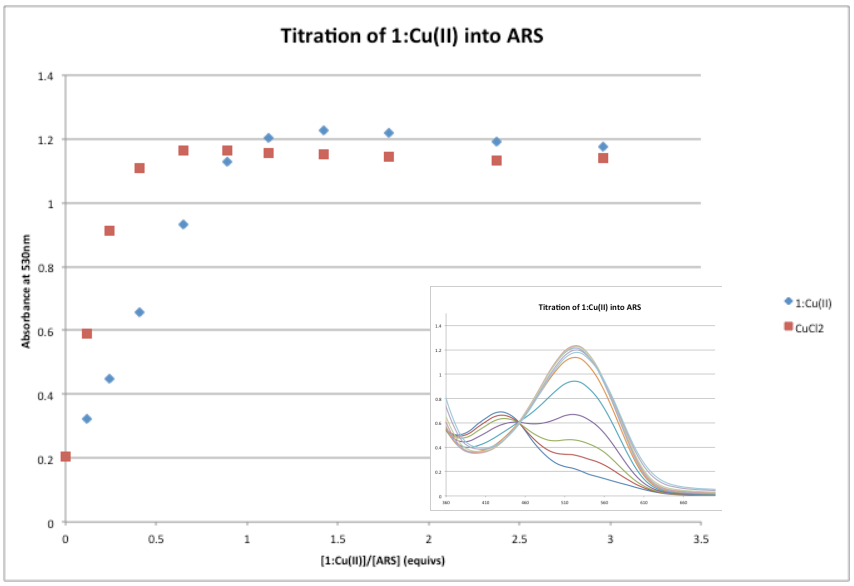


Figure 1.15 Titration of **1.1:Cu(II)** (0-750 μM, blue diamond) in 90% DMSO/H_2O (v/v) with 10 mM HEPES into 250 μM Alizarin Red S in the same buffer.

Although a 1:1 binding curve could not be fit to the data, ARS was estimated to be ideally bound at 1 equivalent of **1.1:Cu(II)** (Figure 1.15, blue diamond). Titration of ARS with a solution of $CuCl_2$ without **1.1** showed ideal ARS binding at 0.5 equivalents of $CuCl_2$, indicating that the presence of the receptor influenced indicator binding (Figure 1.15, red square).

Addition of (*R*)- or (*S*)-phenylbutyrate into **1.1:Cu(II)** complexed with one equivalent of ARS gave a decrease in the indicator-**1.1:Cu(II)** absorbance at 535 nm (Figure 1.16). No increase in the free indicator peak was observed, even after adding 25 equivalents of anion. No appreciable difference in absorbance at 535 nm for (*R*)- or (*S*)-phenylbutyrate was observed (Figure 1.17).

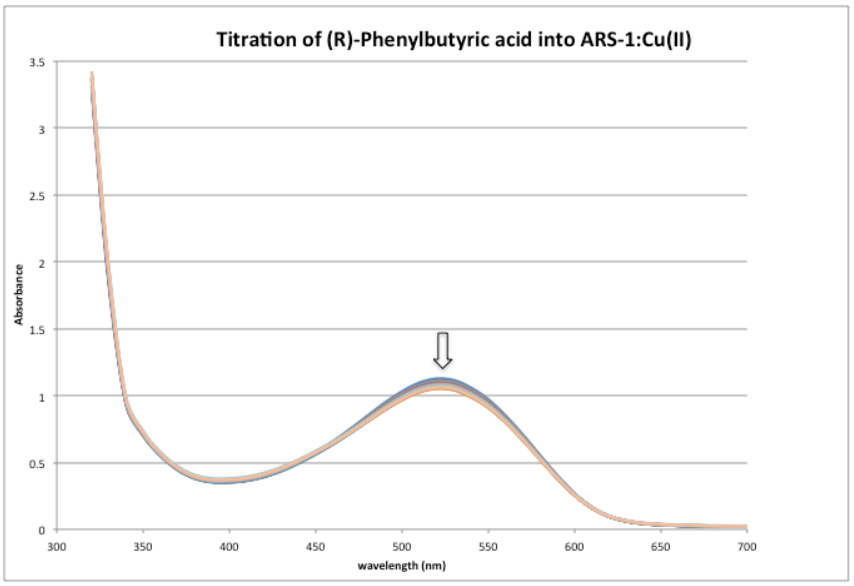


Figure 1.16 Titration of 35.5 mM (*R*)-2-phenylbutyrate in 90%DMSO/H_2O (v/v) with 10 mM HEPES into **ARS-1.1:Cu(II)** in the same buffer.

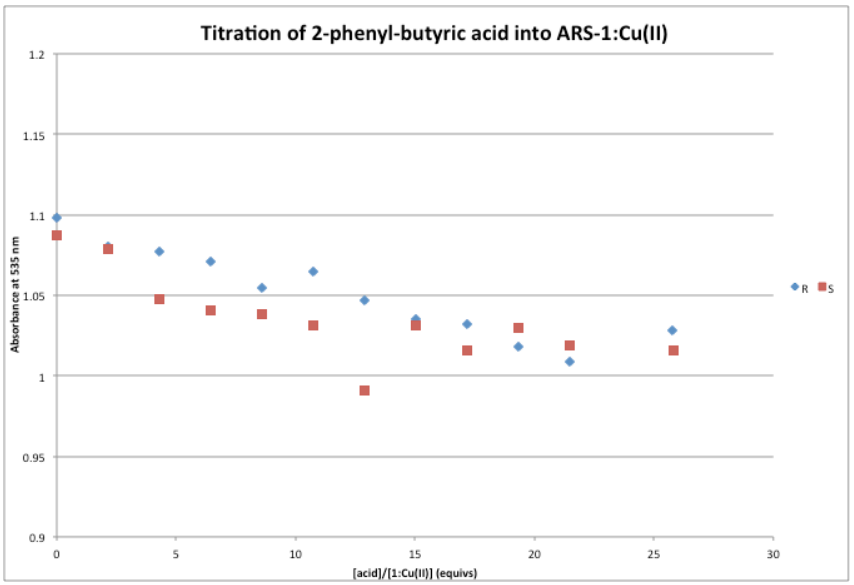


Figure 1.17 Titration of 35.5mM (*R*)- and (*S*)-2-phenylbutyrate in 90% DMSO/H_2O (v/v) with 10 mM HEPES into **ARS-1.1:Cu(II)** in the same buffer at 535 nm.

The same titration with (*R*)- or (*S*)-bromopropanoate also showed a decrease in the 535 nm peak, while a small increase at 435 nm and an isosbestic point at 460 nm were also observed (Figure 1.18). Again, no appreciable difference between absorbance values at 535 nm for the (*R*)- and (*S*)-enantiomers was observed (Figure 1.19).

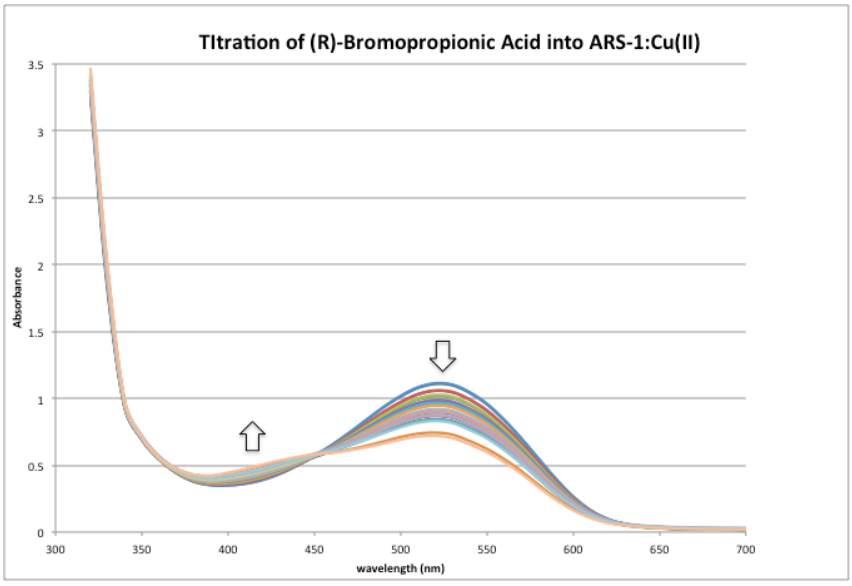


Figure 1.18 Titration of 25.5 mM (*R*)-2-bromopropionate in 90% DMSO/H_2O with 10 mM HEPES into **ARS-1.1:Cu(II)** in the same buffer.

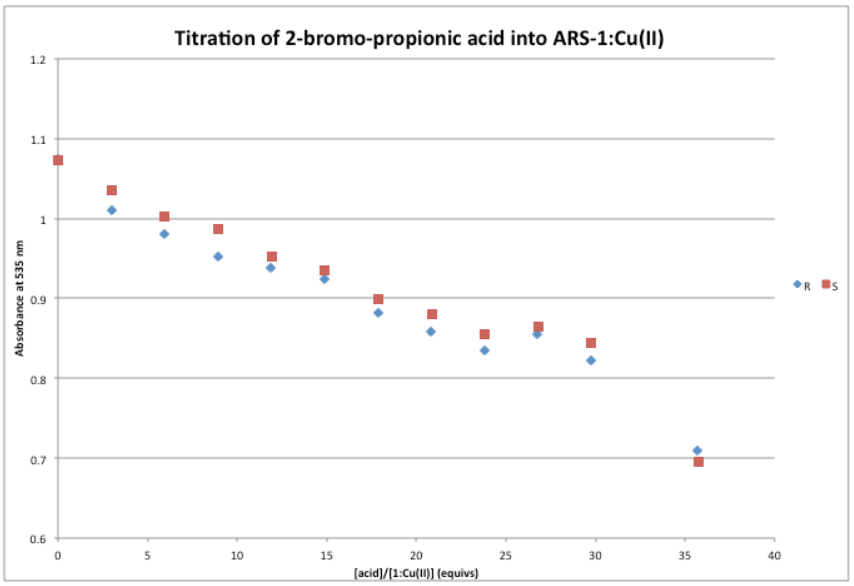


Figure 1.19 Titration of 25.5 mM (*R*)- and (*S*)-2-bromopropionate in 90% DMSO/H_2O with 10 mM HEPES into ARS-1.1:Cu(II) in the same buffer at 535 nm.

These data show that the receptor does not exhibit enantioselectivity for α-chiral carboxylates. However, the receptor did show a difference in absorbance intensity with ARS for the two types of α-chiral carboxylates tested, 2-phenylbutyrate and 2-bromopropionate.

1.4 Luminescent Europium Complexes for Chiral Anion Sensing

1.4.1 Lanthanide Complexes

Europium is a member of the lanthanide series of period six elements in the periodic table. Lanthanide ions can exhibit luminescence that is enhanced upon coordination with organic ligands.[51] The emission lifetimes for such complexes can range from micro- to milliseconds, have characteristically large Stokes shifts, and sharp emission peaks of ~10 nm.[52] Due to long emission wavelengths and long emission

lifetimes, the use of lanthanides in assays helps eliminate background noise from fluorescence of an analyte or media through the use of time-resolved spectroscopy.

All lanthanides form lanthanide(III) ions, which are usually octa- or nonacoordinate. They are hard ions and thus form more stable complexes with oxygen donors compared to nitrogen.[51] Water and hydroxide strongly coordinate with Ln^{3+}, so in aqueous solution, only negatively charged donors or multidentate neutral donors bind strongly.

Lanthanides electronic transitions occur in the 4f shell, which is highly shielded, resulting in narrow bands for f-f transitions.[53] Transitions between f orbitals are Laport-forbidden, resulting in low absorption coefficients and long radiative lifetimes. Population of lanthanide excited states can be achieved by sensitization using antenna ligands.[54] Excitation of ligands bound to the metal results in energy transfer to the lanthanide center. The sensitized responses to ligands occur in the $^5D_4 \rightarrow {}^7F_5$ and $^7F_4$ transitions for Tb^{3+} and $^5D_0 \rightarrow {}^7F_2$ and $^7F_4$ for Eu^{3+} due to the electric dipole character of the transitions.[51] The $^5D_0 \rightarrow {}^7F_2$ europium transition is hypersensitive to ligand environment because of the magnetic dipole change.

Lanthanide(III) complexes of europium and terbium have been prepared with peptides,[55] beta-diketones,[52] pyridyl bis-amidothioureas,[56] and bipyridine ligands.[57] H-bond donating groups have been used for sensing anions with lanthanides, either by interaction with the metal ion to directly perturb its emission or at the ligand antennae giving rise to modulated sensitization and thus changing the population of the metal excited state.[56]

1.4.2 Chiral Lanthanide Complexes

Chiral luminescent europium complexes have been prepared using chiral pyridylamide ligands derived from aminoethylnapthylene[58] and tryptophan[54] as well as with tris(2-pyridylmethyl)amines[59]. These complexes self-assemble in organic solvents such as methanol and acetonitrile. Gunnlaugsson *et al*. found that the chirality of the pyridylamide ligands was transferred to the lanthanide, resulting in the Δ or Λ isomers (Figure 1.20).[58] Yamada and coworkers tested a variety of achiral monoanions with chiral amine complexes and found that the anion sensitivity and selectivity varied depending on the lanthanide and chiral ligand used in the complex.[59] Binding of anions induced a change in the circular dichroism spectrum indicating that the anion coordination influenced the arrangement of the chromophores. When an achiral ligand was complexed with Tb^{3+}, binding to Cl^- was favored. When the (*R*)-enantiomer of the same ligand was used, binding to NO_3^- was preferred.

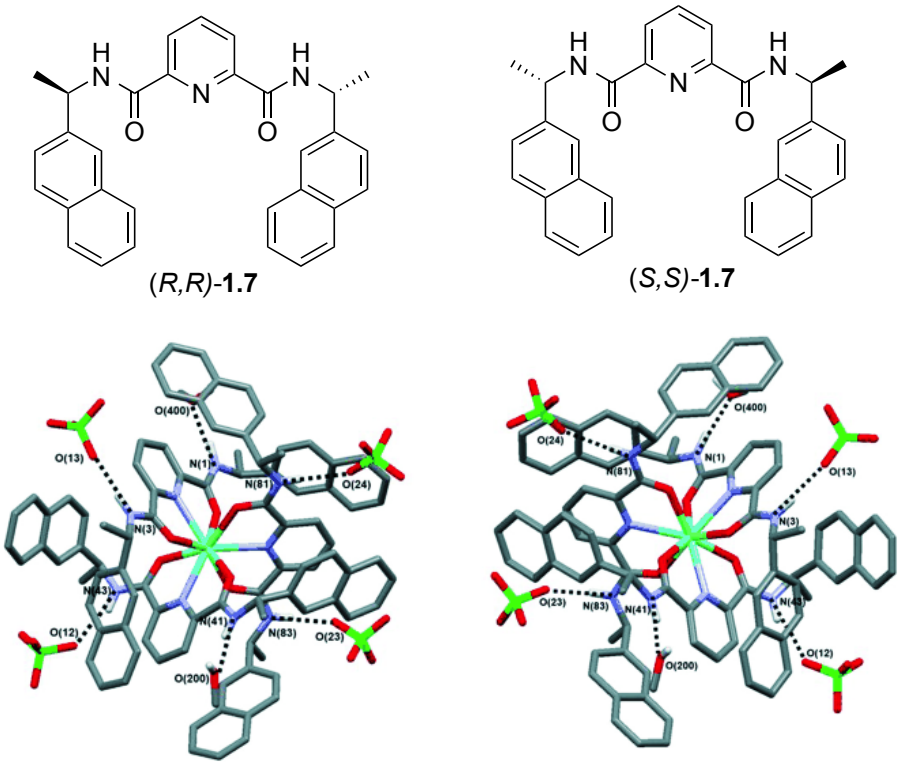


Figure 1.20 (*R,R*)- and (*S,S*)-complexes of aminoethylnaphthalene derivative **1.7** with Europium, adapted from Gunnlaugsson *et al*.[58] (*R,R*)-**1.7** forms the lambda Λ enantiomer, while (*S,S*)-**1.7** forms the Δ.

To date, chiral luminescent complexes have not been employed for sensing of chiral anions. Instead, racemic mixtures of lanthanide complexes have been exposed to chiral analytes to produce a "Pfeiffer-effect" perturbation in the equilibrium of luminescent chiral complexes observable using circularly polarized luminescence (CPL).[60] It is clear that anion binding, either by coordination to the metal or through interaction with the ligand antenna, can modulate lanthanide complex luminescence. Differential binding of chiral anions to chiral lanthanide complexes should therefore result in differential luminescence, dependent on the chirality of the anion.

1.4.3 Development of Chiral Luminescent Complexes with Bicycloguanidinium Receptors for Chiral Carboxylate Sensing

This research aimed to develop a luminescent assay for the detection of chiral carboxylate anions using bicycloguanidinium compounds and europium(III). The guanidinium receptors used in this research have a bicyclo[4.4.0]guanidinium core functionalized with variable peptide sequences (Figure 1.21).[49] The peptidic arms introduce additional H-bond donor functionality to the guanidinium and act as chelation units to Eu^{3+}.

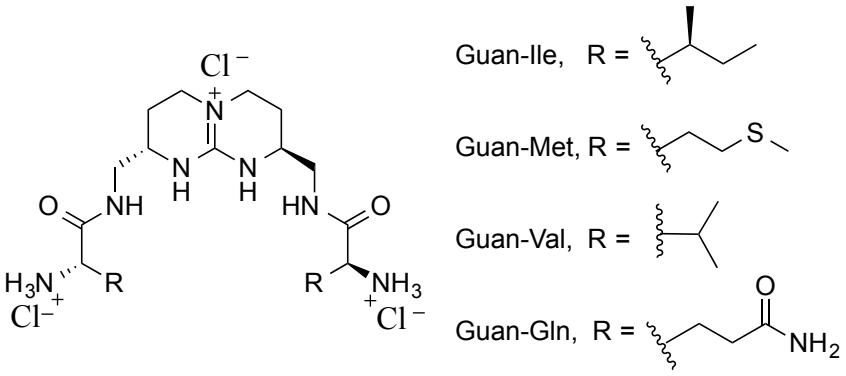


Figure 1.21. The bicycloguanidinium group containing amino acid arms.

Complexes of bicycloguanidinium derivatives with $Eu(CF_3SO_3)$ were prepared following the procedures of Gunnlaugsson.[58] Although crystal structures could not be isolated, the complexes exhibited luminescence that was enhanced over that of $Eu(CF_3SO_3)$ (Figure 1.22). To investigate the binding stoichiometry in solution, a titration of Guan-Gln with $Eu(CF_3SO_3)_3$ was carried out in MeOH. Increasing equivalents of europium(III) resulted in increased emission at 617 nm until saturation of the luminescence signal was reached (Figure 1.23). Although the stoichiometry of the complex in MeOH does not appear to be straightforward, eight equivalents of Eu^{3+} to Guan-Gln were used to ensure the complex signal was saturated.

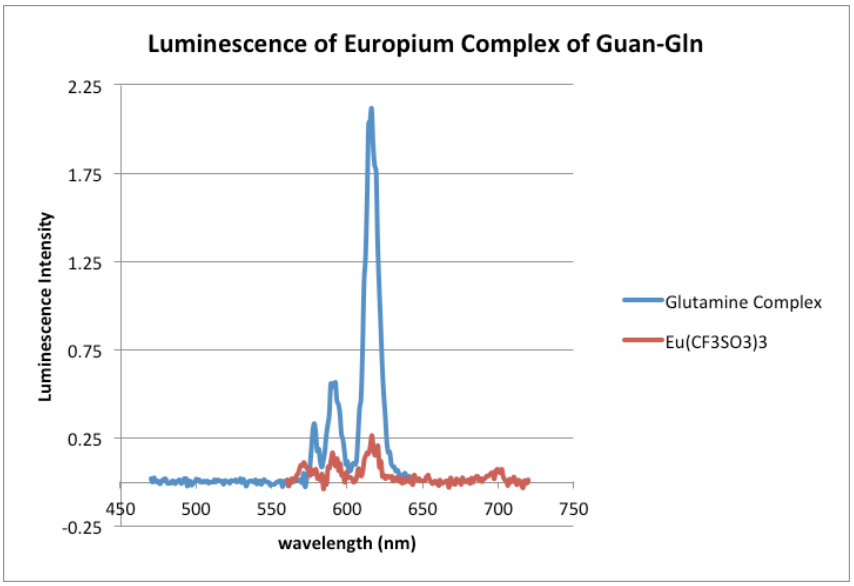


Figure 1.22 Luminescence for the complex obtained after reaction of $Eu(CF_3SO_3)_3$ with Guan-Gln compared to the europium salt alone in MeOH.

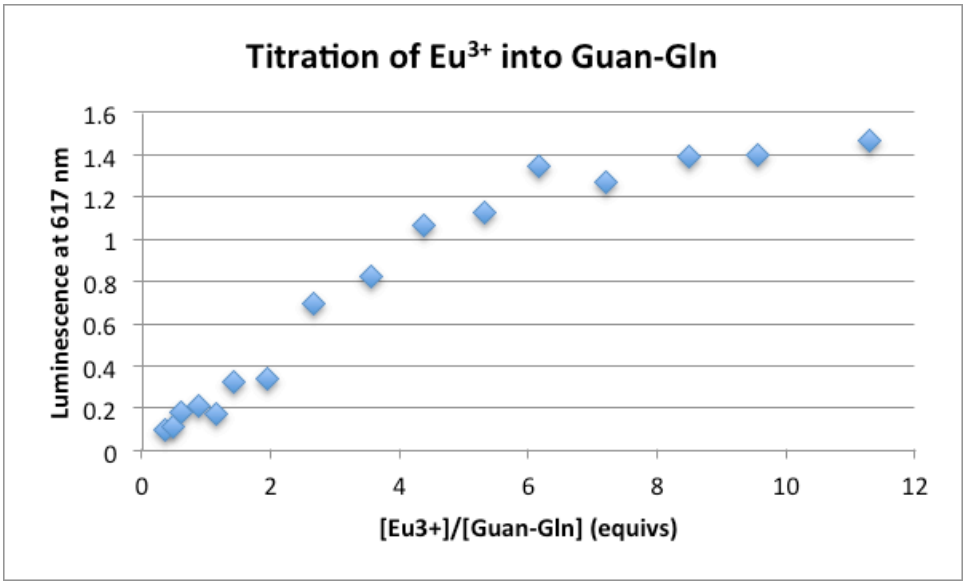


Figure 1.23 Luminescence change at 617 nm upon the addition of increasing equivalents of $Eu(CF_3SO_3)_3$ to Guan-Gln in MeOH. [Guan-Gln] =100 μM

Addition of (*R*)- or (*S*)-2-phenylbutyric acid to the europium complex with Guan-Gln resulted in an initial increase, followed by a decrease of the emission at 617 nm (Figure 1.24). The maximum signal intensity occurs at 0.5 equivalents for the (*R*)-enantiomer but 1.0 equivalents for the (*S*)-enantiomer, after which the emission gradually decreases. The difference in signal intensity between the enantiomers remains constant over 1-6 equivalents of acid and then decreases from 6-12 equivalents. The decrease in emission intensity did not saturate, even after the addition of 13 equivalents of acid.

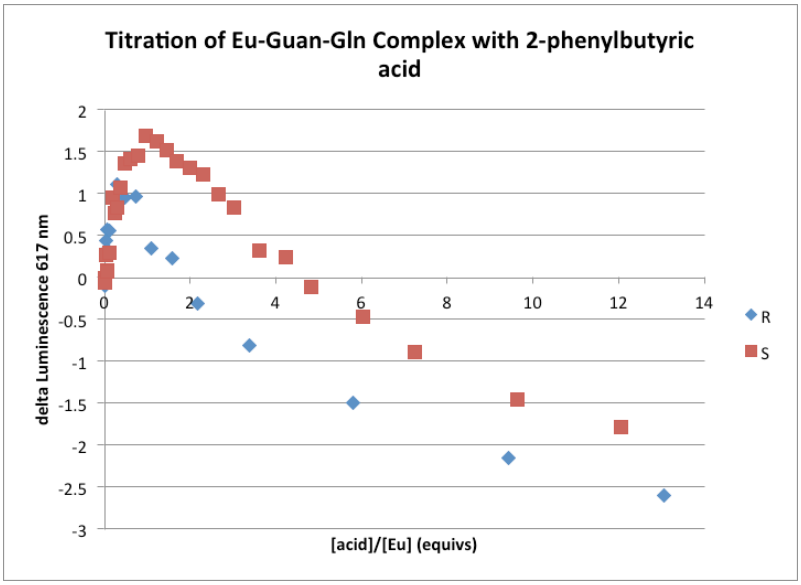


Figure 1.24 Luminescence change at 617 nm upon the addition of increasing equivalents of (*R*)- or (*S*)-2-phenylbutyric acid to Guan-Gln complexed with $Eu(CF_3SO_3)_3$ in MeOH. [Guan-Gln] = 1800 – 500 μM, [$Eu(CF_3SO_3)_3$] = 220 – 60 μM, [2-phenylbutyric acid] = 0 – 750 μM

A major concern was that the difference between enantiomers was due to solution pH differences as the titrations were performed in pure MeOH. For this reason, titrations were repeated in water buffered with 2-(*N*-morpholino)ethanesulfonic acid (MES) at pH 7.4. These titrations were undertaken with $EuCl_3$ in order to use a complimentary counter

ion to that in Guan-Gln. The luminescence at 617 nm increased upon the addition of Eu^{3+} to Guan-Gln, with signal saturation occurring at 0.7 – 0.8 equivalents (Figure 1.25). Based on this titration, the stoichiometry was determined to be 2:1 Guan-Gln:$EuCl_3$. Titration of (R)- or (S)-2-phenylbutyric acid into a solution of 2:1 Guan-Gln:$EuCl_3$ resulted in a signal increase at 617 nm. However, saturation was not reached even after addition of 44 equivalents of acid. There was no significant difference between the luminescence behavior at 617 nm with the addition of (*R*)- or (*S*)-enantiomers (Figure 1.26)

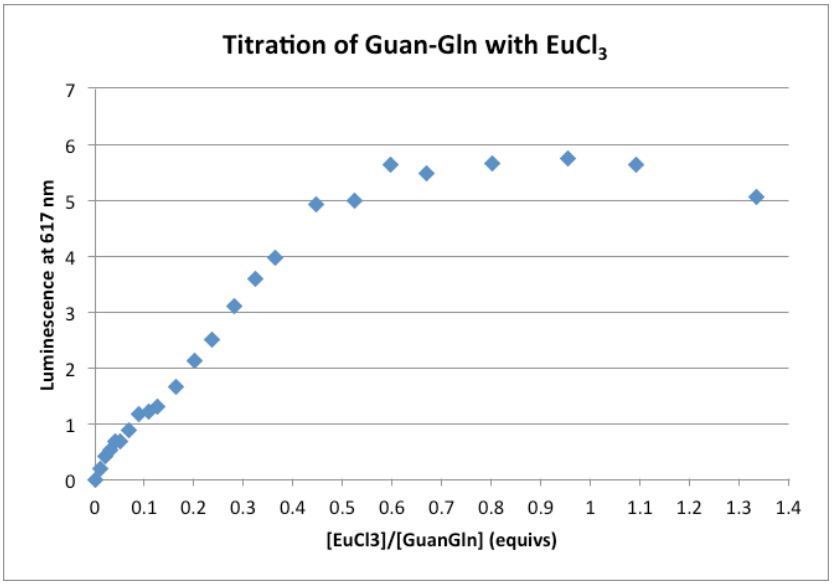


Figure 1.25 Luminescence change at 617 nm upon the addition of increasing equivalents of $EuCl_3$ to Guan-Gln in 10 mM MES pH = 7.4 [Guan-Gln] = 500 μM, [$EuCl_3$] = 0 – 670 μM

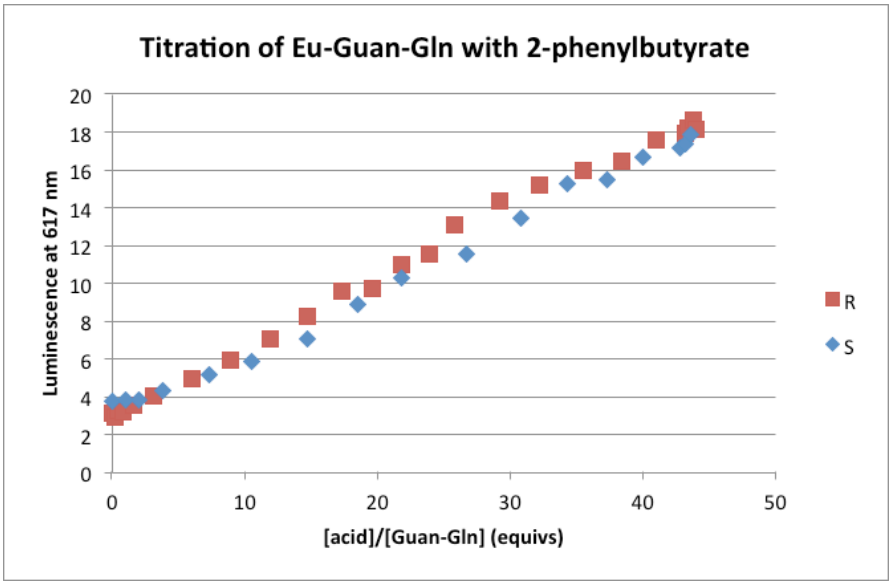


Figure 1.26 Luminescence change at 617 nm upon the addition of increasing equivalents of (R)- or (S)-2-phenylbutyrate to $EuCl_3$ complexed with Guan-Gln in 10 mM MES pH = 7.4 [Guan-Gln] = 500 μM, [$EuCl_3$] = 2 mM, [acid] = 0 – 22 mM.

1.5 Conclusions

Incorporation of the guanidinium group into a bicyclic structure contained within a Cu(II)-complexed macrocycle was hypothesized to give enantioselection for α-chiral carboxylates. Although the receptor was preorganized to provide a sufficiently chiral environment for enantioselection, there was no significant difference in the receptor's UV/Vis response to the tested α-chiral carboxylates when the receptor was bound to Alizarin Red S. Since IDAs have the advantage of reversible interactions between host and indicator, indicator screening was undertaken to optimize the system; however, other suitable indicators were not found. It is possible that the free energy difference between diastereomeric complexes of each enantiomer with the receptor was not significant, or that the indicator was not displaced upon addition of analyte. The former could be due to

enthalpic compensation for the expected entropic differences for enantiomer binding, or perhaps the receptor was not sufficiently rigid.

Many chiral carboxylate chemosensors have been reported that use colorimetric or fluorogenic signaling. The use of sensitized lanthanide luminescence for carboxylate sensing would eliminate potential interferences in samples that may be complex mixtures. Complexes of Eu^{3+} with bicycloguanidinium compounds derived with variable peptide arms were observed using time-resolved fluorescence titration and the emission increased upon addition of carboxylates. Although the tested receptors did not display differential luminescence sensitization with the tested chiral carboxylates in buffer, the complexes did show differences for the enantiomers in MeOH; however, it is unclear if the differences observed in methanol arise from pH effects. Additionally, because water coordinates strongly to lanthanide(III) ions, it is possible that the lack of enantioselectivity in water is due to the weak coordination of bicycloguanidinium cations to Eu^{3+}, which may minimize the free energy differences between expected diastereomeric complexes upon carboxylate coordination.

1.6 EXPERIMENTAL

1.6.1 Chiral Carboxylate Sensing using a Bipyridyl-Functionalized Bicycloguanidinium Compound

Absorbance spectra were measured using a Beckmann Coulter DU 800 UV/Vis Spectrophotometer scanning at 1 nm intervals. Absorbance values for well plate experiments were measured using a BioTek Synergy 2 Well Plate Reader scanning at 5 nm intervals.

1.6.1.1 Titration of Receptor with $CuCl_2$

Stock solutions of **1.1** (1 mM) and $CuCl_2$ (100 mM) were made in 90/10 DMSO/H_2O containing 10 mM HEPES buffer. The copper solution was added in aliquots to a solution of **1.1**, and the UV/Vis absorbance was monitored after each addition. The pH after the titration was monitored and stayed between 7.5-8.0.

1.6.1.2 Titration of Pyrocatechol Violet with $CuCl_2$

From a 100 mM stock solution, 1.5 mM $CuCl_2$ in 90/10 DMSO/H_2O with 10 mM HEPES was prepared. Aliquots of this solution were added to 500 uL of 20 μM pyrocatechol violet in the same buffer, and the absorbance spectrum monitored. After measuring the spectrum from 200-1000 nm, 1 equivalent of $CuCl_2$ was added to fresh pyrocatechol violet solutions, and the absorbance values at 675 nm and 430 nm were monitored over a period of 15 and 30 minutes respectively.

1.6.1.3 Indicator Uptake Experiments

Carboxylate Discrimination

A stock solution of sodium acetate, 15 mM in 100 mM HEPES, was prepared and then diluted to 1.5 mM in 90/10 DMSO/H_2O (10 mM HEPES). Equal amounts of this solution and a 1.5 mM solution of **1.1:Cu(II)** were combined and diluted to make a 75 μM solution of each in 90/10 DMSO/H_2O with 10 mM HEPES. An excess of pyrocatechol violet was added to the **1.1:Cu(II)** and acetate solution, and the UV/Vis absorbance was monitored over time. The above procedure was repeated for monobasic sodium phosphate.

Enantioselective Discrimination of Carboxylates

A 15 mM stock solution of either (*R*)- or (*S*)-phenylbutyric acid in 100 mM HEPES was diluted to 1.5 mM in 90/10 DMSO/H_2O with 10 mM HEPES. One

equivalent of these solutions was added into a solution of **1.1:Cu(II)**, and the solution was pH-adjusted to 8.10. One equivalent of PV was then added to this mixture, and the UV/Vis absorbance changes for the (*R*)- or (*S*)-complexes were observed five seconds later.

1.6.1.4 Titration of Receptor/$CuCl_2$ with Alizarin Red S

In a 384-well plate, variable amounts of **1.1:Cu(II)** were added to 50 uL of 500 μM Alizarin Red S. Each well was diluted with 90/10 DMSO/H_2O containing 10mM HEPES, until the final concentration of Alizarin Red S was 250 μM and the concentration of **1.1:Cu(II)** was 0-750 μM. After sitting for 1.5 hours, the absorbance spectrum for each well was measured. This titration was repeated for $CuCl_2$ without **1.1**.

1.6.1.5 Indicator Displacement Experiment

Stock solutions of (*R*)- or (*S*)-phenylbutyric acid (51 mM) and (*R*)- or (*S*)-bromopropanoic acid (71 mM) were prepared in DMSO. These solutions were diluted by half with 100 mM HEPES and DMSO, such that the final solutions contained 10 mM HEPES and 90/10 DMSO/H_2O along with the (*R*)- or (*S*)-acid. Equal concentrations of Alizarin Red S and **1.1:Cu(II)** were added to a 384-well plate, and then the (*R*)- or (*S*)-acid was titrated in. Each well was diluted to the same total volume using 90/10 DMSO/H_2O 10 mM HEPES buffer, such that the concentration of ARS and **1.1:Cu(II)** remained constant.

1.6.2 Luminescent Europium Complexes for Chiral Anion Sensing

Luminescence spectra were measured using a PTI QuantaMaster™ Fluorimeter equipped with a pulsed xenon flash lamp. Scans were taken at 1 nm intervals, 16 nm slit widths, a 150 usec delay and integration to 500 usec, 20 shots per scan, 300 Hz, and with

two averages. The minimum volume of titrand, 800 uL, was added to a quartz fluorescence cuvette for each titration.

MES buffer was prepared by weighing 488.27 mg MES hydrate into a 250 mL volumetric flask. The volumetric flask was filled ~90% with ddH_2O and the pH was adjusted to 7.4 using saturated NaOH, after which the flask was filled to volume with ddH_2O to give a final concentration of 10 mM MES.

1.6.2.1 Synthesis of bicycloguanidinium complexes with $Eu(CF_3SO_3)_3$

Following the procedure reported by Gunnlaugsson et al (REF), complexes of Guan-Gln and Guan-Met with $Eu(CF_3SO_3)_3$ were prepared using a microwave reactor. 6.51 mg $Eu(CF_3SO_3)_3$ was weighed into a microwave tube, and to this was added 44.32 mg of Guan-Gln (7.3 equivalents). 12.66 mg $Eu(CF_3SO_3)_3$ was weighed, and to this 60.43 mg Guan-Met was added. The microwave tubes were each filled with 7.5 mL HPLC grade MeOH and equipped with a stir bar and a cap. The tube was heated at 70°C for 10 minutes. Ether was added to the solutions cooled on ice, which resulted in formation of a white precipitate. The ether was decanted and the solutions air-dried to isolate the solids.

1.6.2.2 Titration of Guan-Gln with $Eu(CF_3SO_3)_3$

Stock Guan-Gln was prepared by weighing 2.50 mg into a 5 mL volumetric and filling to volume with MeOH to give a solution of 888 μM. 25 uL of this stock was added to a 1 mL volumetric flask and filled to volume with MeOH to give a 22.2 μM solution of Guan-Gln to be used as titrand.

Stock $Eu(CF_3SO_3)_3$ was prepared by weighing 2.10 mg into a 5 mL volumetric flask and filling to volume with MeOH to give a 701 μM solution. 715 uL of this stock was added to 25 uL of stock Guan-Gln in a 1 mL volumetric flask and then diluted to volume to give a titrant solution of 500 μM $Eu(CF_3SO_3)_3$ with 22.2 μM Guan-Gln.

1.6.2.3 Titration of Guan-Gln/Eu(CF_3SO_3)$_3$ complex with (R)- and (S)-2-phenylbutyric acid

Stock solutions of acid were prepared by weighing 2.99 mg of (*R*)-2-phenylbutyric acid and diluting with 1 mL MeOH to give a solution of 18.21 mM. 3.00 mg of (*S*)-2-phenylbutyric acid was diluted with 1 mL MeOH, giving a solution of 18.27 mM. These solutions were diluted as follows: 549 uL of 18.21 mM (*R*)-2-phenylbutyric acid was added to a 10 mL volumetric flask and diluted to volume with MeOH; 547 uL of 18.27 mM (*S*)-2-phenylbutyric acid was added to a 10 mL volumetric flask and diluted to volume with MeOH.

Titrand solutions were prepared with 250 μM Eu(CF_3SO_3)$_3$ with 22.2 μM Guan-Gln. Over the course of the titration, the titrand was diluted ~4x.

1.6.2.4 Titration of Guan-Gln with $EuCl_3$

To prepare the titrand, 2.41 mg of Guan-Gln was weighed into a 1 mL volumetric flask and filled to volume with 10 mM MES buffer, resulting in a solution of 4.28 mM Guan-Gln. 116.8 mL of the 4.28 mM Guan-Gln solution was transferred into a 1 mL volumetric flask and diluted to volume with 10 mM MES to give a 500 μM Guan-Gln solution.

To prepare the titrant, a solution of 8.62 mM $EuCl_3$ was prepared by weighing 53.5 mg $EuCl_3$ into a 25 mL volumetric flask. The flask was filled to volume with 10 mM MES buffer. To a 2 mL volumetric with 233.6 uL 4.28 μM Guan-Gln, 464 uL 8.62 mM $EuCl_3$ was added. The flask was filled to volume using 10 mM MES to give a solution with 500 μM Guan-Gln and 2 mM $EuCl_3$.

1.6.2.5 Titration of Guan-Gln/$EuCl_3$ complex with (R)- and (S)-2-phenylbutyric acid

(*R*)-2-phenylbutyric acid (100 mM) was prepared by weighing 165.26 mg into a 10 mL volumetric flask. The volumetric flask was filled ~90% with 10 mM MES. The pH was adjusted to 7.31 using saturated NaOH and 5 M HCl. Addition of NaOH beyond pH = 13 caused MES to precipitate, while at pH < 5 the acid was not completely soluble. After pH adjustment, the flask was filled to volume with 10 mM MES. 100 mM (*S*)-2-phenylbutyric acid was prepared in a similar manner to the R- enantiomer by weighing 164.82 mg into a 10 mL volumetric flask. The pH was adjusted to 7.75.

Guan-Gln (2.61 mM) was prepared by weighing 1.47 mg into a 1 mL volumetric flask and diluting to volume with 10 mM MES. This solution was used to prepare a 2:1 Guan-Gln:$EuCl_3$ 1 mM: 500 μM solution by adding 766 uL of 2.61 mM Guan-Gln and 116 uL of 8.62 mM $EuCl_3$ to a 2 mL volumetric flask and diluting to volume with 10 mM MES. This solution was diluted by a factor of two to give the titrand solution of 500 μM:250 μM Guan-Gln:$EuCl_3$.

Titrant solutions were prepared by adding 398 uL of 100.38 mM (*S*)-2-phenylbutyrate or 397 uL of 100.65 mM (*R*)-2-phenylbutyrate to a 1 mL volumetric flask with 500 uL of 1 mM:500 μM Guan-Gln:$EuCl_3$ and filling to volume with 10 mM MES. This gave a solution of 500 μM:250 μM Guan-Gln:$EuCl_3$ with 40 mM either (*R*)- or (*S*)-2-phenylbutyrate.

1.7 REFERENCES

1. Anslyn, E. V.; Dougherty, D. A., *Modern Physical Organic Chemistry*. 2nd ed.; University Science: Sausalito, CA, 2006.
2. Lichtenthaler, F. W., *Angewandte Chemie International Edition in English* **1995,** *33*, 2364-2374.

3. RoyChowdhury, A.; Ghosh, P.; Saha, S. K.; Mitra, P.; Banerjee, P., *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2014,** *124*, 492-499.

4. Esipenko, N. A.; Koutnik, P.; Minami, T.; Mosca, L.; Lynch, V. M.; Zyryanov, G. V.; Anzenbacher, P., *Chemical Science* **2013,** *4*, 3617-3623.

5. Hoffmann, E. K., *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes* **1986,** *864*, 1-31.

6. Bruzzoniti, M. C.; Sarzanini, C.; Mentasti, E., *Journal of Chromatography A* **2000,** *902*, 289-309.

7. Beer, P. D.; Gale, P. A., *Angewandte Chemie International Edition* **2001,** *40*, 486-516.

8. Palacios, M. A.; Nishiyabu, R.; Marquez, M.; Anzenbacher, P., *Journal of the American Chemical Society* **2007,** *129*, 7538-7544.

9. Hummer, G.; Pratt, L. R.; García, A. E., *The Journal of Physical Chemistry* **1996,** *100*, 1206-1215.

10. dos Santos, C. M. G.; Gunnlaugsson, T., *Supramol Chem* **2009,** *21*, 173-180.

11. Kobiro, K.; Inoue, Y., *Journal of the American Chemical Society* **2003,** *125*, 421-427.

12. Gale, P. A.; Sessler, J. L.; Kral, V., *Chemical Communications* **1998**, 1-8.

13. Gale, P. A. a. H., Cally J. E., Anion Receptors Containing Heterocyclic Rings. In *Supramolecular Chemistry: from Molecules to Nanomaterials*, First ed.; Steed, P. A. G. a. J. W., Ed. John Wiley & Sons, Ltd: West Sussex, 2012; Vol. 3, pp 1125-1151.

14. Md. Alamgir Hossain, R. A. B., Victor W. Day, Kristin Bowman-James, Amide and Urea-Based Receptors. In *Supramolecular Chemistry: from Molecules to Nanomaterials*, First ed.; Steed, P. A. G. a. J. W., Ed. John Wiley & Sons, Ltd: West Sussex, 2012; Vol. 3, pp 1153-1178.

15. Caltagirone, C.; Gale, P. A., *Chemical Society Reviews* **2009,** *38*, 520-563.

16. Stibor, I.; Zlatu≈°kov√°, P., Chiral Recognition of Anions. In *Anion Sensing*, Stibor, I., Ed. Springer Berlin Heidelberg: 2005; Vol. 255, pp 31-63.

17. Kano, K.; Kamo, H.; Negi, S.; Kitae, T.; Takaoka, R.; Yamaguchi, M.; Okubo, H.; Hirama, M., *Journal of the Chemical Society, Perkin Transactions 2* **1999**, 15-22.

18. Zhu, L.; Anslyn, E. V., *Journal of the American Chemical Society* **2004,** *126*, 3676-3677.

19. Wright, A. T.; Anslyn, E. V., *Chemical Society Reviews* **2006,** *35*, 14-28.

20. (a) Wright, A. T.; Griffin, M. J.; Zhong, Z. L.; McCleskey, S. C.; Anslyn, E. V.; McDevitt, J. T., *Angew Chem Int Edit* **2005,** *44*, 6375-6378; (b) Zhou, H. C.; Baldini, L.; Hong, J.; Wilson, A. J.; Hamilton, A. D., *J Am Chem Soc* **2006,** *128*, 2421-2425.

21. Albert, K. J.; Lewis, N. S.; Schauer, C. L.; Sotzing, G. A.; Stitzel, S. E.; Vaid, T. P.; Walt, D. R., *Chemical Reviews* **2000,** *100*, 2595-2626.

22. (a) Wiskur, S. L.; Anslyn, E. V., *Journal of the American Chemical Society* **2001,** *123*, 10109-10110; (b) Wiskur, S. L.; Ait-Haddou, H.; Lavigne, J. J.; Anslyn, E. V., *Accounts of Chemical Research* **2001,** *34*, 963-972.

23. Zhu, L.; Zhong, Z. L.; Anslyn, E. V., *Journal of the American Chemical Society* **2005,** *127*, 4260-4269.

24. Metzger, A.; Anslyn, E. V., *Angewandte Chemie International Edition* **1998,** *37*, 649-652.

25. Theodore J. Leitereg, D. G. G., Jean Harris, Thomas R. Mon, Roy Teranishi, *Journal of Agricultural and Food Chemistry* **1971,** *19*, 785-787.

26. Polak, E. H.; Fombon, A. M.; Tilquin, C.; Punter, P. H., *Behavioural Brain Research* **1989,** *31*, 199-206.

27. Nelson, G.; Chandrashekar, J.; Hoon, M. A.; Feng, L.; Zhao, G.; Ryba, N. J. P.; Zuker, C. S., *Nature* **2002,** *416*, 199-202.

28. Folmer-Andersen, J. F.; Kitamura, M.; Anslyn, E. V., *Journal of the American Chemical Society* **2006,** *128*, 5652-5653.

29. Leung, D.; Folmer-Andersen, J. F.; Lynch, V. M.; Anslyn, E. V., *Journal of the American Chemical Society* **2008,** *130*, 12318-12327.

30. Shabbir, S. H.; Joyce, L. A.; da Cruz, G. M.; Lynch, V. M.; Sorey, S.; Anslyn, E. V., *Journal of the American Chemical Society* **2009,** *131*, 13125-13131.

31. Willener, Y.; Joly, K. A.; Moody, C. J.; Tucker, J. H. R., *J Org Chem* **2008,** *73*, 1225-1233.

32. (a) Alfonso, I.; Dietrich, B.; Rebolledo, F.; Gotor, V.; Lehn, J. M., *Helv Chim Acta* **2001,** *84*, 280-295; (b) Stibor, I.; Holakovsky, R.; Mustafina, A. R.; Lhotak, P., *Collect Czech Chem C* **2004,** *69*, 365-383; (c) Xu, K. X.; Wang, Y. X.; Jiao, S. Y.; Zhao, J.; Wang, C. J., *Can J Chem* **2010,** *88*, 367-374.

33. (a) Yang, D.; Li, X.; Fan, Y. F.; Zhang, D. W., *Journal of the American Chemical Society* **2005,** *127*, 7996-7997; (b) Qing, G. Y.; He, Y. B.; Wang, F.; Qin, H. J.; Hu, C. G.; Yang, X., *Eur J Org Chem* **2007**, 1768-1778.

34. (a) Chen, Z. H.; He, Y. B.; Hu, C. G.; Huang, X. H., *Tetrahedron-Asymmetr* **2008,** *19*, 2051-2057; (b) Folmer-Andersen, J. F.; Lynch, V. M.; Anslyn, E. V., *Journal of the American Chemical Society* **2005,** *127*, 7986-7987.

35. Blondeau, P.; Segura, M.; Perez-Fernandez, R.; de Mendoza, J., *Chem Soc Rev* **2007,** *36*, 198-210.

36. Miyaji, H.; Hong, S.-J.; Jeong, S.-D.; Yoon, D.-W.; Na, H.-K.; Hong, J.; Ham, S.; Sessler, J. L.; Lee, C.-H., *Angewandte Chemie International Edition* **2007,** *46*, 2508-2511.

37. Joyce, L. A.; Maynor, M. S.; Dragna, J. M.; da Cruz, G. M.; Lynch, V. M.; Canary, J. W.; Anslyn, E. V., *Journal of the American Chemical Society* **2011,** *133*, 13746-13752.

38. Schmuck, C., *Chem-Eur J* **2000,** *6*, 709-718.

39. Müller, G.; Riede, J.; Schmidtchen, F. P., *Angewandte Chemie International Edition in English* **1988,** *27*, 1516-1518.

40. Berger, M.; Schmidtchen, F. P., *Journal of the American Chemical Society* **1996,** *118*, 8947-8948.

41. Berger, M.; Schmidtchen, F. P., *Angew Chem Int Edit* **1998,** *37*, 2694-2696.

42. Jadhav, V. D.; Schmidtchen, F. P., *Org Lett* **2005,** *7*, 3311-4.

43. Echavarren, A.; Galan, A.; Lehn, J. M.; Demendoza, J., *Journal of the American Chemical Society* **1989,** *111*, 4994-4995.

44. Galan, A.; Andreu, D.; Echavarren, A. M.; Prados, P.; Demendoza, J., *Journal of the American Chemical Society* **1992,** *114*, 1511-1512.

45. Gleich, A.; Schmidtchen, F. P.; Mikulcik, P.; Muller, G., *J Chem Soc Chem Comm* **1990**, 55-58.

46. Jadhav, V. D.; Schmidtchen, F. P., *Org Lett* **2006,** *8*, 2329-32.

47. Jadhav, V. D.; Herdtweck, E.; Schmidtchen, F. P., *Chemistry* **2008,** *14*, 6098-107.

48. Jadhav, V. D.; Schmidtchen, F. P., *J Org Chem* **2008,** *73*, 1077-1087.

49. Collins, B. E. A kinetic investigation of boronic acid/diol interactions and pattern-based recognition of [alpha]-chiral carboxylates. PhD. Dissertation, The University of Texas at Austin, 2010.

50. Garribba, E.; Micera, G.; Sanna, D.; Strinna-Erre, L., The Cu(II)-2,2'-bipyridine system revisited. *Inorganica Chimica Acta* **2000,** *299* (2), 253-261.

51. Richardson, F. S., *Chemical Reviews* **1982,** *82*, 541-552.

52. Yuan, J.; Wang, G., *J Fluoresc* **2005,** *15*, 559-568.

53. Vogler, A.; Kunkely, H., *Inorganica Chimica Acta* **2006,** *359*, 4130-4138.

54. Lincheneau, C.; Leonard, J. P.; McCabe, T.; Gunnlaugsson, T., *Chemical Communications* **2011,** *47*, 7119-7121.

55. Bonnet, C. S.; Devocelle, M.; Gunnlaugsson, T., *Organic & Biomolecular Chemistry* **2012,** *10*, 126-133.

56. Lincheneau, C.; Duke, R. M.; Gunnlaugsson, T., *Organic & Biomolecular Chemistry* **2012,** *10*, 6069-6073.

57. Butler, C.; Goetz, S.; Fitchett, C. M.; Kruger, P. E.; Gunnlaugsson, T., *Inorganic Chemistry* **2011,** *50*, 2723-2725.

58. Kotova, O.; Kitchen, J. A.; Lincheneau, C.; Peacock, R. D.; Gunnlaugsson, T., *Chemistry – A European Journal* **2013,** *19*, 16181-16186.

59. Yamada, T.; Shinoda, S.; Tsukube, H., *Chemical Communications* **2002**, 1218-1219.

60. Muller, G.; Riehl, J., *J Fluoresc* **2005,** *15*, 553-558.

Chapter 2: Introduction to Multivalent Biological Analyte Sensing

Biological targets, such as proteins and cells, are challenging analytes owing to their complex structures. Protein macromolecules exhibit secondary and tertiary structures as well as post-translational modifications, such a glycosylation, that influence their function and shape. Cell surfaces contain of a variety of proteins, lipids, and carbohydrates, which differ in composition according to cell type.[1] Cancer cells exhibit changes in the number and distribution of these macromolecules on their surface membranes compared to normal cells.[2] Cell surface macromolecules use multivalent interactions for signaling, conformational contact between surfaces, and strong binding.[3] Analytes such as proteins and cells are currently detected by bioimaging,[4] sequencing,[5] or highly selective receptors, such as antibody arrays.[6] Differential arrays for multivalent biological analytes have emerged more recently, and such receptor arrays are also useful for detection and differentiation of multivalent analytes such as proteins and cells.

2.1 Multivalency

Multivalency, or polyvalency, is the simultaneous interaction of multiple binding units on one entity with multiple binding units on a complementary entity, where binding units are typically molecules.[3] The number of shared, complementary interactions between a host and guest defines the valency of a complex,[7] and a multivalent complex is defined when the separation of the complex requires the dissociation of at least two interactions between partners with multiple binding units (Figure 2.1).[7] Compared to monovalent interactions, multivalent associations have unique thermodynamic and kinetic properties, resulting in advantages in molecular association compared to an equivalent number of monovalent interactions or a single strong monovalent interaction.

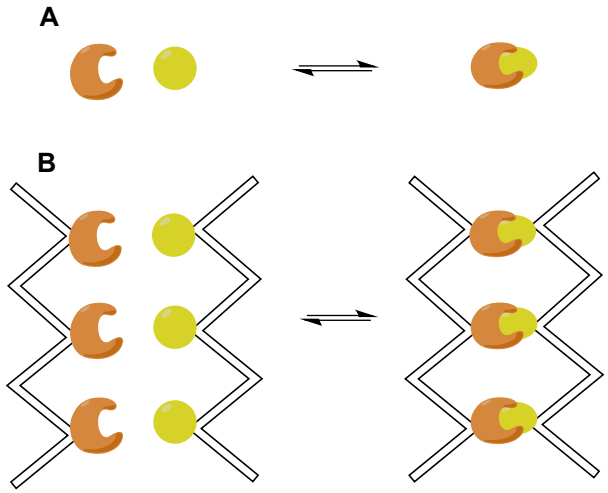


Figure 2.1 Illustration of the valency of complexes. A monovalent complex formed from a monovalent host and a monovalent guest (**A**) and a multivalent complex (trivalent) formed from a trivalent host and a trivalent guest (**B**).

Multivalent receptors can also form complexes with multiple multivalent ligands in an intermolecular fashion, which leads to the formation of aggregates (Figure 2.2). After the first association between complementary units on a multivalent host and a multivalent guest, the probability of subsequent inter- or intramolecular binding to a neighboring host site depends on the effective concentration of the now tethered guest compared to the concentration of a second guest free in solution.[7] Intermolecular binding is favored at high concentrations and with small, rigid entities with low directionality, while formation of a multivalent complex is favored at low concentrations. Receptors can be designed with high directionality and some flexibility in order to favor formation of multivalent complexes.

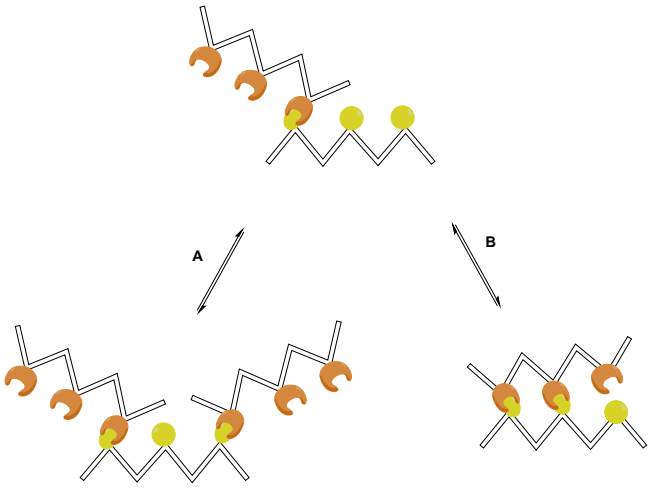


Figure 2.2 After the first association of multivalent hosts and guests, subsequent binding results in either intermolecular binding to form aggregates (**A**) or intramolecular binding to form a multivalent complex (**B**), depending on the effective concentration of a second ligand.

Statistically, multivalent complex formation is more favorable than monovalent complex formation because there are multiple, simultaneous interactions, and as such multivalent interactions can be much stronger than a single monovalent one. The energy of association of a multivalent complex is described by its Gibbs free energy (equation 2.1). Association of ligands and receptors to form multivalent complexes can occur with favorable or unfavorable enthalpic and entropic components. The entropy and enthalpy can be enhanced compared to the monovalent interaction if the average multivalent interaction has a more negative enthalpy and more positive entropy. Rigid linkers can lead to spacial mismatches between multivalent receptors and ligands, which results in enthalpically diminished binding. Flexible linkers may lead to entropically diminished associations due to loss of conformational entropy upon binding; however, the dependence of the free energy of multivalent interactions on linker length is weak, and strong associations with flexible linkers have been observed, likely due to their higher effective concentrations.[8,9]

$$N\Delta G_{avg}^{poly} = \Delta G_N^{poly} = \Delta H_N^{poly} - T\Delta S_N^{poly} \quad \text{(equation 2.1)}$$

The association for two interacting units can be represented as an affinity constant, K_{ass}^{mono}, determined from the Gibbs free energy. This constant differs from the association constant of a multivalent interaction, referred to as avidity, which is defined by the average of the association constants multiplied over *N* interactions between *N* ligands and *N* receptors on two entities (equation 2.2). When $K_N^{poly} > K_{ass}^{mono}$ the multivalent interaction exhibits enhanced affinity over the monovalent interaction. This can be associated with positive, negative, or non-cooperativity.

$$K_N^{poly} = (K_{avg}^{poly})^N \quad \text{(equation 2.2)}$$

Cooperativity (α) of multivalent ligand association can be evaluated by comparing the average free energy of the multivalent interactions to that of the monovalent interaction. If α = 1, there is no enhancement to subsequent association of a ligand in the multivalent complex. If α > 1, there is an enhancement. If α < 1, subsequent binding is thermodynamically unfavorable compared to the first association (equation 2.3). Often the valency, *N*, of a complex is not known, which makes it impossible to determine cooperativity.[3] Additionally, a multivalent interaction can be much stronger than a single monomeric interaction that contributes to the complex, even if the monomeric interactions interfere with one another (α < 1). The parameter β describes the enhanced affinity of a multivalent interaction compared to a monovalent one (equation 2.4). A multivalent association will be useful if β > 1 regardless of the cooperativity. For such a case, $K_N^{poly} > K^{mono}$ even if $K_N^{poly} \leq (K^{mono})^N$.

$$\Delta G_{avg}^{poly} = \alpha\Delta G^{mono} \quad \text{(equation 2.3)}$$

$$\beta = \frac{K_N^{poly}}{K^{mono}} \quad \text{(equation 2.4)}$$

Compared to monovalent interactions, the kinetics of multivalent interactions is advantageous for maintaining complex association. The kinetics of association, k_{on}, for

multivalent interactions is the same as that of the monovalent case. In contrast, the rate of dissociation, k_{off}, depends on the stepwise dissociation of all of the interactions in the multivalent complex, and the complex is fully dissociated when dissociation of the monovalent interaction between two entities has occurred.[7] If the local concentration is high, then partially dissociated complexes will further bind, leading to low k_{off}.

Heteromeric multivalency can result in greater strength and specificity than the equivalent monovalent interactions. This type of interaction involves a mixture of ligand-receptor pairs of different types. Additional ligand types may increase the number of interactions between two multivalent entities and therefore increase the avidity of the interaction.[3] Heteromeric multivalency is more broadly applicable than the homomeric analog because it gives the ability to target multivalent ligands that do not have multiple, closely associated identical binding sites. Monomeric proteins, for example, are multivalent ligands that often display a main binding site with secondary binding sites, such as adjacent hydrophobic pockets, which are best targeted with heteromeric multivalent receptors.[9]

2.2 Designed and Evolved Receptors for Multivalent Biological Analytes

Multivalent interactions are common in biological systems and come with a variety of functional advantages. The capability of multiple interactions allows for a range of signal strengths, which is more useful than a binary on/off response from a single interaction.[3] Multiple interacting ligands on two entities allows for conformal contact not achievable by a single, strong interaction. This contact can also lead to conformational changes that act as signaling mechanisms. Cell adhesion represents a multivalent interaction that is regulated by dynamic adhesive contacts from adherent

molecules diffusely distributed over the cell surface, such as cadhedrin, immunoglobulin, selectin, proteoglycan, and integrin.[10] Interaction of cell surfaces with multivalent molecules can lead to reorganization and redistribution of molecules on the cell surface.[3]

Biological receptors exhibit high specificity, high affinity, and reversibility for their ligands as a result of evolution, lending recognition and differentiation that allow for the complexity of life.[11] Hence, it makes good sense that biopolymers can be used to bind particular biological targets *in lieu* of small molecules, taking advantage of the properties afforded from multivalent interactions. Nucleic acids, proteins, peptides, and antibodies are some of the multivalent receptors that have been developed for biological analyte recognition, through the use of highly specific interactions.

2.2.1 Highly Specific Arrays

Nucleic acids have recently gained popularity as sensing platforms, particularly in biosensing. For example, DNA microarrays contain an assortment of specific DNA sequences that hybridize with cyanine dye-labeled complementary sequences in a sample, and can be used for genotyping, detection of single nucleotide polymorphisms (SNPs), and measuring gene expression.[12] DNA hybridization is a highly specific, multivalent interaction governed by Watson and Crick base pair complementarity.[13] The specificity of the interaction does, however, depend on effects like matching between probe and target, probe length, and probe GC content. At the microarray level, specificity can be affected by sample purity and probe quantity. Nevertheless, such gene analysis is useful for comparison and monitoring of normal and disease states.

Gene activity and regulation is a complex, dynamic process and not necessarily linearly related to protein expression.[14] Therefore, array technologies to profile, detect, and quantitate proteins have been developed that consist of proteins, antibodies, tissues,

or small molecules that are designed for highly specific interactions with analytes. Identification of protein sequence, structure, post-translational and structural modifications, and interaction partners is useful for characterization of biological states.

To generate protein-profiling arrays, biological or biologically relevant molecules are attached to support surfaces, referred to as chips, in high density and then tested in order to determine a sample's composition. Such arrays have been used for antibody specificity, cross-reactivity, and immune marker identification; protein-protein, protein-drug, and protein-ligand interactions; and sera analysis and epitope mapping. These arrays utilize specific receptors that are not necessarily cross-reactive, but are useful for their high information-density and fast readouts. The ultimate goal of specific array sensing is systemic analysis of biomolecules and their roles in disease.

2.2.2 Phage Display and SELEX

Instead of using proteins, antibodies, or small molecules designed for specificity, receptors for biomolecules can be discovered in an analogous manner to natural selection by screening large libraries of diverse chemical structures. These libraries can be generated through combinatorial chemical synthesis or by utilizing biological machinery. Biopolymers are particularly suited for generating diverse libraries to discover receptors for biological analytes as they can be easily replicated and therefore artificially evolved.

The use of building blocks to generate diverse structures parallels what arose through nature and billions of years of evolution.[15] Using nucleotide, amino acid, and sugar building blocks, complex structures like oligonucleotides, peptides and proteins, and carbohydrates can be generated.[16] Diverse biopolymer libraries can be created using chemical synthesis,[17] by harnessing biology,[18] or using a biosynthetic combination of both.[19]

Both the generation of molecular diversity and the selection of molecules that have high affinity for a biological target can be achieved using biological machinery. Exposure of biopolymers to a target analyte or analyte mixture in combination with a selection pressure, such as an affinity assay, winnows libraries of random biopolymer sequences to those sequences that are best suited to the target. Such "artificial evolution" mimics the diversity created through evolution but occurs at an accelerated rate due to the increased rate of diversity generation compared to that of nature.

Two major methods of artificial evolution are phage display and SELEX (*S*ystematic *E*volution of *L*igands by *EX*ponential enrichment). Phage display is an *in vivo* way of generating diverse libraries where a viral DNA programs the expression of a peptide of a particular sequence on a bacteriophage coat protein.[18] Combinatorial libraries can be made from a heterogeneous mixture of phages, each with a unique DNA plasmid and therefore expressing a particular peptide. Exposure of the phage library to a target or mixture of targets and subsequent affinity purification leads to capture of the bound phages. The bound members of the library are copied, passing along any mutations in the encoding DNA, and amplified when the bacteriophage infects a bacterial host and replicates. The replicated library is used for further affinity purification and the library's peptide sequences are identified using the DNA sequences of the plasmids in the selected phages. Such libraries have been used to generate human antibody fragments *in vitro*,[20] and to generate peptide ligands for drug discovery.[21] Weiss et al. developed a biosensor for prostate specific membrane antigen (PSMA) by screening a library of phage-displayed peptides in order to isolate an unnatural polypeptide that selectively binds PSMA.[22] The chosen peptide was displayed on a virus and used for the fabrication of nanowires as part of a PSMA detecting device.

Another popular example of artificial evolution is SELEX, an *in vitro* method where large libraries of oligonucleotides are synthesized chemically, amplified using polymerase enzymes, and exposed to a biological target in an iterative process (Figure 2.3).[23] A library of single-stranded DNA or RNA sequences is screened against a biological target for affinity and the binding sequences retained are replicated using PCR. The double strands resulting from PCR are separated to give single strands. For RNA libraries, the resulting dsDNA is reverse transcribed, and RNA single strands are isolated. The new library of binding sequences is used for further affinity purification, until multiple cycles with increasing selection pressure and subsequent amplification lead to purification of a small number of sequences specific to the biological target. These oligonucleotides can be sequenced to determine sequence homologies of the winnowed pool to identify binding structures. Single stranded nucleic acids that show high affinity and specificity for a target are known as aptamers. Using SELEX, aptamers have been generated for a variety of applications requiring high sensitivity and selectivity, such as affinity purification, biosensors, and therapeutics.[24]

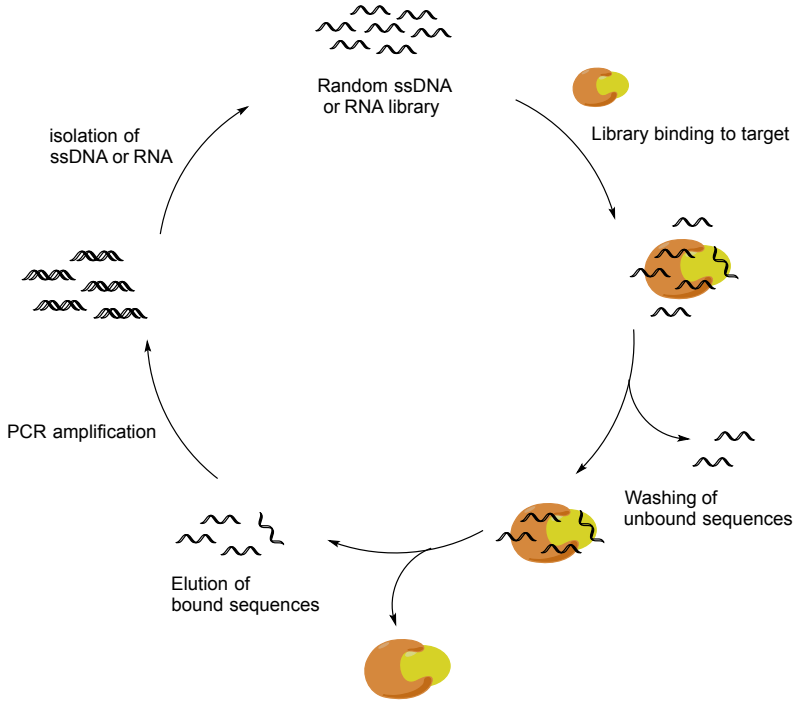


Figure 2.3 The process of SELEX.

Aptamers have been generated for both bacterial and eukaryotic cells.[25] Some aptamers are generated against specific epitopes on cell surfaces, while others have been selected against whole cells using a process known as cell-SELEX.[26] Cell-SELEX can generate aptamers specific to particular cell types by using negative selections against control cells to remove sequences that do not specifically bind the target cell. This allows for the generation of receptors that preferentially bind one type of cell over another without characterizing the cell surface features that differ between analytes.

2.3 Differential Sensing of Multivalent Biological Analytes

Molecular sensors, such as DNA microarrays and aptamers, have traditionally been designed for high affinity and specificity to a particular analyte.[27] In contrast, differential sensing is inspired by biological recognition of analytes by receptors in the

nose and tongue and utilizes an array of cross-reactive receptors that do not necessarily require specificity or high affinity.[28] Instead, analytes can be discriminated by their diagnostic patterns of interaction with an array of receptors, where each receptor responds to each analyte to a different degree.[29]

Differential arrays have been developed for multivalent biological analytes, such as proteins and cells, using multivalent receptors. These arrays are capable of analyte detection at nanomolar concentrations due to multivalent interactions.

2.3.1 Differential Sensors for Peptides and Proteins

Proteins are highly complex molecules; therefore, the design of synthetic receptors that target proteins with high affinity and specificity is challenging. As a result, differential sensing methods have been developed using receptors of good affinity that are capable of responding to many analytes. Using a resin-bound array based on a combinatorial library consisting of 19^3 members of 1,3,5-triethylbenzene substituted with binding arms composed of amino acids and boronic acids, Wright *et al*. were able to distinguish classes of proteins through the ion pairing, hydrogen bonding, reversible boronate ester formation, and hydrophobic interactions built into the receptor hosts (Figure 2.4).[30] The proteins tested varied in molecular weight (MW), glycosylation, and isoelectric point (pI). Using chemometric techniques to reduce the high dimensionality of the measured indicator uptake rates across the array, the tested proteins were effectively grouped into classes based on their pI and were well separated according to their glycosylation state.

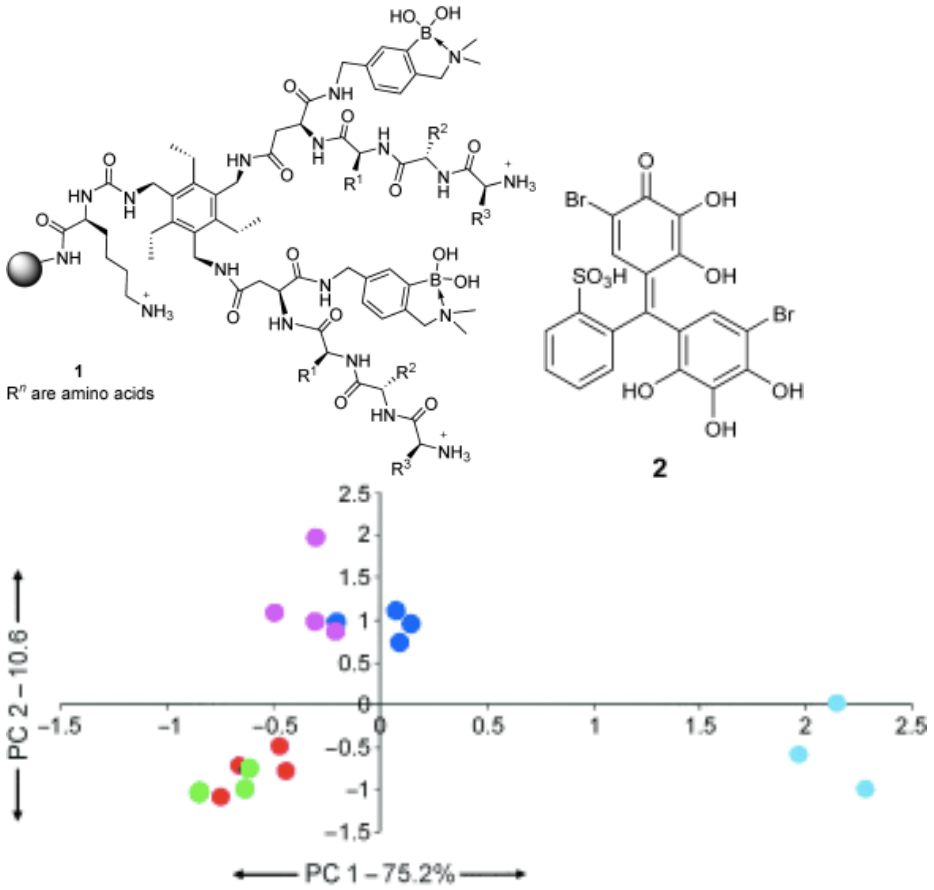


Figure 2.4 The receptors and indicator used by Wright *et al*. (top) and the resulting PCA score plot of proteins using optical data from the resin-bound array (bottom).

You *et al*. employed a "chemical nose" sensory composed of nanoparticles and fluorescent polymers.[31] The researchers chose nanoparticles as receptor scaffolds because their size and surface areas, especially compared to small molecules, are appropriate for conforming contact with protein surfaces and their surface structure can be manipulated. Cationic gold nanoparticles functionalized with six end groups, four of which were hydrophobic, one aromatic, and one alcoholic, quenched the fluorescence of bound polymer. When the polymer-bound particles were exposed to solutions of protein, the polymer was displaced from the nanoparticle and its fluorescence restored. The

fluorescence signal differed for the seven proteins tested, and the values were input into chemometric analysis routines for classification (Figure 2.5).

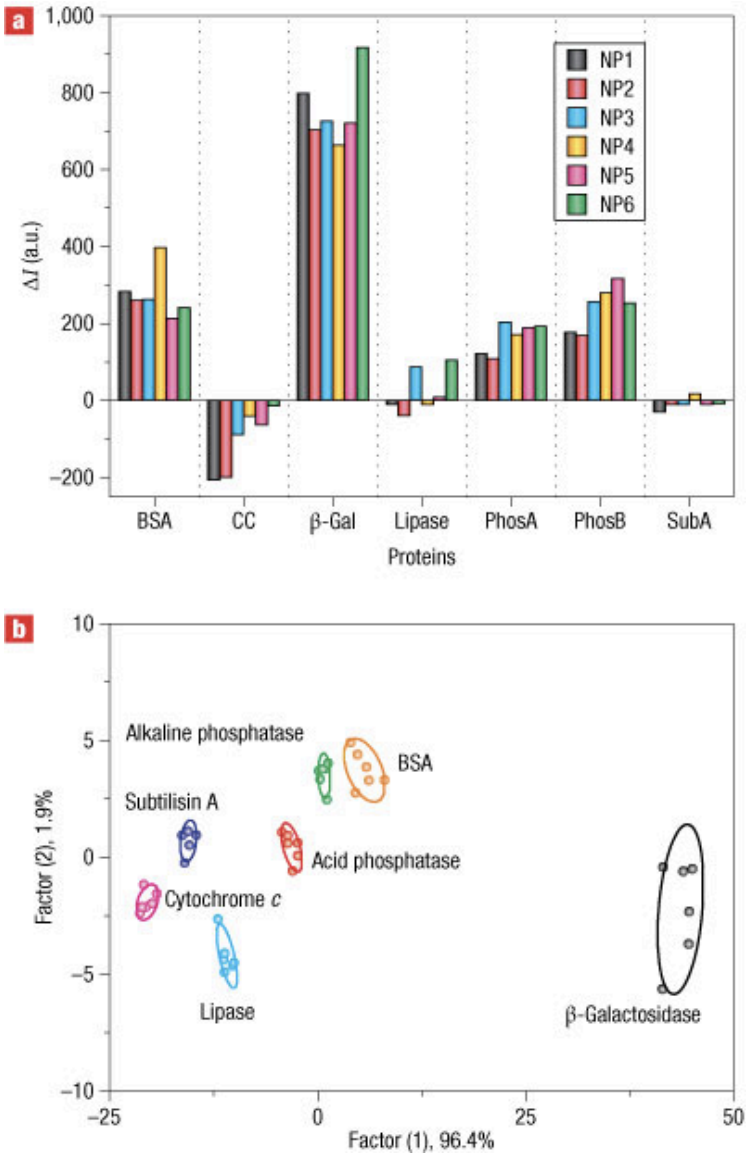


Figure 2.5 Fluorescence patterns generated from displacement by proteins of a quenched polymer bound to functionalized nanoparticles (top). LDA classification of proteins based on the fluorescence response (bottom).

The linear discriminant analysis (LDA) score plot shows separation and classification of the different proteins and proteins tested, and most of the samples fall within the 95% confidence ellipses. However, 96.4% of the classification is captured on the first axis, while a paltry 1.9% is captured on the second axis. This indicates that the nanoparticles used in the array all behaved in a similar manner, and the end group functionalization had minimal impact on the interaction of the nanoparticle with the tested proteins. The plot of fluorescence response of each nanoparticle to each protein indeed shows that, for the majority of the responses, each nanoparticle gives a fluorescence response of similar direction and magnitude. Such responses suggest that a single nanoparticle could have been employed to achieve protein differentiation and that an array of nanoparticles is superfluous.[32]

Despite this, the group continued employing the array for detection of proteins in complex media such as human serum with good differentiation but poor cross-reactivity.[33] Introduction of additional aromatic functionalities to the nanoparticle array and signal generation by the activity of β-galactosidase on a sugar-quenched fluorophore lead to a fluorescence response that exhibited greater cross-reactivity and still allowed for protein classification using LDA classification.[34]

2.3.2 Differential Sensors for Cells

Differential sensing arrays are capable of discriminating complex mixtures, as a unique pattern is produced in the array that is diagnostic for the mixture.[35] Cell surfaces are complex mixtures of which all of the component structures, various lipids, carbohydrates, and proteins, are not thoroughly defined. The number and type of these structures can differ between cells of different tissue origins, different cancer states, and in response to non-native effectors such as transfection.

Rotello and Bunz expanded the use of arrays of functionalized gold nanoparticles to patterning cancer cells. Their first array consisted of three nanoparticles functionalized with an aryl, alkyl, or alcohol residue and was used to discriminate cancer cells of different tissue origin, breast cancer cells of different metastatic states, and isogenic murine cells of different metastatic states.[36] Except in the latter case, greater than 96% of the classification was captured on the first factor. Discrimination of isogenic murine cancer cells that had been transfected to become cancerous (TD) and metastatic (V14) was achieved using the nanoparticle array, resulting in significant classification on the first two factor axes (Figure 2.6).

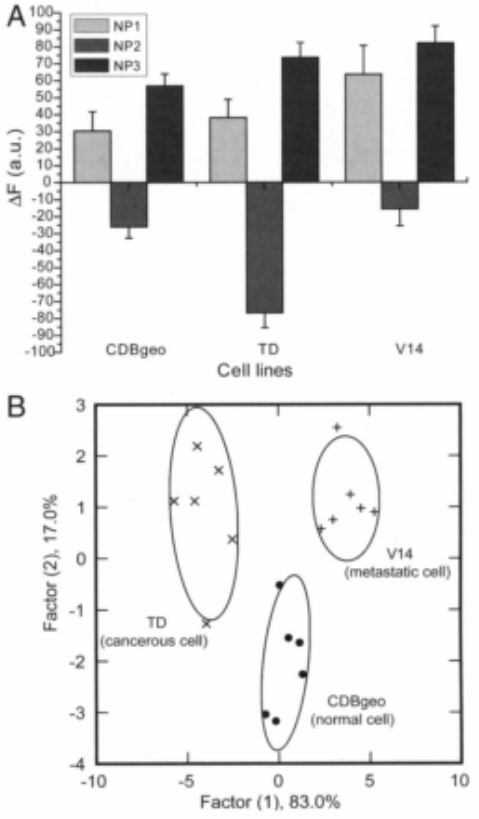


Figure 2.6 Fluorescence patterns generated from displacement by cells of a quenched polymer bound to functionalized nanoparticles (top). LDA classification of cells based on the fluorescence response (bottom)

Modifications of this concept using only the fluorescent polymers themselves,[37] or using the nanoparticles with green fluorescent protein (GFP)[38] for signaling led to classification of cancer cell states using LDA, again with >93% of classification achieved by the first factor. This method was, however, proven effective for measurement on samples of approximately 5000 cells, which was a four-fold increase in sensitivity compared to previous reports. Application of this method to lysates of tissues that were metastatic growths from one tumor type was also pursued. The metastatic sites could be discriminated from each other and from normal tissues (Figure 2.7).

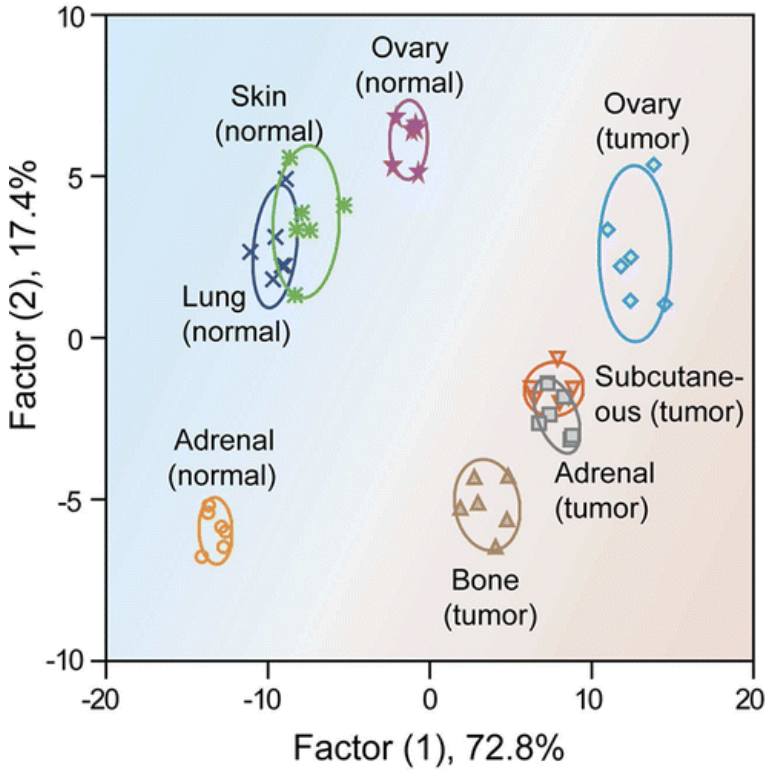


Figure 2.7 LDA classification of tissues from metastatic sites using nanoparticle arrays with GFP.

Zhou *et al.*[39] found that nanoparticles displaying dual-ligands gave better discrimination for different cancer cell types compared to nanoparticles expressing only one ligand, presumably due to heteromeric multivalent interactions. Gold nanoparticles functionalized with folic acid ligands, dual-functionalized secondary ligands, or both were exposed to cancer cells and the gold content measured. More nanoparticles bound when both FA and ligands were present, compared to FA or ligands alone, and this effect was more than additive (Figure 2.8). This illustrates the power of multivalent interactions for discrimination of cells.

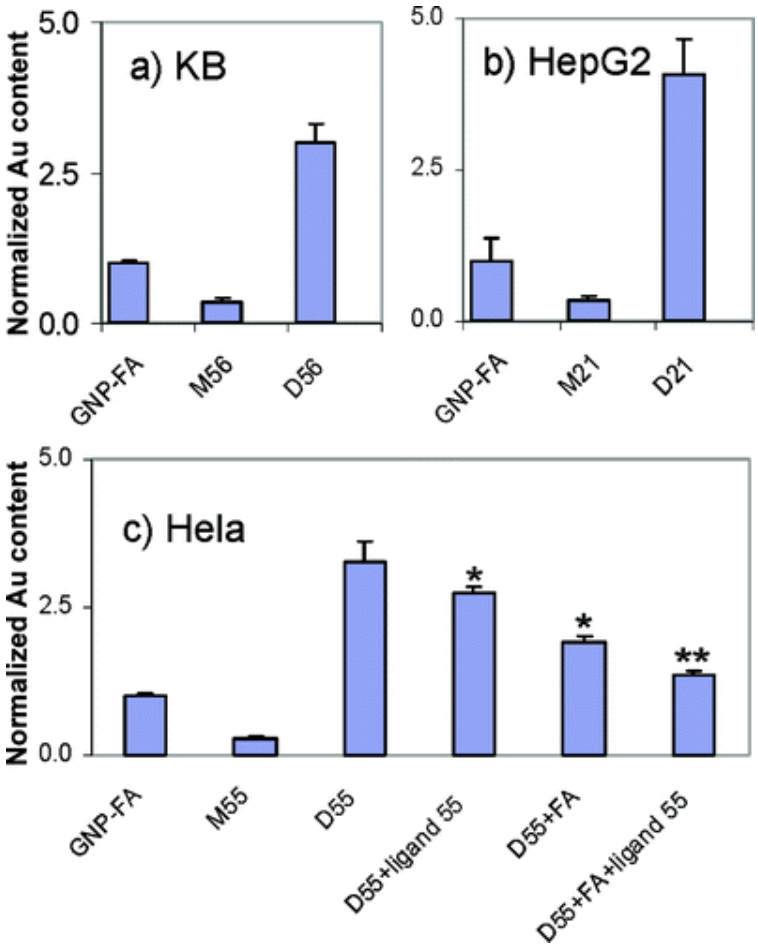


Figure 2.8 Binding of gold nanoparticles to different cell types. Dual functionalized ligands (D) give higher response than folic acid (FA) or secondary ligands (M) alone.

Rotello and Bunz have recently developed differential arrays composed of nanoparticles functionalized with variable sensing groups electrostatically bound to a signaling molecule for patterning proteins[31, 33-34, 40] and cells[36, 38, 41]. These arrays work by displacement of a fluorescent molecule[38] or activation of an enzyme[34] upon nanoparticle binding to the protein or cell and are capable of detection at nanomolar concentrations due to multivalent interactions with the analyte. However, chemometric results show that discrimination is primarily along only one factor axis, suggesting that all the receptors act similarly, possibly eliminating the need for an array.[37] Zhou *et al*.[39] found that nanoparticles displaying dual-ligands gave better discrimination for different cancer cell types compared to nanoparticles expressing only one ligand, presumably due to heteromeric multivalent interactions.

2.4 Chemometric Analysis of Multivariate Data

Part of the power of array sensing arises from the use of multiple receptors; the response of each of these represents a new variable. Although the addition of variables improves analyte classification and discrimination, the data sets generated can often be very large and hard to interpret because each additional variable increases the dimensionality of a data set. In differential arrays, multiple receptors respond to a single analyte and a single receptor is capable of responding to multiple analytes, which leads to multiple variables correlated to a single response and to other variables. In order to reduce the dimensionality of large data sets and improve their interpretability, multivariate data analysis techniques have been developed. When such analysis techniques are applied to chemical data sets, they are considered chemometrics, which is a term that encompasses data- and computer-driven chemical analysis.[42]

2.4.1 Principal Component Analysis

Principal component analysis (PCA) and discriminant analysis (DA) are statistical analysis techniques often employed for interpreting multivariate chemical data. PCA is an exploratory technique that takes p variables and transforms them to q new variables using a variance maximizing rotation; often the goal is to reduce the variables such that $p >> q$ while retaining the essential information of p and removing redundant information.[43] PCA reduces data set dimensionality and reduces variable redundancy by first finding a vector, named a principal component, in the multidimensional variable space that describes the most variance in the data set.[32] Subsequent principal components are defined by vectors that maximize the remaining variability that is uncorrelated, or orthogonal, to that described by previous components. These orthogonal axes are the latent variables and describe decreasing amounts of the variance in the data set. PCA thus represents a basis set transformation where the principal components extracted from the data set are a linear combination of the original variables that describe the maximum variance. This can be represented by:

$$PC_i = b_{i1}x_1 + b_{i2}x_2 + \cdots + b_{ip}x_p \qquad \text{(equation 2.5)}$$

where PC_i is the i^{th} principal component extracted from the data and is formed from the sum of each original variable, x, multiplied by the contribution of that variable, b, to the principal component.

For PCA, either the correlation or covariance matrix is used to describe the variance for the p values of m variables measured across n cases. The variance of a variable $\boldsymbol{X}$ across n cases is given by the square of the standard deviation:

$$var(\boldsymbol{X}) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}_{\boldsymbol{X}})^2 \qquad \text{(equation 2.6)}$$

where x_i is an instance of n on variable $\boldsymbol{X}$ and $\bar{x}_{\boldsymbol{X}}$ is the mean of n measurements of variable $\boldsymbol{X}$. The covariance of two variables $\boldsymbol{X}_a$ and $\boldsymbol{X}_b$ is given by:

$$covar(\boldsymbol{X}_a, \boldsymbol{X}_b) = \frac{1}{n-1}\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i - \bar{x}_{\boldsymbol{X}a})\left(y_j - \bar{y}_{\boldsymbol{X}b}\right) \qquad \text{(equation 2.7)}$$

where x_i and y_j are instances of n on the two variables, $\boldsymbol{X}_a$ and $\boldsymbol{X}_b$, respectively. This equation shows that variance is actually a special case of covariance where the two variables tested are identical. The covariance matrix is square and consists of the covariance values of pairs of variables $\boldsymbol{X}_1$,$\boldsymbol{X}_1$ through $\boldsymbol{X}_m$,$\boldsymbol{X}_m$ for m total variables:

$$\begin{bmatrix} covar(\boldsymbol{X_1}, \boldsymbol{X_1}) & \cdots & covar(\boldsymbol{X_1}, \boldsymbol{X_m}) \\ \vdots & \ddots & \vdots \\ covar(\boldsymbol{X_m}, \boldsymbol{X_1}) & \ldots & covar(\boldsymbol{X_m}, \boldsymbol{X_m}) \end{bmatrix}$$

It is therefore expected that the covariance value of [$\boldsymbol{X}_1$,$\boldsymbol{X}_m$] will be identical to [$\boldsymbol{X}_m$,$\boldsymbol{X}_1$], resulting in a symmetric matrix across the diagonal. Additionally, the diagonal of the matrix is simply the variance of each variable.

The correlation matrix is created in a similar manner to the covariance matrix; however, the correlations are normalized by the product of the standard deviations, according to Pearson's method, resulting in values between -1 and 1.

$$corr(\boldsymbol{X}_a, \boldsymbol{X}_b) = \frac{covar(\boldsymbol{X}_a, \boldsymbol{X}_b)}{\sigma_{\boldsymbol{X}a}\sigma_{\boldsymbol{X}b}} \qquad \text{(equation 2.8)}$$

The diagonal of this matrix represents the correlation of each variable to itself, which should be perfectly positive, resulting in a value of 1.

The variances described by each principal component and the contribution of each variable to that principal component are calculated using singular value decomposition (SVD), which is an eigenvalue problem.[32] Any symmetric matrix can be decomposed into the product of three matrices:

$$A = USU^T \qquad \text{(equation 2.9)}$$

where A is the symmetric matrix, U is a matrix of eigenvectors, S is a diagonal matrix of eigenvalues, and U^T is the transposed matrix of U. The symmetric matrix A is either the correlation or the covariance matrix calculated from a data set consisting of measurements across multiple variables. The eigenvectors in matrix U are the weights, b,

that create new, uncorrelated variables from the original variables, x. These new variables are the principal components. The eigenvectors define a new coordinate system and transform the data in A to reference positions on the new axes.

Scores for each case along each component are calculated by inputting the values across each of the original variables x into equation 2.5. The scores for all of the cases in the new coordinate system can be represented by a score plot of chosen PCs (Figure 2.9). Score plots are useful when PCA is used as a classification tool, as similar cases will be grouped together. The contributions of the original variables to the new, latent variables in the extracted coordinate system are the variable loadings. The loadings for all of the variables across chosen PCs can be represented in a loading plot (Figure 2.9).

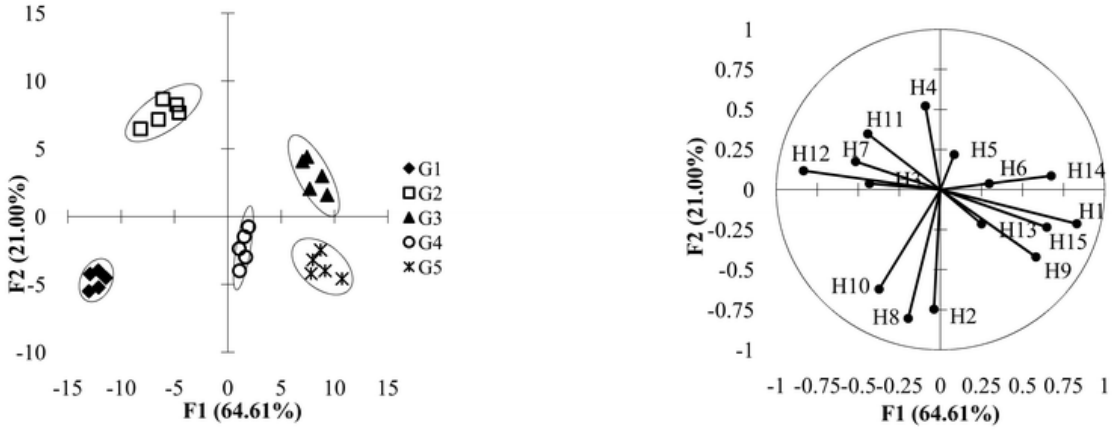


Figure 2.9 *Left:* PCA score plot for measurements on five analytes (G1-G5). *Right:* PCA loading plot of the contribution of receptors (H) to the first two PCs.

Each eigenvector, b, is associated with an eigenvalue, λ_b. The sum of the diagonal eigenvalue matrix S is equivalent to the sum of the diagonal of A because the eigenvalues capture the variance of the new variables, which describe the variance in the original data. The variable space has, however, been transformed such that each eigenvalue corresponds to the variance along a principal component. The importance of that

principal component can be evaluated by calculating the amount of variance it describes by dividing its eigenvalue by the sum of the eigenvalue matrix. Through the use of eigenvectors and eigenvalues, SVD can retain the essential information from a data set while transforming and simplifying it.

2.4.2 Discriminant Analysis

In PCA, each data point is treated in the same manner as other data points. In contrast, for discriminant analyses a dataset is transformed in order to maximize classification of pre-defined groups identified in the data.[32] This is achieved by a combination of variables that define axes that maximize the distances between groups and minimize the distances within groups. Variables are added to the classification model in a stepwise manner if they contribute to the prediction of group membership, using the group means and variances.[44] The mean values of the pre-defined groups across each of the measured variables are calculated, the statistical significances between the means are tested, and if these are significantly different according to a variable then that variable discriminates the groups and is added to the model.

Linear discriminant functions can be determined using Bayes or Fisher discriminant analysis.[45] Bayesian discrimination calculates a posterior probability, $P(l|\boldsymbol{x})$, for a data distribution $\boldsymbol{f}_j$ of k groups that depends on the mean, $\boldsymbol{\mu}_j$, and covariance, $\boldsymbol{\Sigma}_j$. The probability that a measurement $\boldsymbol{x}$ belongs to a particular group l is based on a prior probability, p_j, shown in equation 2.6.

$$P(l|\boldsymbol{x}) = \frac{\boldsymbol{f}_l(\boldsymbol{x})p_l}{\sum_{j=1}^{k}\boldsymbol{f}_j(\boldsymbol{x})p_j} \qquad \text{(equation 2.10)}$$

For linear discriminant functions, an assumption is made that the covariances of different groups are equal. This assumption can be tested using Box's test with the F distribution for more than five classes and/or variables. If this assumption is not valid, quadratic

functions could be determined instead to give an optimal classification rule. (Figure 2.10) Prior probabilities are estimated from the group size divided by *n* total samples, means are calculated by the arithmetic means of the groups, and the covariances are estimated from the pooled sample covariance matrix, S_p, which is a weighted sum of the group covariance matrices, S_k.

$$S_p = \frac{(n_1-1)S_1+\cdots+(n_k-1)S_k}{n_1+\cdots+n_k-k} \qquad \text{(equation 2.11)}$$

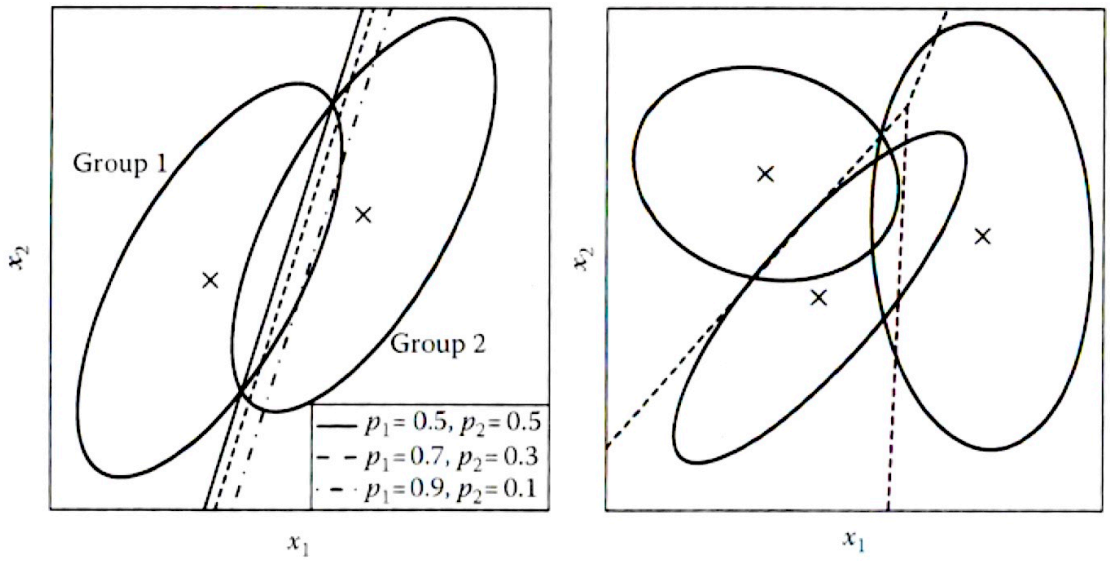


Figure 2.10 *Left:* Covariances of different groups are equal. Different prior probabilities affect the decision boundary. *Right:* Covariances of different groups are not equal; quadratic functions should be defined.

Linear discriminant scores are calculated from classification rules derived by plugging the data into the posterior probability equation 2.10. The linear functions depend on multiplication of a loading vector, $\boldsymbol{b}_{\text{BAYES}}$, with a measurement $\boldsymbol{x}$, and adjusting this term for the mean and the prior probabilities (second and third terms, respectively).

$$d_j(\boldsymbol{x}) = \boldsymbol{b}_{\text{BAYES}}\boldsymbol{x} - \frac{\boldsymbol{b}_{\text{BAYES}}\boldsymbol{\mu_j}}{\mathbf{2}} + \log(p_j) \qquad \text{(equation 2.12)}$$

The loading vector depends on the population means and the pooled covariance. Objects are assigned to a group that gives the largest value of $P(l|\boldsymbol{x})$, which is the largest discriminant score $d_j(\boldsymbol{x})$.

Fisher discriminant analysis does not use the prior probabilities to determine linear functions for classification, and linear functions can be derived for groups without equal covariance matrices, although in such a case the functions will not give optimal minimization of misclassification. Fisher rules are determined when multivariate data is transformed into univariate discriminant variables using linear combinations of the original variables that give maximum separation between groups, similar to PCA (see equation 2.5).

$$y = b_1x_1 + b_2x_2 + \cdots + b_mx_m \qquad \text{(equation 2.13)}$$

Here x-values are the values of a measurement on a variable x of m total variables, and b-coefficients are loadings; $b_1 \ldots b_m$ determine a loading vector $\boldsymbol{b}$ that defines a direction in m-dimensional variable space. This is very similar to PCA, however the Fisher rule uses the variation between the groups, $\boldsymbol{B}$, and the covariance within a group, $\boldsymbol{W}$, to determine the loadings.

$$Separation = \frac{b^T B b}{b^T W b} \qquad \text{(equation 2.14)}$$

This leads to an eigenvalue problem, where $\boldsymbol{b}$ is an eigenvector of $\boldsymbol{W^{-1}B}$ that gives a corresponding eigenvalue of separation. Fisher discriminant analysis therefore allows for dimension reduction by plotting the data along the eigenvectors that give the highest eigenvalues, or the dimensions that result in the best separation of groups.

Assignment of a measurement to a group is based on its Fisher discriminant score. Typically, the Fisher discriminant scores are calculated using Mahalanobis distances to the group means. Mahalanobis distances calculated from each observation to a data center are independent from variable scaling and account for the covariance structure, $\boldsymbol{C}$.

$$d(x_i) = [(x_i - \overline{x})^T C^{-1}(x_i - \overline{x})]^{0.5} \qquad \text{(equation 2.15)}$$

For LDA, the matrix of eigenvectors ***b*** that maximize the separation between groups is used as the covariance structure.[42, 46] Other distance methods can be used with the matrix of eigenvectors, such as Euclidean distance, to calculate discriminant scores.[46] A measurement is assigned to the group that gives a minimum Fisher score.

2.4.3 Summary

PCA is useful for dimensionality reduction of a data matrix in a manner that uncovers the latent variables, which describe the variance in the data structure. The score plots and loading plots can be evaluated visually to give an understanding of the data. For pattern recognition, DA can be utilized to separate objects by a decision surface that is defined by a discriminant variable. For LDA, this surface is a plane that defines a linear latent variable that seeks to maximize the variance between pre-defined groups and minimize the variance within groups. LDA can be used for data sets that have highly correlated variables, and Fisher LDA can be used for data sets where there are more variables than objects per group, as it results in data reduction. These chemometric methods are useful tools for analysis of multivariate data sets, which are produced when collecting data from differential arrays.

2.5 REFERENCES

1. Garrett, R.; Grisham, C. M., *Biochemistry*. 2nd ed.; Saunders College Pub.: Fort Worth, 1999.
2. Iyer, S.; Gaikwad, R. M.; Subba-Rao, V.; Woodworth, C. D.; Sokolov, I., *Nat Nanotechnol* **2009,** *4*, 389-393.
3. Mammen, M.; Choi, S. K.; Whitesides, G. M., *Angew Chem Int Edit* **1998,** *37*, 2755-2794.
4. Ambrosi, G.; Ciattini, S.; Formica, M.; Fusi, V.; Giorgi, L.; Macedi, E.; Micheloni, M.; Paoli, P.; Rossi, P.; Zappia, G., *Chem Commun* **2009**, 7039-7041.

5. Wigler, M.; Navin, N.; Kendall, J.; Troge, J.; Andrews, P.; Rodgers, L.; McIndoo, J.; Cook, K.; Stepansky, A.; Levy, D.; Esposito, D.; Muthuswamy, L.; Krasnitz, A.; McCombie, W. R.; Hicks, J., *Nature* **2011,** *472*, 90-U119.

6. Wingren, C.; Borrebaeck, C. A. K., *Current Opinion in Biotechnology* **2008,** *19*, 55-61.

7. Mulder, A.; Huskens, J.; Reinhoudt, D. N., *Org Biomol Chem* **2004,** *2*, 3409-3424.

8. Kane, R. S., *Langmuir* **2010,** *26*, 8636-8640.

9. Krishnamurthy, V. M.; Estroff, L. A.; Whitesides, G. M., Multivalency in Ligand Design. In *Fragment-based Approaches in Drug Discovery*, Wiley-VCH Verlag GmbH & Co. KGaA: 2006; pp 11-53.

10. Gumbiner, B. M., *Cell* **1996,** *84*, 345-357.

11. Bridgham, J. T.; Carroll, S. M.; Thornton, J. W., *Science* **2006,** *312*, 97-101.

12. Gresham, D.; Dunham, M. J.; Botstein, D., *Nat Rev Genet* **2008,** *9*, 291-302.

13. Koltai, H.; Weingarten-Baror, C., *Nucleic Acids Research* **2008,** *36*, 2395-2405.

14. Cahill, D. J., *Journal of Immunological Methods* **2001,** *250*, 81-91.

15. Osborne, S. E.; Ellington, A. D., *Chemical Reviews* **1997,** *97*, 349-370.

16. *Nat Biotech* **2000,** *18*, IT50 - IT52.

17. (a) Geysen, H. M.; Meloen, R. H.; Barteling, S. J., *Proceedings of the National Academy of Sciences* **1984,** *81*, 3998-4002; (b) Houghten, R. A., *Proceedings of the National Academy of Sciences* **1985,** *82*, 5131-5135.

18. Smith, G. P.; Petrenko, V. A., *Chemical Reviews* **1997,** *97*, 391-410.

19. Uphoff, K. W.; Bell, S. D.; Ellington, A. D., *Current Opinion in Structural Biology* **1996,** *6*, 281-288.

20. Winter, G.; Griffiths, A. D.; Hawkins, R. E.; Hoogenboom, H. R., *Annual Review of Immunology* **1994,** *12*, 433-455.

21. Kay, B. K.; Kurakin, A. V.; Hyde-deruyscher, R., *Drug Discovery Today* **1998,** *3*, 370-378.

22. Arter, J. A.; Diaz, J. E.; Donavan, K. C.; Yuan, T.; Penner, R. M.; Weiss, G. A., *Analytical Chemistry* **2012,** *84*, 2776-2783.

23. (a) Ellington, A. D.; Szostak, J. W., *Nature* **1990,** *346*, 818-22; (b) Tuerk, C.; Gold, L., *Science* **1990,** *249*, 505-10.

24. (a) Bunka, D. H. J.; Stockley, P. G., *Nat Rev Micro* **2006,** *4*, 588-596; (b) Lorsch, J. R.; Szostak, J. W., *Biochemistry* **1994,** *33*, 973-982; (c) Ng, E. W. M.; Shima,

D. T.; Calias, P.; Cunningham, E. T.; Guyer, D. R.; Adamis, A. P., *Nat Rev Drug Discov* **2006,** *5*, 123-132.

25. Hamula, C. L.; Zhang, H.; Guan, L. L.; Li, X. F.; Le, X. C., *Anal Chem* **2008,** *80*, 7812-9.

26. Sefah, K.; Shangguan, D.; Xiong, X.; O'Donoghue, M. B.; Tan, W., *Nat. Protocols* **2010,** *5*, 1169-1185.

27. Fabbrizzi, L.; Poggi, A., *Chemical Society Reviews* **1995,** *24*, 197-202.

28. Wright, A. T.; Anslyn, E. V., *Chemical Society Reviews* **2006,** *35*, 14-28.

29. Zhou, H. C.; Baldini, L.; Hong, J.; Wilson, A. J.; Hamilton, A. D., *J Am Chem Soc* **2006,** *128*, 2421-2425.

30. Wright, A. T.; Griffin, M. J.; Zhong, Z.; McCleskey, S. C.; Anslyn, E. V.; McDevitt, J. T., *Angewandte Chemie International Edition* **2005,** *44*, 6375-6378.

31. You, C. C.; Miranda, O. R.; Gider, B.; Ghosh, P. S.; Kim, I. B.; Erdogan, B.; Krovi, S. A.; Bunz, U. H. F.; Rotello, V. M., *Nat Nanotechnol* **2007,** *2*, 318-323.

32. Stewart, S.; Ivy, M. A.; Anslyn, E. V., *Chemical Society Reviews* **2014,** *43*, 70-84.

33. De, M.; Rana, S.; Akpinar, H.; Miranda, O. R.; Arvizo, R. R.; Bunz, U. H. F.; Rotello, V. M., *Nat Chem* **2009,** *1*, 461-465.

34. Miranda, O. R.; Chen, H. T.; You, C. C.; Mortenson, D. E.; Yang, X. C.; Bunz, U. H. F.; Rotello, V. M., *J Am Chem Soc* **2010,** *132*, 5285-5289.

35. Lavigne, J. J.; Anslyn, E. V., *Angewandte Chemie International Edition* **2001,** *40*, 3118-3130.

36. Bajaj, A.; Miranda, O. R.; Kim, I. B.; Phillips, R. L.; Jerry, D. J.; Bunz, U. H. F.; Rotello, V. M., *P Natl Acad Sci USA* **2009,** *106*, 10912-10916.

37. 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.

38. Bajaj, A.; Rana, S.; Miranda, O. R.; Yawe, J. C.; Jerry, D. J.; Bunz, U. H. F.; Rotello, V. M., *Chem Sci* **2010,** *1*, 134-138.

39. Zhou, H. Y.; Jiao, P. F.; Yang, L.; Li, X.; Yan, B., *J Am Chem Soc* **2011,** *133*, 680-682.

40. You, C. C.; Agasti, S. S.; Rotello, V. M., *Chem-Eur J* **2008,** *14*, 143-150.

41. Phillips, R. L.; Miranda, O. R.; You, C. C.; Rotello, V. M.; Bunz, U. H. F., *Angew Chem Int Edit* **2008,** *47*, 2590-2594.

42. Introduction. In *Introduction to Multivariate Statistical Analysis in Chemometrics*, CRC Press: 2009.

43. Wold, S.; Esbensen, K.; Geladi, P., *Chemometrics and Intelligent Laboratory Systems* **1987,** *2*, 37-52.

44. StatSoft, I., Electronic Statistics Textbook. StatSoft: 2013.

45. Classification. In *Introduction to Multivariate Statistical Analysis in Chemometrics*, CRC Press: 2009.

46. Dixon, S. J.; Brereton, R. G., *Chemometrics and Intelligent Laboratory Systems* **2009,** *95*, 1-17.

Chapter 3: Spectroscopic Discrimination of Multivalent Analytes

The utility of multivalent interactions is illustrated by the evolution of such interations in biological systems. Synthetic multivalent receptors for biological analytes have been developed and have also proven particularly useful for analyte discrimination. Fluorescence signals generated as the response of an array of receptors to a multivalent analyte can be analyzed using chemometrics in order achieve classification based on the pattern of response of the receptor array. We have chosen to develop a dynamic templating process for our receptor array in order to discriminate cell surfaces, a representative multivalent biological analyte. Differential array sensing has been shown to be a powerful tool for discrimination of complex mixtures. The cross-reactive array developed herein was applied to the discrimination of cell surfaces because they are complex mixtures for which the exact composition is not well defined and is variable across cell types. Development of an array of multivalent receptors capable of dynamically arranging on a scaffold was hypothesized to provide better analyte complementarity, which would lead to better discrimination.

3.1 Background

The Rotello and Bunz collaboration resulted in the development of arrays of multivalent receptors for the discrimination of biological analytes such as proteins and cells.[1] These receptors, which consist of functionalized gold nanoparticles or polymers, are static; the recognition groups are not capable of rearrangement in response to a particular analyte. Additionally, the receptors consist of simple recognition units that span basic molecular properties, such as variable hydrophobicity, but do not target known motifs on the chosen analytes. Although the Rotello group has incorporated more

complex recognition units by using peptide-functionalized nanoparticles for protein discrimination, these peptides were simply used to increase the hydrophobicity of their nanoparticles and to allow for chiral surface interactions (Figure 3.1).[2] Nevertheless, such peptides were able to differentiate α-chymotrypsin and cytochrome c, two proteins with variable hydrophobicity, without developing peptides to target specific motifs on those proteins.

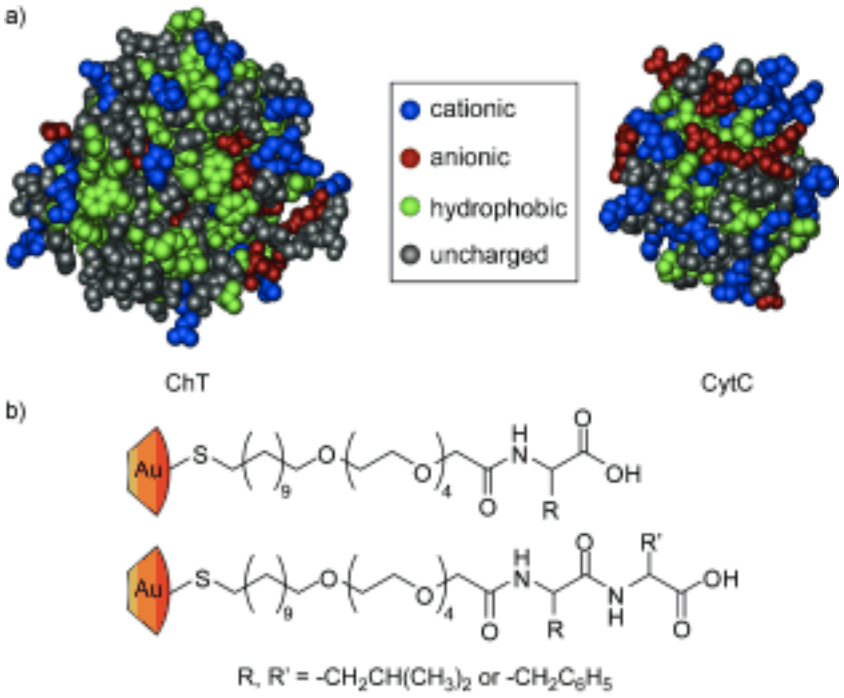


Figure 3.1 Proteins differentiated by peptide-functionalized gold nanoparticles, based on hydrophobic and chiral surface interactions.

3.1.1 Cancer Cells as Multivalent Analytes

Cancer cells exhibit differences in expression of surface proteins, sugars, and lipids between different tissue types and different metastatic states within the same tissue type.[3] Much work has been done to identify specific receptors and expression changes at both the protein- and gene-level that occur to give rise to the altered growth and adhesion properties associated with cancer tissues. Targeting with highly specific receptors, such

as antibodies, has led to the identification of malignant and tumor-associated receptors or changes to receptor expression. However, characterization of the entire cell surface composition and cancer associated surface changes has not yet been achieved. Cell surfaces, therefore, represent complex mixtures that can be differentiated using cross-reactive differential arrays.

Cancer cells of different tissue origin available from the NCI-60 panel were chosen for this study. The NCI-60 is a panel of cancer cells that is typically used for in vitro screening of anti-cancer drugs.[4] The 60 different cancer cell lines comprise leukemia, melanoma, and cancers of lung, colon, brain, ovary, breast, prostate, and kidney tissues. By comparing the response pattern of a tested compound across the panel to those of standard prototype compounds, it is possible to identify a compound's mechanism of action. As molecular targets displayed on each of the cell lines in the panel become better characterized, identification of compounds that interact with these targets can be achieved.

Nine cell lines of different tissue origin were chosen for study, eight of which are available in the NCI-60. These cell lines are: MOLT-4 (blood), SK-MEL-28 (skin), A-549 (lung), HCT-15 (colon), U87-MG-VIII (brain), SK-OV-3 (ovary), MDA-MB-231 (breast), DU-145 (prostate), 786-O (kidney) (Table 3.1). MOLT-4 is a leukemia line isolated from peripheral blood of a 19-year-old male in relapse for acute lymphoblastic leukemia; it is a suspension cell line of T lymphoblasts. SK-MEL-28 is a malignant melanoma isolated from a 51-year-old male. A-549 is a non-small cell lung carcinoma consisting of alveolar basal epithelial cells isolated from a 58-year-old male. These cells are type II pneumocytes that secrete pulmonary surfactant via exocytosis,[5] and are capable of synthesizing lecithin with a high percentage of desaturated fatty acids. They test positive for keratin by immunoperoxidase staining and express glycosaminoglycans

(GAGs) on the cell surface, which are capable of inhibiting coagulation. HCT-15 is a colorectal adenocarcinoma isolated from colon tissue; it is Dukes' type C, meaning the tumor had spread to a lymph node close to the bowel. The cells produce keratin. U-87 MG is a glioblastoma isolated from the brain of a 44-year-old male; U-87 MGΔVIII has been transfected with EGFRvIII to emulate endogenous epidermal growth factor receptor 1 (EGFR) expression *in vitro*.[6] SK-OV-3 is an adenocarcinoma isolated from ovarian ascites of a 64-year-old female. It is considered a "typical ovarian cancer line". The cells are tumorigenic, and although they express estrogen receptors, they are growth resistant to estrogen.[7] They overexpress human epidermal growth factor receptor 2 (erbB-2).[8] MDA-MB-231 is an adenocarcinoma of mammary gland tissue derived from a metastatic site from a 51-year-old female; it expresses EGFR and transforming growth factor alpha receptors (TGFαR). DU-145 is a carcinoma of prostate tissue derived from a metastatic site from a 69-year-old male. 786-O is a renal cell adenocarcinoma derived from the kidney of a 58-year-old male.

Cell Name	Cancer	Cell Type	Invasiveness
MOLT-4	Leukemia	T Lymphoblast	tumorigenic
SK-MEL-28	Melanoma	melanocyte	tumorigenic
A-549	NSCLC	epithelial	tumorigenic
HCT-15	Colorectal	epithelial	metastatic
U87MGVIII	Glioblastoma	glial	malignant
SK-OV-3	Ovarian	epithelial	metastatic
MDA-MB-231	Breast	epithelial	metastatic and malignant
DU-145	Prostate	epithelial	metastatic
786-O	Renal	epithelial	tumorigenic

Table 3.1 Cancer cell lines of various tissue origin, cell type, metastatic potential, and tumorigenicity used in the following studies.

3.1.2 Multicomponent Receptor Design

The goal of this project is to create a dynamic, multicomponent sensing ensemble for the detection of multivalent analytes. The self-assembling multivalent receptor in this system consists of a DNA strand and a peptide-conjugated DNA intercalator, thiazole orange. DNA has been traditionally used as a highly specific sensor, such as the development of aptamers[9] or use of complementary sequences in DNA microarrays,[10] but in this project its utility as a scaffold for creating a differential array of cross-reactive receptors was explored. Several DNA intercalators can dynamically insert along the DNA backbone to provide the most favorable arrangement of the peptides for binding the target analyte. (Figure 3.2) Penta-peptides of variable sequence are attached to thiazole orange through standard coupling reactions to serve as cross-reactive receptors. In addition to the dynamic nature of the sensing ensemble, the peptide motifs were chosen to bind known

structures on the chosen biological multivalent analytes and were also expected to interact in a cross-reactive manner.

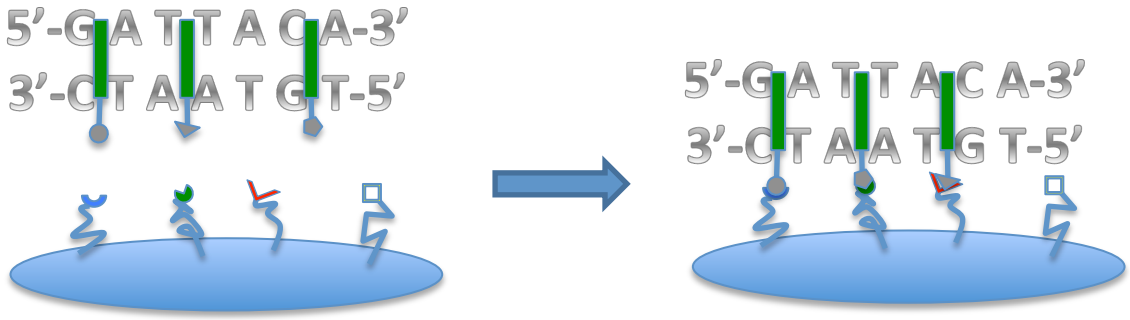


Figure 3.2 A dynamic, self-assembled multivalent receptor consisting of a DNA strand and peptide recognition units, employed for the discrimination of multivalent analytes.

3.1.2.1 Thiazole Orange

Thiazole orange (**TO**) is a cyclic cyanine dye with a monomethine bridge that acts as a DNA intercalator (Figure 3.3). Its mode of intercalation has been characterized as insertion between base pairs, leading to stacking of the benzothiazolium portion with the pyrimidine bases thymine and cytidine and the quinolinium portion with the purine bases adenine and guanine.[11] Substituents on the quinolinium nitrogen protrude from the minor groove of the DNA double strand. Recognition units attached to this position should therefore exhibit directionality, which has been shown to improve multivalent interactions.[12] Thiazole orange has an association constant with double-stranded DNA (dsDNA) in aqueous buffer at physiological pH of 10^6 M^{-1}.[13] The dye has low background fluorescence due to non-radiative relaxation through torsional modes about the methine bridge, but fluorescence emission is enhanced upon intercalation in dsDNA where rotation is restricted, preventing non-radiative relaxation.[14] Fluorescence enhancements of

up to 18,900x upon DNA intercalation have been observed with **TO**,[15] making it an excellent probe for fluorometric studies.

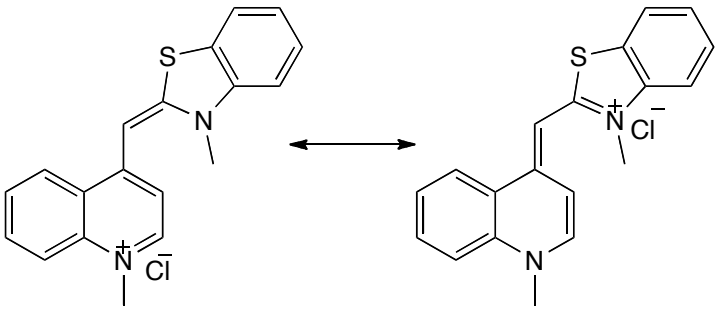
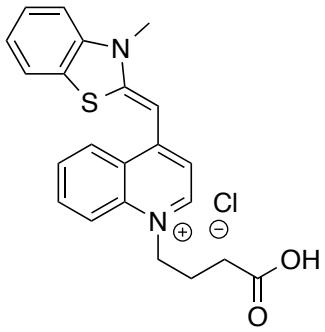


Figure 3.3 The cyanine dye thiazole orange (**TO**).

Conjugates of DNA and cyanine dye intercalators, such as **TO**, have been used to probe DNA hybridization[16] and sequence,[17] elucidate nucleosome formation,[18] and detect aptamer binding to small molecules.[19] Peptide conjugates of **TO** have been synthesized previously and have also been used for nucleic acid studies.[16-17, 20] The properties of thiazole orange intercalation into linear DNA have been well characterized.[13, 21] However, no studies on the interaction of **TO** with supercoiled DNA have been reported, although binding of dyes with similar structures to supercoiled DNA has been studied.[22]

The first goal in the larger scheme of cellular differentiation using dynamic DNA-intercalator complexes was to synthesize and characterize the binding stoichiometry of a carboxylic acid thiazole orange derivative, **TO1** (Figure 3.4), and its peptide conjugates with dsDNA. The carboxylic acid unit of **TO1** serves to create a spacer between the site of intercalation and site of recognition that can serve as a flexible linker while providing a functional handle for peptide coupling. DNA strands of different length and shape were explored and these effects on intercalator stoichiometry were characterized.



TO1

Figure 3.4 The carboxylic acid derivative of thiazole orange (**TO1**).

3.1.2.2 Sequences of Peptide Conjugates

Rather than synthesizing peptide conjugates to possess variability in hydrophobicity, charge, and acidity by incorporating amino acids of specific functionality, peptide sequences were chosen from motifs that had been previously identified in the literature as targeting the test-bed multivalent analytes chosen for this study, cancer cell surfaces.

The Kolonin group developed phage display peptides of random sequence and tested them against the NCI-60 panel of cancer cells.[23] The sequences of peptides after exposure to the cancer cells were determined, and tripeptide motifs were identified. Using the heat map in Figure 3.5, motifs were chosen that showed high binding to one of the chosen cell and displayed cross-reactivity for the others. (Table 3.2)

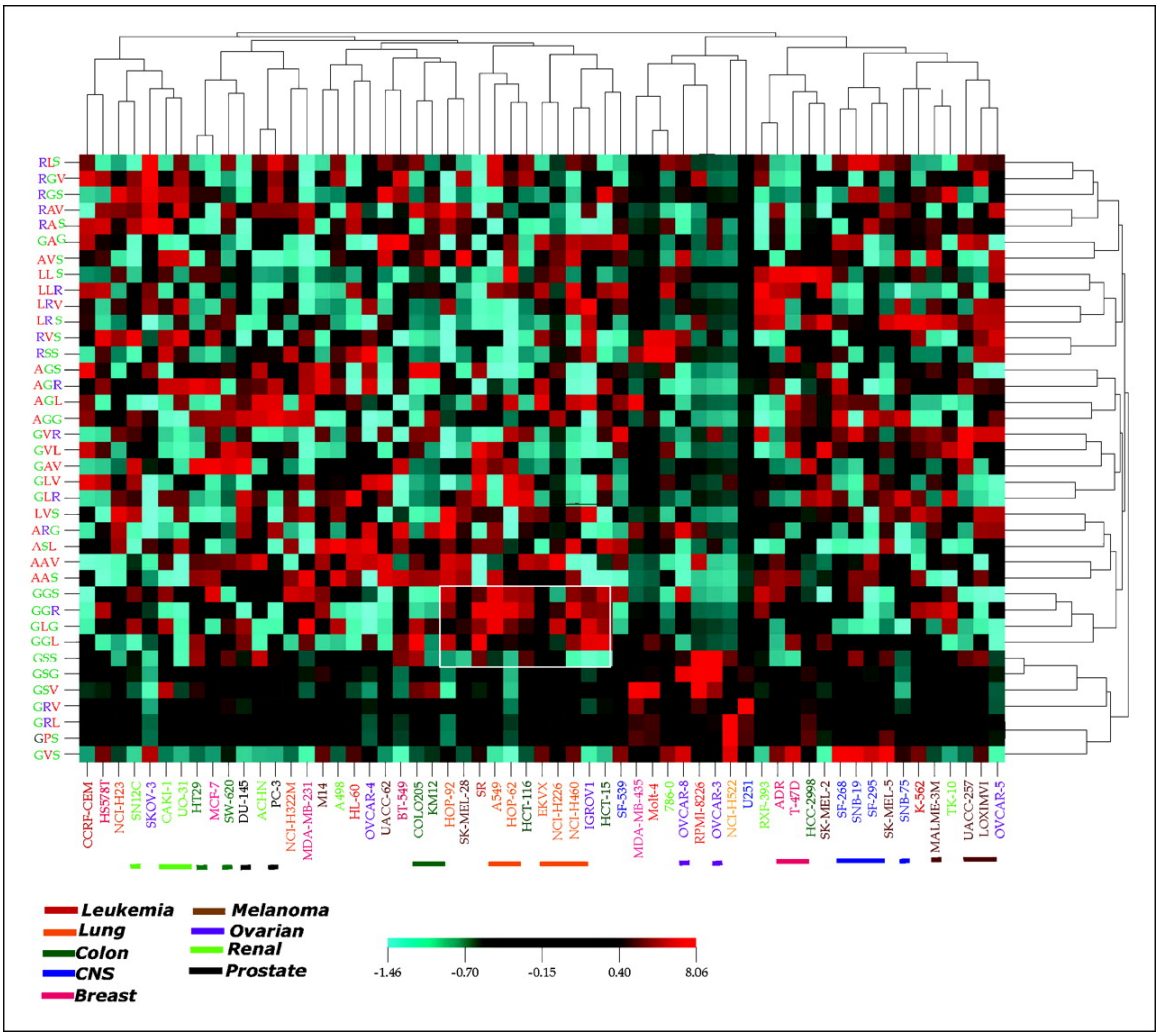


Figure 3.5 Tripeptide motifs identified as highly associated or not associated with particular cancer cell lines from the NCI-60 panel. Figure from Kolonin *et al*.[23]

Peptide (dye-N→C)	Cell Target	Binding Motif
VVKLK	hyaluronic acid	VVKLK[24]
KGGRA	A-549, SK-MEL-28	GGR[23]
KRGSA	SKO-V-3, HCT-15	RGS[23]
PRGDK	integrin	PRGD[25]/RGDK[26]
KDGRC	integrin	DGR[27]
GYPYG	mannose mimic	GYPYG[28]
KRSSA	MOLT-4, 786-O	RSS[23]
KAGLA	MDA-MB-231, DU-145	AGL[23]
KGARC	erbB	GXRC[29]

Table 3.2 Peptide sequences developed for the differential array based on motifs previously identified to bind different cellular targets.

With the discovery of the arginine-glycine-aspartic acid (RGD) sequence involved in cell adhesion, it was established that three amino acids could form an essential recognition site for cells within the context of a large protein.[30] Many adhesive proteins found in the extracellular matrix (ECM), such as fibronectin and vitronectin, use the RGD tripeptide motif to bind to integrins. Integrins are heterodimeric transmembrane proteins that are responsible for cell-cell and cell-ECM interactions, and that provide signals for cell motility, adhesion, proliferation, and apoptosis (Figure 3.6). Integrins are composed of α and β subunits; there are eighteen types of α and eight types of β subunits found in mammals, and these can be combined to form a variety of heterodimers. For

example, the αvβ3 integrin is a vitronectin receptor that is capable of binding to a large variety of proteins that express RGD. Its expression is increased in epithelial cells undergoing angiogenesis, making it a target for tumor inhibition. The α5β1 integrin is a more selective receptor that binds fibronectin and is also associated with angiogenesis.

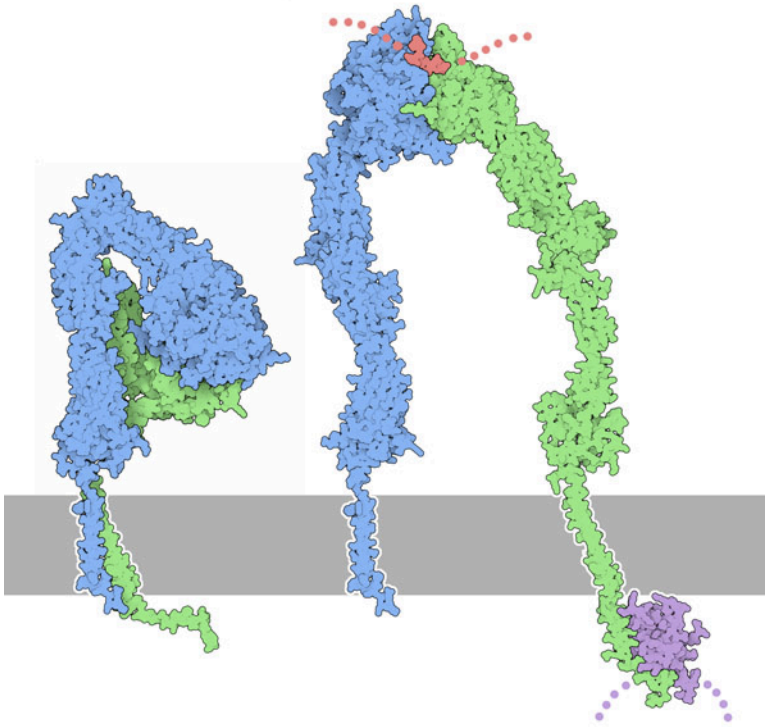


Figure 3.6 The transmembrane heterodimeric proteins known as integrins are composed of α (blue) and β (green) subunits (left). The protein unfolds upon activation, which leads to downstream signaling (right).

Peptide mimics of adhesion proteins can easily be synthesized by incorporation of RGD into short peptides. The smallest active unit is the RGD motif itself, and only a methyl group on the C-terminus of the aspartic acid is required to retain activity. The amino acids that flank the RGD motif can influence the specificity of integrin recognition, resulting in differential binding to different integrin subtypes. For example, PRGD was a conserved motif found in several RGD-containing 15-mers isolated from

phage display against αvβ3, and a 15-mer peptide containing this sequence showed differential binding to α5β1 and αvβ3 type integrins with preference for the latter.[25a] Additionally, PRGD is found in the integrin-binding site of the disintegrins kistrin and dendroaspin, both derived from snake venom, which bind to the fibrinogen receptor αIIbβ3 and inhibit platelet aggregation and blood clotting.[25b] The tetrapeptide motif RGDK was found to provide specificity for α5β1 integrin when appended to a cationic lipophilic headgroup.[26] The specificity was altered by both disruption of the RGD motif and elimination of the lysine.

Non-RGD containing peptides have been found to bind integrins.[30] One of these is the inverted sequence, DGR, which inhibits fibronectin binding to α5β1 to a lesser extent than RGD.[31] However, DGR has been shown to be more effective than RGD at inhibiting the binding of laminin and collagen to fibroblasts.[27] Additionally, DGR is conserved in the collagen domain of surfactant collectin proteins.[32]

Small peptide sequences are capable of acting as recognition sites for proteins, as evidenced in integrin-binding by the tripeptide RGD. Phage display has been used as a tool for identification of short, conserved sequences for other cellular targets. Short peptides have been shown to target common cell surface features that are expressed differently or associated with malignancy, and some of these were chosen from the literature for our differential array. The YPY sequence was identified from phage display studies as a consensus sequence for mimicking the sugar mannose. The authors hypothesize that the tyrosine hydroxyls combined with the hydrophobicity of the sequence provide analogs to the hydroxyls and carbons of the sugar ring.[28] Mannose is a C2 epimer of glucose that is found in post-translational glycosylated proteins. [33]

The sequence VVKLK was identified as a common binding motif in hyaluronan-binding proteins.[24] Hyaluronan, or hyaluronic acid, is an unbranched polysaccharide

classified as a glycosaminoglycan and is composed of repeating disaccharide units of glucuronic acid and *N*-acetylglucosamine (Figure 3.7). Hyaluronic acid is a major component of the ECM and can be found on cell surfaces where it contributes to cell adhesion, motility, and growth. Such characteristics are modified in the development of tumors and metastasis; increased levels of hyaluronic acid are associated with malignancy and poor prognosis of tumors such as breast and prostate cancers, while decreased levels are found in malignant melanomas.[34]

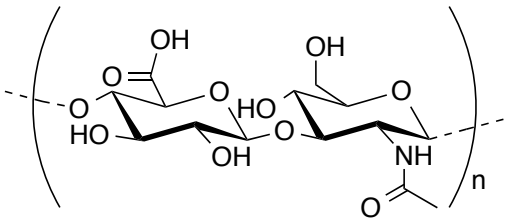


Figure 3.7 The structure of hyaluronic acid, which is composed of repeating units of a disaccharide of glucuronic acid and *N*-acetylglucosamine. Alkyl hydrogens have been omitted for clarity.

GXRC is part of a conserved motif identified in ligands for the tyrosine kinase receptors erbB, of which epidermal growth factor receptor (EGFR or erbB-1) is a member.[29] ErbB-family receptors consist of an extracelluar domain, a transmembrane domain, and a cytoplasmic tyrosine kinase domain (Figure 3.8).[35] Activation of erbB by extracelluar ligand binding leads to receptor dimerization and autophosphorylation, which allows for phosphotyrosine-binding proteins to associate with erbB leading to signal transduction cascades that ultimately modulate cell proliferation, migration, and adhesion. EGFR is amplified, overexpressed, and/or mutated in many cancers, which exhibit altered proliferation, migration, and adhesion. Mutations that activate or amplify EGFR are found with a high frequency in glioblastomas. Deletions in the EGFR genes

are also observed, and a common mutant is EGFRvIII, which is a deletion variant that is constantly active, capable of downstream phosphorylation, and less susceptible to degradation. Effective chemotherapies, such as afatenib, have been developed that inhibit EGFR activity.

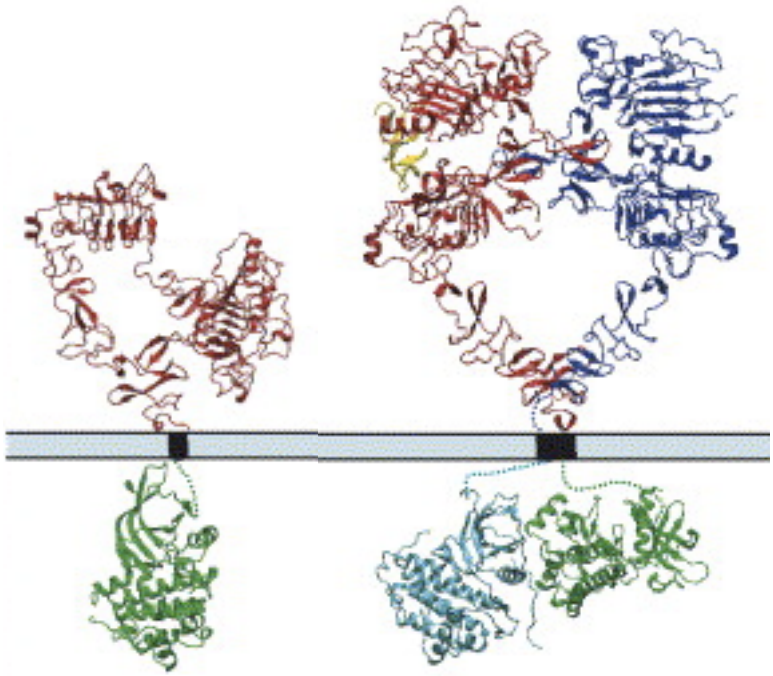
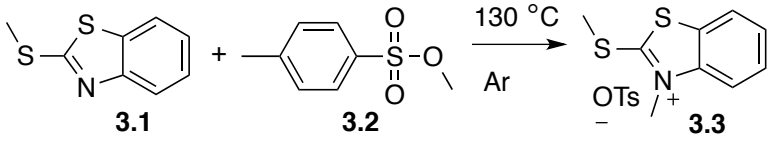


Figure 3.8 ErbB monomer (left) and active heterodimer (right).

3.2 Results and Discussion

3.2.1 Synthesis of a Carboxylic Acid Thiazole Orange Derivative

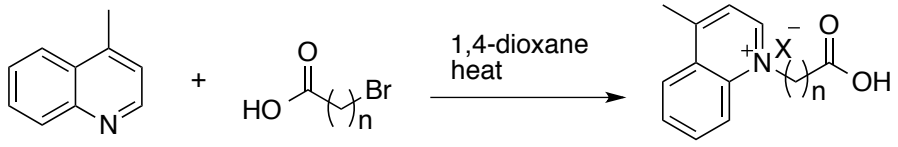
Benzothiazolium salt **3.3** was synthesized in good yield from commercially available starting materials according to the procedures of Bethge *et al*.[17] (Scheme 3.1).



3.1 **3.2** **3.3**

Scheme 3.1 The reaction of **3.1** and **3.2** to form thiazolium salt **3.3**.

The synthesis of quinolinium salt **3.6** proved to be more challenging (Scheme 3.2). Following the procedures of Thompson,[20d] Carreon and Kelley,[20b] and Bethge *et al*.[17] using 5-bromovaleric acid (**3.5d**) and lepidine (**3.4**) resulted respectively in a black tar, the protonated lepidine salt, and no reaction. Modification to the procedure of Carreon and Kelley using dry 1,4-dioxane as a solvent yielded a small amount of desired product **3.6d**, though the isolated solid was mostly protonated lepidine.



3.4 **3.5a-d** **3.6a-d**

	n	product
a	1	1,4-dioxane-2,5-dione
b	2	acrylic acid
c	3	γ-butyrolactone
d	4	δ-valerolactone

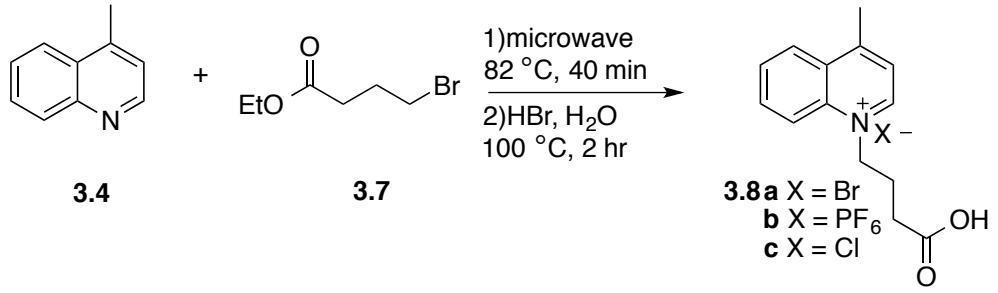
Scheme 3.2 The attempted synthesis of quinolinium salts **3.6a-d**.

To expedite the reaction, the modified procedure of Carreon and Kelley was adapted for microwave use, but the isolated solid still showed low formation of **3.6d**.

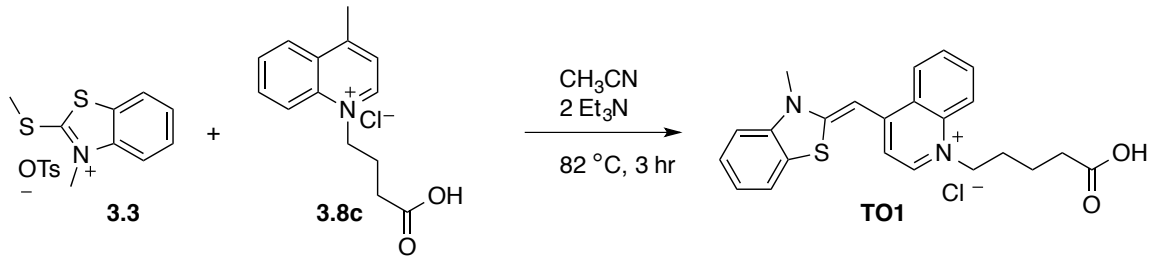
Repeating this reaction using bromoacetic acid as the electrophile also resulted in low formation of product **3.6a**; the major product was the protonated lepidine salt, excluding β-elimination as the sole mechanism for the proton exchange. The crude product from the microwave reaction using bromoacetic acid showed a singlet at 5.52 ppm by ^{1}H-NMR not found in either the starting material or the product **3.6a**. This signal corresponds to a lactone dimer, 1,4-dioxane-2,5-one, formed by intermolecular cyclization of two carboxylates. Both lepidine and bromoacetic acid have similar pK_a values (pK_a 4-5) and undergo proton exchange in solution. Once the carboxylate is formed, cyclization to form a 6-membered ring by displacement of bromide is possible. Byproduct formation is driven by the stability of the lactone dimer and ion pairing of protonated lepidine with the bromide leaving group, which prevents substitution by **3.4** to form **3.6a**. Five and six-membered lactone products can also be observed by ^{1}H-NMR when 4-bromobutyric acid and 5-bromovaleric acid are used, respectively. When 3-bromopropionic acid is used, β-elimination occurs to produce acrylic acid; β-elimination is driven in this case by the stable α,β-unsaturated configuration.

In order to prevent protonation of lepidine, the reaction was run with Hünig's base, but the resulting solid was identified as the protonated salt of Hünig's base. Therefore, ethyl 4-bromobutyrate (**3.7**) was used as the electrophile in order to prevent lactone formation (Scheme 3.3). The ethyl ester of **3.8a** was obtained and hydrolyzed to the acid, but the resulting salt **3.8a** was soluble in water, so the product **3.8b** was obtained by precipitation with KPF_6. This salt was reacted with **3.3** according to the procedures of Thompson and resulted in both the PF_6^- salt of **TO1** and its methyl ester. Due to esterification and the sparing solubility of **3.8b** in methanol, the reaction was repeated in acetonitrile and resulted in the desired product; however, **TO1**-PF_6 had limited solubility in acetonitrile. **TO1** was prepared instead by ion exchanging the quinolinium-PF_6 salt

3.8b for chloride, and then reacting the chloride salt **3.8c** with **3.3** in acetonitrile (Scheme 3.4). Finally, **TO1** was purified by recrystallization from methanol/water.



Scheme 3.3 The synthesis of quinolinium salts **3.8a-c** by reaction of 4-methylquinoline and ethyl 4-bromobutyrate.



Scheme 3.4 Reaction of thiazolium salt **3.3** with quinolinium salt **3.8c** to form a carboxylic acid thiazole orange derivative **TO1**.

3.2.2 Titrations of TO1 and DNA

Characterization of the intercalation properties of the new thiazole orange derivative, **TO1**, with a variety of DNA strands was undertaken. For our purposes, ascertaining the average number of intercalators bound to each strand of dsDNA was important for preparing the dynamic arrays. Therefore, initial fluorescence titrations of **TO1** were undertaken with different lengths of DNA, and the intercalation of **TO1** into supercoiled DNA was characterized. We hypothesized that intercalation of **TO1** into

supercoiled DNA would relax the strand,[36] and thus the supercoiled strand would be able to incorporate more dyes per base pair relative to linear strands.

3.2.2.1 Titration with 53 bp dsDNA

When a 53 bp strand was titrated into a solution of **TO1**, saturation was observed at 0.2 base pair equivalents (bp equivs), corresponding to 1 dye for every 5 base pairs (Figure 3.9). The theoretical maximum of intercalators per strand occurs at 1 dye per 2 base pairs, or intercalation every other base pair. The measured value is lower than the theoretical limit, indicating that the zwitterionic **TO1** may have a repulsive interaction between the carboxylate and the phosphate backbone, resulting in decreased intercalation. Performing the inverse titration, where the concentration of 53 bp dsDNA was held constant, resulted in higher fluorescence values and saturation at 0.35 bp equivs, or 1 dye every 3 base pairs (Figure 3.9). This titration was performed at a higher concentration than the previous titration, indicating that the emission intensity and equivalents are concentration dependent. Alternatively, the increase in fluorescence may be due to contribution from a fluorescent DNA-templated aggregate of **TO1,** as the DNA concentration in the titrant was lower than that of **TO1** for this experiment. DNA-templated aggregates have been reported for cyanine dyes of similar structure.[37]

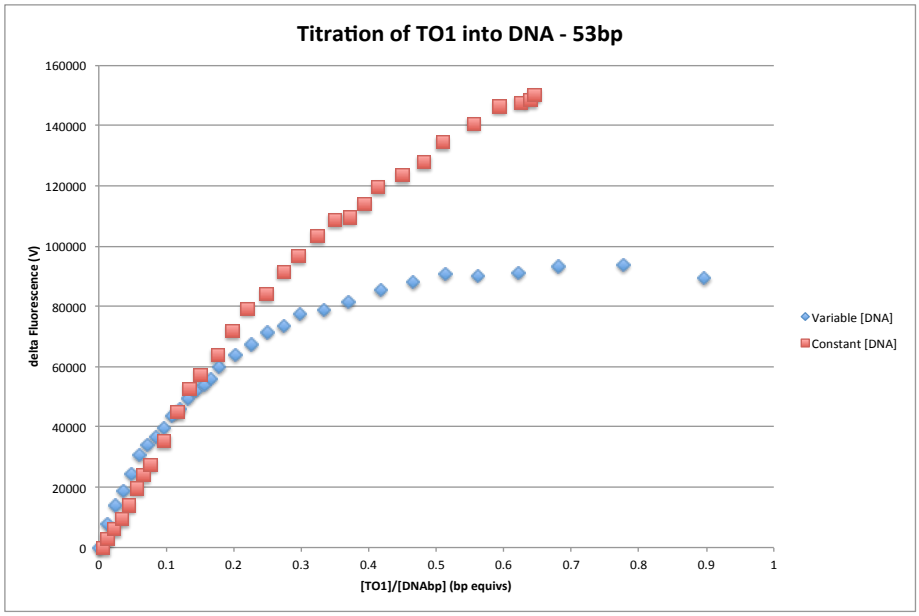


Figure 3.9 Titration of **TO1** into 53 bp dsDNA in DPBS, pH =7.4. [**TO1**]=0-3.4 uM, [DNA]=5.3 uM (red); [**TO1**]=0-2.3 uM, [DNAbp]=5.3-2.6 uM (blue). λ_{ex}=498 nm, λ_{em}=528 nm**.**

3.2.2.2 Titration with 202 bp dsDNA

Using a longer dsDNA consisting of 202 bps, saturation of emission at 528 nm was observed at 0.2 bp equivs, corresponding to 1 dye every 5 base pairs (Figure 3.10). Since this strand is a double stranded aptamer template, it is possible that intercalation was not as favorable due to the potential non-linear structure of the duplex.

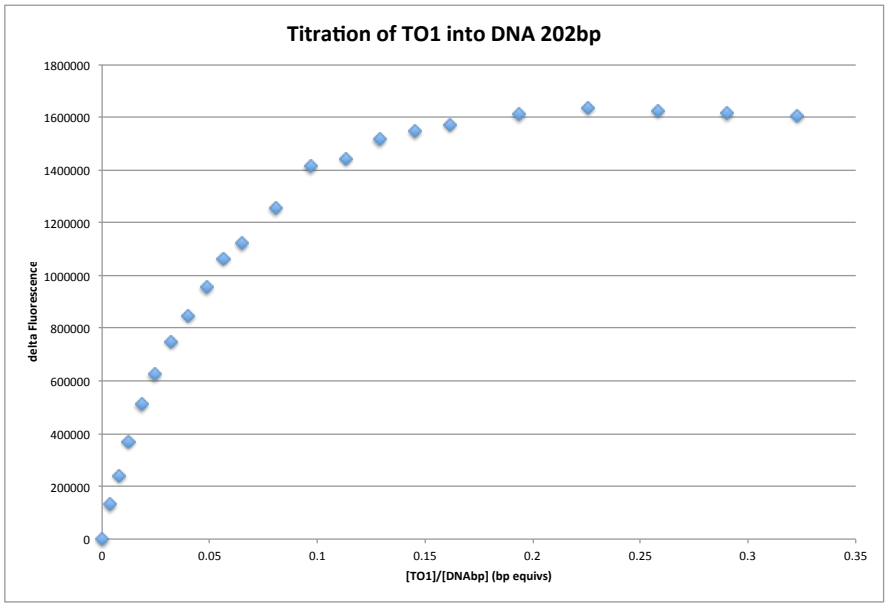


Figure 3.10 Titration of **TO1** into 202 bp dsDNA in DPBS, pH = 4.4. [**TO1**]=0-1.9 uM, [DNAbp]=11.3-6.0 uM, pH=7.0, λ_{ex}=498 nm, λ_{em}=528 nm.

3.2.2.3 Titration with 1000 bp dsDNA

Addition of **TO1** to a solution of 1 kb dsDNA led to an emission increase at 528 nm (Figure 3.11). Saturation behavior was observed at 0.28 bp equivs, corresponding to 1 dye every 3.5 base pairs (Figure 3.12). A continuous variation analysis was performed where the total concentration of **TO1** and base pairs of the 1 kb dsDNA were held constant (4 uM) and the mole fraction of each component was varied. The highest fluorescence was observed when $\chi_{TO1} = 0.50$, which corresponds to 1 dye per every base pair (Figure 3.13); this is an unexpected result since the theoretical limit of intercalation for a mono-intercalator is one dye every two base pairs.[15] Once again, this increase over the theoretical limit may be due to DNA-templated aggregation, as the total concentration of dsDNA was 4 mM, much higher than that used in the titration experiments.

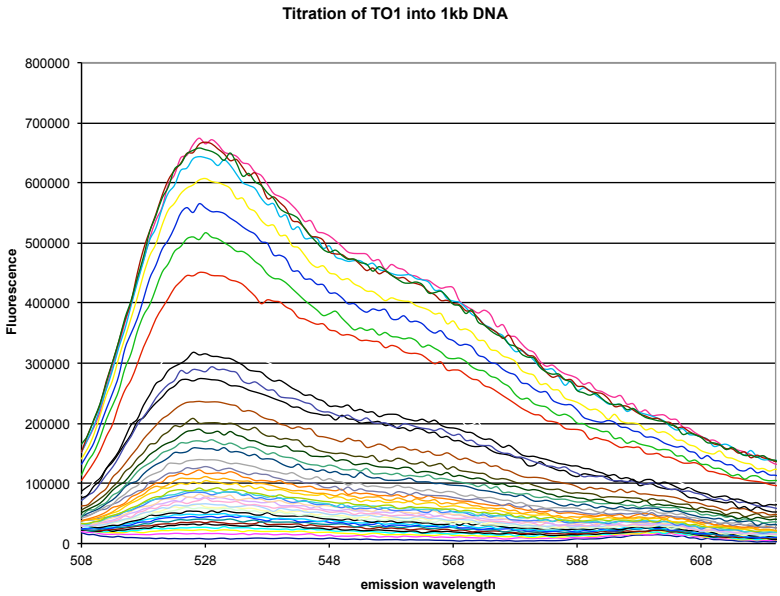


Figure 3.11 Titration of **TO1** into 1 kb dsDNA in DPBS, pH = 7.4. [**TO1**]=0-1.8 uM, [DNAbp]=3.2-2.0 uM, λ_{ex}=498 nm.

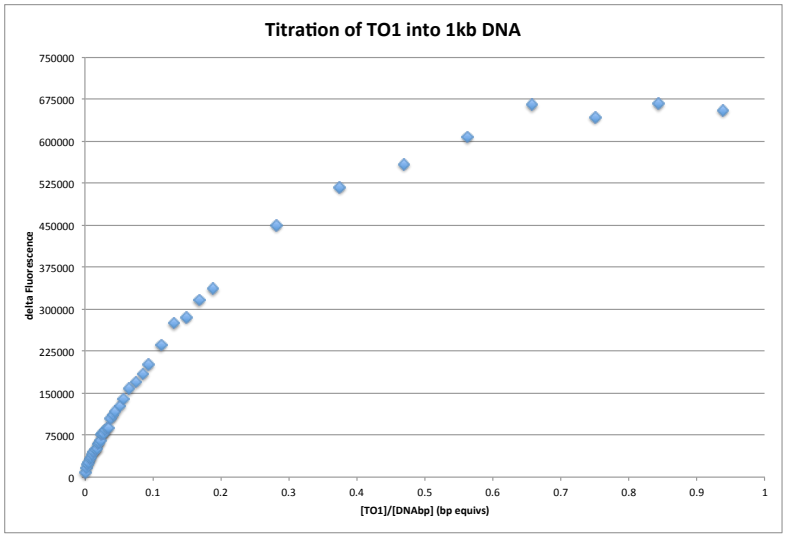


Figure 3.12 Titration of **TO1** into 1 kb dsDNA in DPBS, pH = 7.4. [**TO1**]=0-1.8 uM, [DNAbp]=3.2-2.0 uM, λ_{ex}=498 nm, λ_{em}=528 nm.

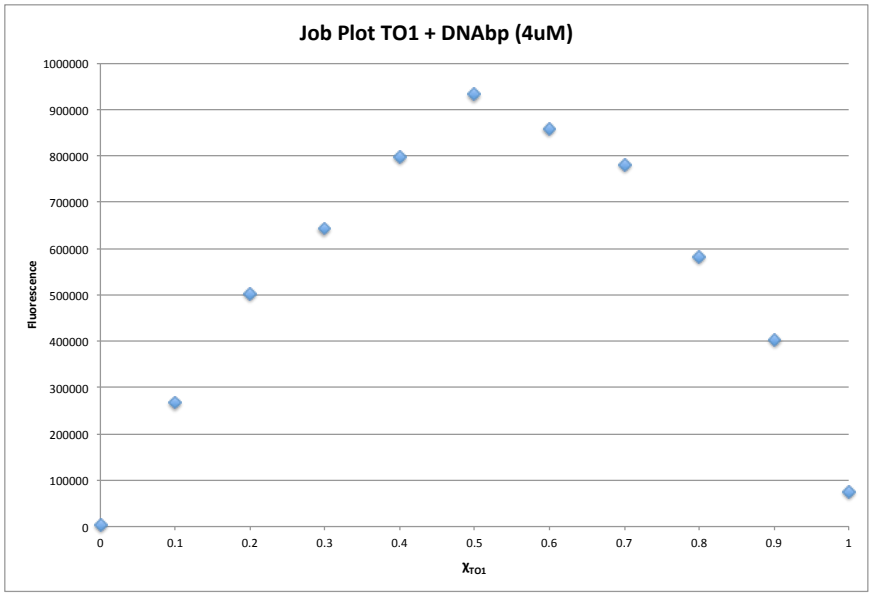


Figure 3.13 Job Plot of **TO1** and 1 kb dsDNA in DPBS, pH = 7.4, [DNAbp] + [**TO1**] is 4.5 uM, λ_{ex}=498 nm, λ_{em}=528 nm.

3.2.2.4 Titration with Supercoiled dsDNA

The intercalation of **TO1** into supercoiled plasmid DNA was investigated. Saturation of the fluorescence signal was observed at 0.25 equivalents, or 1 dye every 4 base pairs (Figure 3.14). The amount of **TO1** required for maximum fluorescence was lower with the supercoiled DNA compared to other strands. This indicates that the supercoiled DNA did not incorporate more **TO1** compared to linear strands.

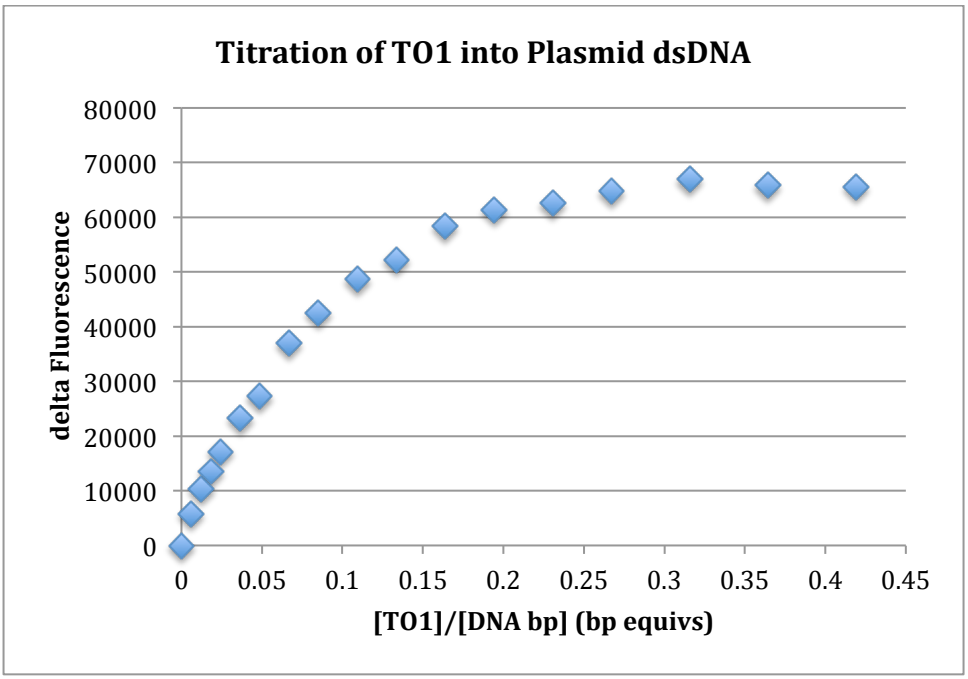


Figure 3.14 Titration of **TO1** into 2686 bp Supercoiled dsDNA. [**TO1**] = 0-1.1 uM, [DNAbp] = 5.0 – 2.6 uM, DPBS, pH = 7.4.

3.2.2.5 Mass Spectrometric Characterization of TO1 Interalation with 14 bp dsDNA

The intercalation of **TO1** into 14 bp dsDNA was analyzed using ESI-MS. These titrations showed that at a 1:1 ratio of dye and DNA, one intercalator is loaded for every DNA strand, which corresponds to 1 intercalator for every 14 base pairs (Figure 3.15). Increasing the ratio of **TO1**:DNA from 1:1 to 5:1 led to an increase in the signal corresponding to the 1:1 dye:DNA complex. At a ratio of 3:1, a peak corresponding to two dyes per DNA strand is evident, or 1 dye for every 7 base pairs, and this peak also increases with increasing dye concentration. At a ratio of 5:1, a peak can be observed for three dyes bound per DNA strand, or 1 dye every 5 base pairs. These results are in good agreement with the solution phase studies, especially considering that anion repulsion between **TO1** and the polyanionic DNA backbone would be stronger in the gas phase.

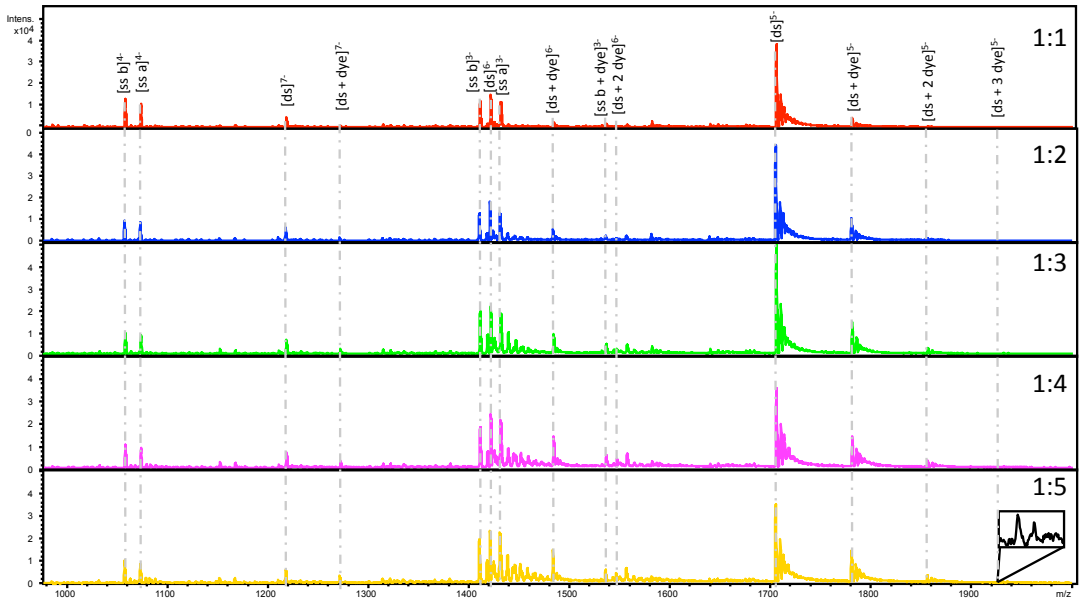
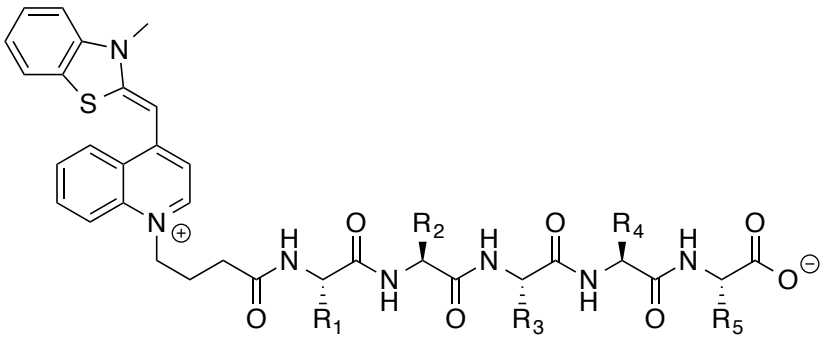


Figure 3.15 ESI-MS of 5 uM dsDNA intercalated with concentrations of **TO1** ranging from 5 – 25 uM in 30 mM NaOAc buffer, pH = 7.0. Duplex DNA is represented by ds and the two single strands are represented by ss a and ss b. The inset provided in the spectrum for 1:5 duplex-to-dye shows an enlarged view of the triply intercalated duplex in the 7- charge state.

3.2.3 Peptide Conjugates of Thiazole Orange

3.2.3.1 Synthesis of TO1-peptides

Solid-phase Fmoc synthesis was used to prepare the nine peptides with sequences shown in Table 3.2. After the peptides were synthesized, the C-terminus was deprotected and the amine terminus was capped with **TO1** via an amide linkage. The peptides were cleaved from the Wang resin and side chains were simultaneously deprotected. HPLC purification was performed on the **TO1-peptides** and the fraction containing the desired sequence was verified using LC/MS (Figure 3.16).



peptide	R_1	R_2	R_3	R_4	R_5
TO1-1	Val	Val	Lys	Leu	Lys
TO1-2	Lys	Gly	Gly	Arg	Ala
TO1-3	Lys	Arg	Gly	Ser	Ala
TO1-4	Pro	Arg	Gly	Asp	Lys
TO1-5	Lys	Asp	Gly	Arg	Cys
TO1-6	Gly	Tyr	Pro	Tyr	Gly
TO1-7	Lys	Arg	Ser	Ser	Ala
TO1-8	Lys	Ala	Gly	Leu	Ala
TO1-9	Lys	Gly	Ala	Arg	Cys

Figure 3.16 Structure of **TO1-peptides** with sequences as shown in Table 3.2.

3.2.3.2 Titrations of DNA and TO1-peptides

Titration of each peptide derivative of **TO1** was undertaken with 1 kb dsDNA in DPBS. Differences in equivalences of peptide required to saturate the fluorescence signal, as well as differences in the saturated emission intensity, were observed for the variable peptide sequences. For example, titrations of **TO1-4** and **TO1-7** with 4 nM 1000mer dsDNA lead to different emission intensities and saturation behaviors. The emission of **TO1-4** reached saturation at 2000-2500 equivalents (2 – 2.5 bp equivs) with a maximum relative emission at 250,000 (Figure 3.17). In contrast, the titration with **TO1-7** reached saturation at approximately 1000 equivalents (1 bp equivalent) with a maximum relative emission of 350,000, which is 1.4x higher than **TO1-4** (Figure 3.18). Although both peptides contain a single lysine, which allows for greater association with the negatively charged DNA, the placement of lysine in **TO1-7** is closer to the site of

intercalation than that of **TO1-4**. The additional positive charge closer to the intercalator may allow for greater intercalation of **TO1-7** compared to **TO1-4**, resulting in greater fluorescence emission and saturation at a lower concentration.

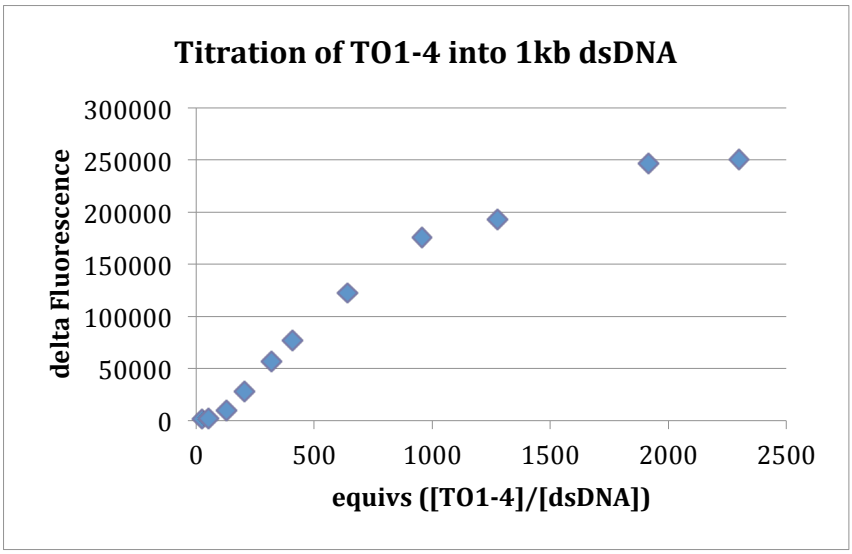


Figure 3.17 Titration of 4 nM 1000 bp dsDNA with **TO1-4**.

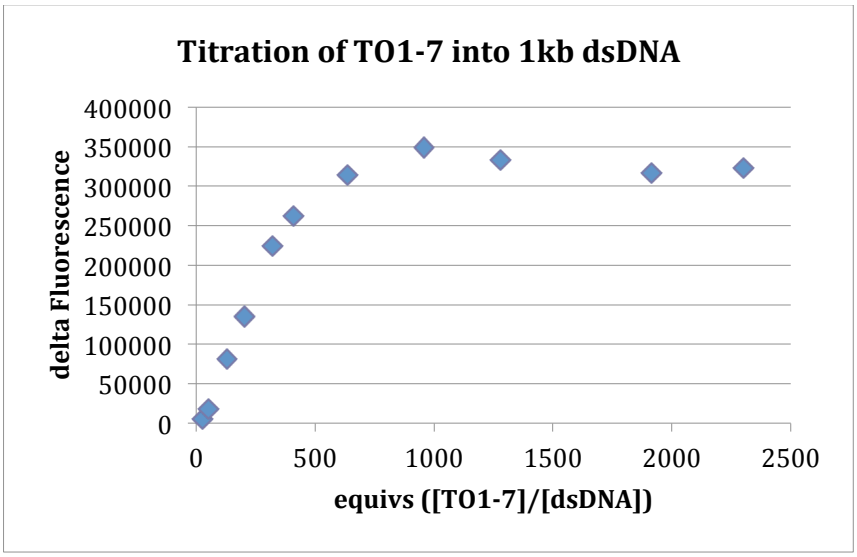


Figure 3.18 Titration of 4 nM 1000 bp dsDNA with **TO1-7**.

3.2.4 Patterning of Three Cancer Cell Lines with a 9-Peptide Array

The next goal of this study was to determine the ability of an array of receptor assemblies to discriminate multivalent analytes using a small subset of analytes. To achieve this, the multivalent analytes used for these studies were three cancer cell lines from the NCI-60 panel of different tissue origin and, therefore, different surface characteristics. SK-OV-3, A-549, and MDA-MB-231 cells were chosen for initial discrimination studies using a differential array consisting of a 1000-mer strand of dsDNA and nine **TO1-peptide** conjugates. The fluorescence response of peptide conjugates **TO1-1** through **TO1-9**, both intercalated into DNA and free in solution, upon incubation with cells was measured, and the values were analyzed using chemometric statistical techniques.

3.2.4.1 Fluorescence Response of the Array

The fluorescence response of the array of peptides to cells was measured after incubation at 37 °C for 40 minutes and subsequent washing with DPBS. These values were compared to the fluorescence values of the solutions that had not been exposed to cells. The DNA intercalated peptides (DNA-peptide) solutions were prepared at concentrations where saturation of the fluorescence signal was observed, resulting in a maximal fluorescence value for each peptide. The absolute fluorescence values were variable, as each peptide sequence has sequence-dependent emission intensity. After incubation with cells and subsequent washing, the emission intensities for all sequences decreased compared to the solutions not incubated with cells. In order to compare the response of the peptides in the array and to eliminate emission intensity bias, the fluorescence values of each peptide were expressed as percentage contributions to the cumulative emission of the array (Figure 3.19). As shown in Figure 3.19, the unexposed peptides have different percent contributions due to their varying emissions, resulting in

variable percent contributions of the cell-exposed peptides. In general, those peptides that had a large percent contribution before exposure to cells continue to have a large contribution after exposure. However, it is clear that the percentage contributions have changed upon exposure to cells and that there are subtle differences between the percentages that vary according to cell type. In order to elucidate these subtle differences, the percentage values were compared between solutions not exposed to cells and those exposed to cells by calculating the difference, resulting in a delta percent (Figure 3.20).

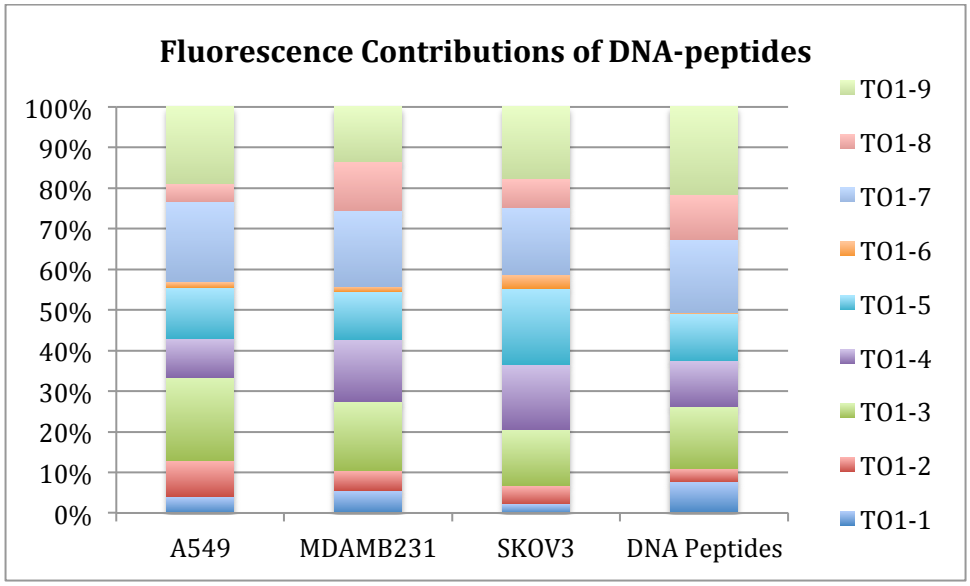


Figure 3.19 Average percent contribution of each DNA-peptide for each cell line, as compared to the contribution of the DNA-peptides not exposed to cells.

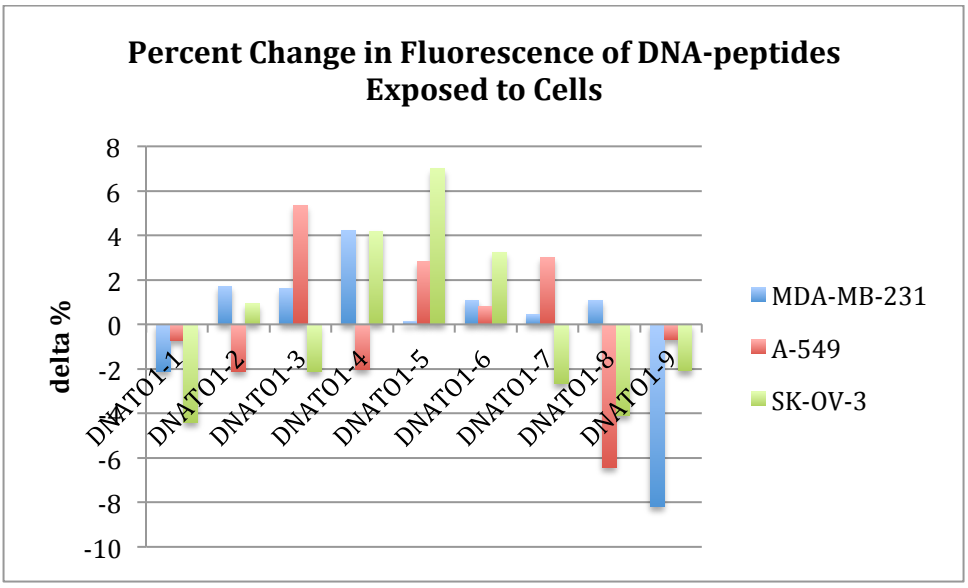


Figure 3.20 Average delta percent response of the three cell lines to the DNA-peptide array.

The free peptide solutions were prepared at the same concentrations as the DNA-intercalated solutions, only omitting DNA. The fluorescence values of these solutions varied slightly based on the concentration used but were much lower than those of the DNA containing solutions. Upon exposure to cells and subsequent washing, the emission intensity of the peptides increased greatly. This is due to the ability of the cationic peptides to associate with the anionic cell surface, resulting in restriction of rotation of the thiazole orange and thus fluorescence emission. The free peptide emissions were also expressed as percent contributions (Figure 3.21), and delta percent values were calculated from the percentages of the non-exposed and cell-exposed peptides (Figure 3.22). The free peptide percentages shown in Figure 3.22 show that emissions of peptides that have not been exposed to cells are more similar to each other than those of unexposed peptides intercalated into DNA. The emission intensities after exposure to cells are not dependent on the original emission intensities of the free peptides, which are representative of the

concentrations. The delta percent values show that there are slight differences in the response of each peptide to the different cells; however, the response across the peptide array is similar for each cell line.

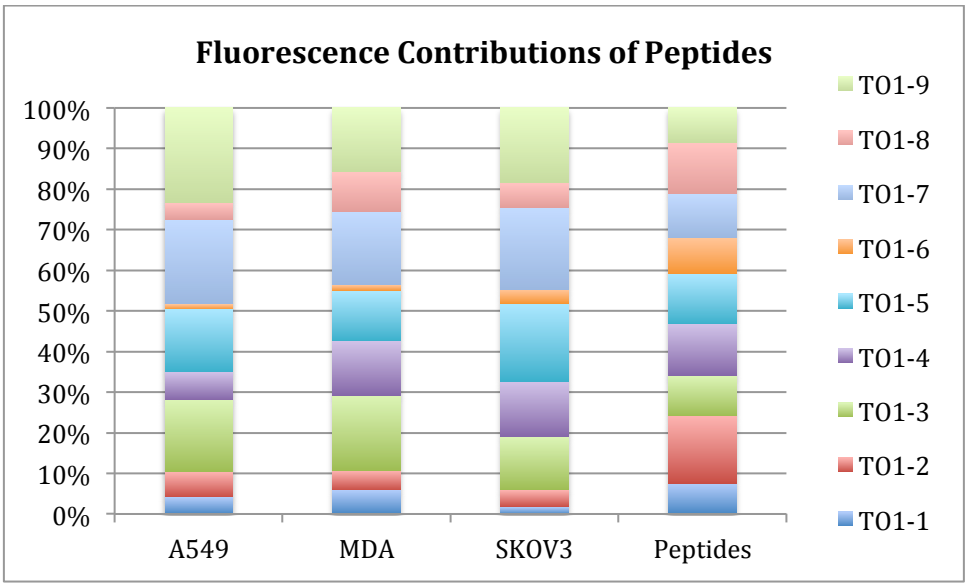


Figure 3.21 Average percent contribution of each peptide for each cell line, as compared to the contribution of the peptides exposed to cells.

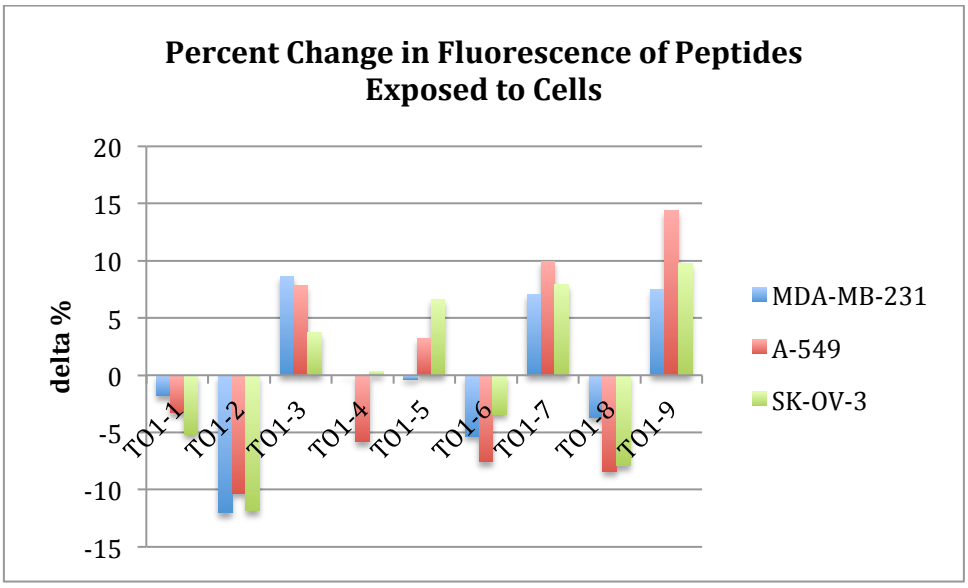


Figure 3.22 Average delta percent response of the three cell lines to the peptide array.

3.2.4.2 Chemometric Analysis

Classification of three cell types MDA-MB-231, SK-OV-3, and A-549 was undertaken in order to validate the discriminatory ability of the array. The response of each peptide in the array, represented as a delta percent value, was used as a variable in Principal Component Analysis (PCA). One of the major goals of PCA is to reduce the dimensionality of multivariate data in order to eliminate redundancy and make the data more interpretable. PCA can also be used as a classification tool, although it is not biased towards grouping cases of similar class and any clusters that appear in the score plot arise naturally from the data.

First, the 18 total variables were separated into two sets in order to investigate the cell-classification ability of the DNA-peptides and the peptides alone. The delta fluorescence values across each variable for nine repeats of measurements on different cell lines were used as the cases. Using the correlation between the variables, the dimensionality of the experimental data sets could consistently be reduced such that

>70% of the original variance was described by two principal components (PCs), with >40% of the variance on the first PC and >30% on the second PC. Figure 3.23 shows the contribution of the variables to the PCs and the classification of the three cell lines using the responses of DNA-peptides in a combined loading plot and score plot, referred to as a biplot. Many variables contribute to each PC but to varying extents, indicating the receptors are behaving in a cross-reactive manner. However, DNA**TO1-1**, DNA**TO1-3**, and DNA**TO1-7** all have similar contributions to the PCs indicating they respond in a similar manner. Importantly, the three cell lines are well separated along the two PCs and repeats of each cell type are clustered, although there is some scatter.

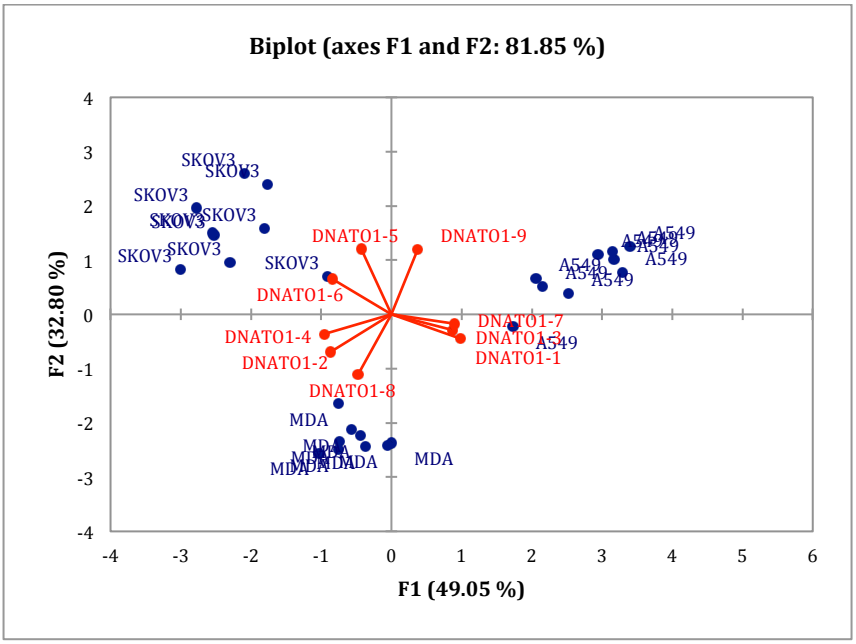


Figure 3.23 Biplot of delta percent response of DNA-peptide array to the three cell lines tested, SK-OV-3, A-549, and MDA-MB-231.

The biplot in Figure 3.24 shows the contribution of the variables to the PCs and the classification of the three cell lines using the responses of free peptides. Again, the

variables have different yet significant contributions to each PC, although **TO1-7** and **TO1-9**, and to a lesser extent **TO1-2**, respond in a similar manner. The three cell lines are also well classified and repeats of measurements are clustered, although again with some scatter.

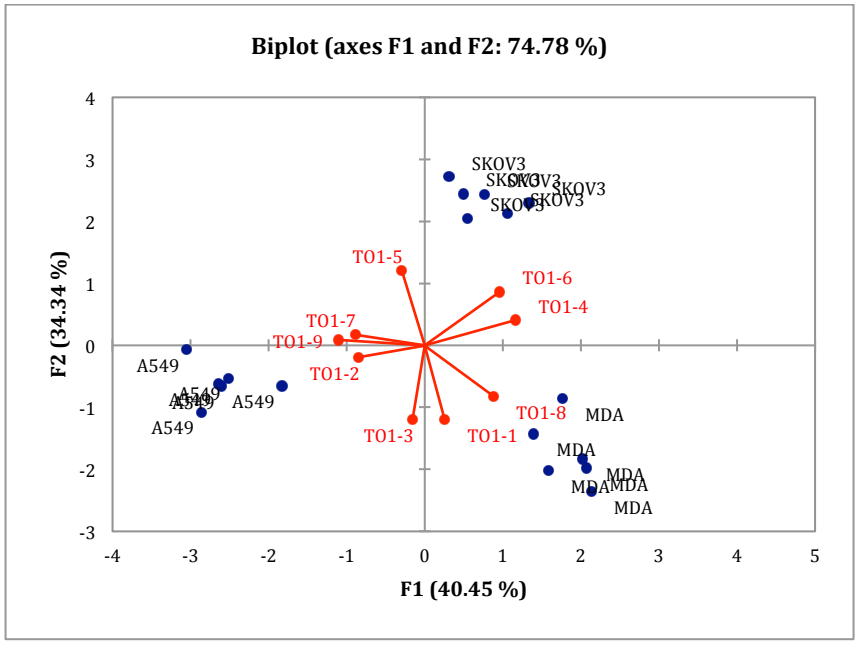


Figure 3.24 Biplot of delta percent responses of peptide array without DNA to the three cell lines tested, SK-OV-3, A-549, and MDA-MB-231.

The responses of the peptides with and without DNA were combined, resulting in 18 variables. For this combination 77% of the variance in the original data was described by two PCs. It is evident from the loading plot that many variables contribute to each of the two PCs to varying extents, indicating the receptors behave in a cross-reactive manner (Figure 3.25). Additionally, combination of the two variable sets did not result in perfect correlation of the response of the free peptides to that of the DNA-peptides, although the responses of **TO1-4**, **TO1-5**, **TO1-6**, and **TO1-8** are similar for both variable sets.

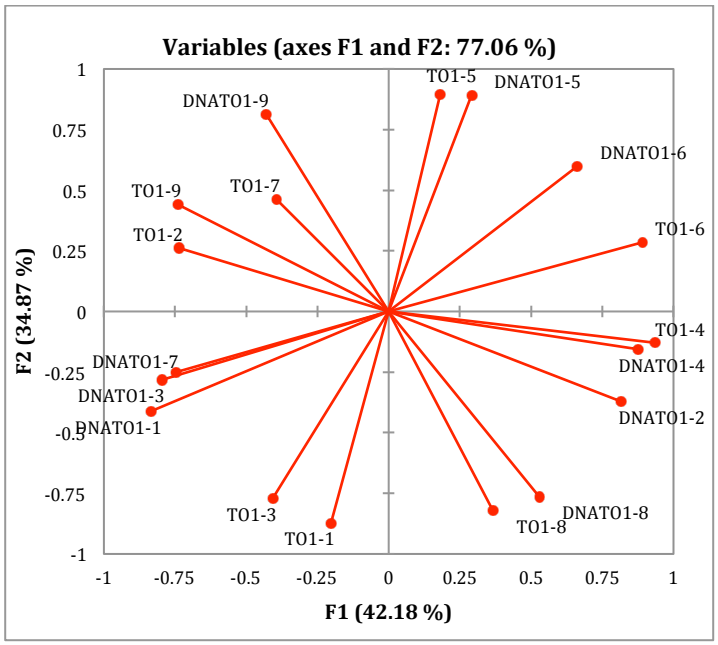


Figure 3.25 Loading plot for discrimination of three cell lines, MDA-MB-231, SK-OV-3, A-549, using the array of nine peptides, with and without 1000mer dsDNA.

Using the 18-variable array, the cell lines were again successfully classified according to cell type on the PCA score plot (Figure 3.26). The three cell lines are well separated on both the first and second PCs. The score of a replicate of a cell line measurement on a PC is a sum of its scores on each variable, which is dependent on each variable's contribution to that PC. A positive correlation of a peptide with a PC indicates that the response of a case to that peptide variable will result in a positive contribution to that case's score on that PC. Similarly, a negative correlation will result in a case having a negative contribution to its score. If a case does not have a strong response to a variable, that variable will not have a significant contribution to its score. **TO1-4**, **TO1-6**, DNA**TO1-4**, and DNA**TO1-2** are all highly positively correlated with the first principal component, F1 (Table 3.2). **TO1-4** contains the RGD motif and was designed to bind to integrin with potential specificity for α5β1 and αvβ3. **TO1-6** contains a conserved

sequence found in peptide mimics of the sugar mannose. **TO1-2** contains the GGR motif that was selected in phage display to bind the cell lines differentially, with preference for A-549 and SK-MEL-28. The positive end of this PC appears to be associated with adhesion properties of the cells. Of the three cell lines, SK-O-V3 scores the highest on this PC, with MDA-MB-231 scoring the second highest. In contrast, A-549 has a negative score on the first PC; DNA**TO1-1**, DNA**TO1-3**, DNA**TO1-7**, **TO1-2**, and **TO1-9** are all highly negatively correlated with this PC (Table 3.3). **TO1-1** was selected to bind to hyaluronan. **TO1-3**, **TO1-7**, and **TO1-2** were all selected to bind cells differentially with preferences for SK-OV-3/ HCT-15, MOLT-4/786-O, and A-549/SK-MEL-28, respectively. **TO1-9** contains the GXRC motif for binding the erbB family of receptors. The negative end of this PC may be associated with cell surface characteristics associated with cell proliferation and survival, as well as specific cell interactions.

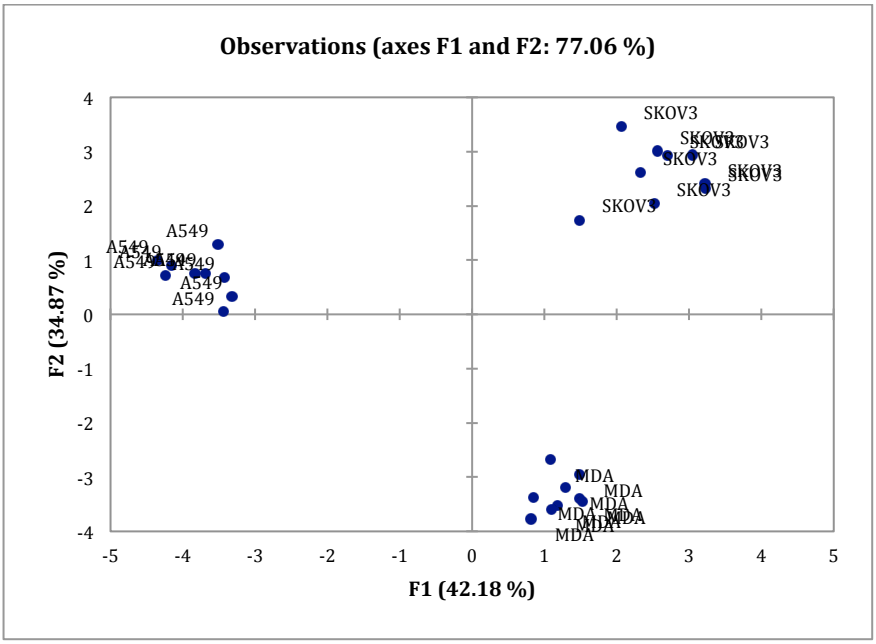


Figure 3.26 The PCA score plot showing classification of three chosen cell lines using the 18-variable peptide array.

	F1		F2
pep4	0.937	pep5	0.894
pep6	0.891	DNApep5	0.892
DNApep4	0.876	DNApep9	0.813
DNApep2	0.816	DNApep6	0.598
DNApep6	0.663	pep7	0.463
DNApep8	0.529	pep9	0.440
pep8	0.366	pep6	0.284
DNApep5	0.291	pep2	0.261
pep5	0.181	pep4	-0.130
pep1	-0.203	DNApep4	-0.156
pep7	-0.394	DNApep7	-0.250
pep3	-0.406	DNApep3	-0.282
DNApep9	-0.429	DNApep2	-0.371
pep2	-0.735	DNApep1	-0.411
pep9	-0.741	DNApep8	-0.765
DNApep7	-0.746	pep3	-0.772
DNApep3	-0.796	pep8	-0.821
DNApep1	-0.834	pep1	-0.873

Table 3.3 Correlations of the peptides with the first two PCs. Variables with a high, positive correlation (>0.7) are colored green, while variables with a high, negative correlation (<-0.7) are colored red.

TO1-5, DNA**TO1-5**, and DNA**TO1-9** are highly positively correlated with the second PC, while **TO1-1**, **TO1-8**, **TO1-3**, and DNA**TO1-8** are highly negatively correlated (Table 3.3). **TO1-5** contains the inverse integrin binding sequence, DGR, which was hypothesized to bind to integrin receptors that recognize laminin and collagen. DNA**TO1-9** contains the GXRC motif for binding erbB-family receptors. **TO1-1** was designed to bind to hyaluronan. These peptides all target cell surface features that have been associated with malignant transformations and are associated with a variety of cell types. The positive end of this PC is therefore likely associated with general malignant transformation. **TO1-8** and **TO1-3** contain motifs for differential cell binding, with

respective preferences for MDA-MB-231/DU-145 and SK-OV-3/HCT-15. The negative end of this PC appears to be associated with specific cell responses. On this axis, SK-OV-3 has the highest score, while MDA-MB-231 has the lowest score.

TO1-4, **TO1-5**, and **TO1-6** contribute to the position of SK-OV-3 on the score plot, indicating that this cell line may express receptors for laminin and/or collagen, such as β1 integrins and the mannose receptor. Studies have shown that SK-OV-3 significantly expresses β1 integrins (96%) and expresses α1-3 (85-92%), α5-6 (94, 79%) and αv (93%) integrins to a greater extent than other ovarian cancers.[38] The cell line has been shown to bind significantly to collagen types I and IV, laminin, and fibronectin, Migration of this cell line towards these ECM adhesion molecules was determined to be mediated by α2β1, α6β1, and α5β1 integrins, respectively.[39] SK-OV-3 also shows a positive response for **TO1-2** containing the GGR motif in the context of DNA, although the response is less positive than for A-549. This motif was selected to bind A-549 and SK-MEL-28, which indicates that this peptide is cross-reactive. Without DNA, **TO1-3** and **TO1-7** both result in positive fluorescence responses to SK-OV-3, though the responses are negative in the multivalent receptor. **TO1-3** contains the RGS motif that was selected for binding SK-OV-3. Although its response is positive, it is the weakest for SK-OV-3 compared to the other cell lines tested and it does not bind well in the context of the multivalent DNA assembly. This indicates that the peptide is not behaving as hypothesized, but is behaving in a cross-reactive manner. **TO1-9** which contains the GXRC motif for binding erbB receptors also contributes to the position of SK-OV-3 on the score plot, although only in the DNA-free case; SK-OV-3 overexpresses erbB-2.[40]

TO1-8 and DNA**TO1-8**, which contain the AGL motif selected for binding to MDA-MB-231, both contribute to the classification of MDA-MB-231 on the PCA plot. The response of DNA**TO1-8** results in a positive delta percent for MDA-MB-231, while

TO1-8 alone results in a negative delta percent. However, this delta percent value is the least negative for MDA-MB-231 compared to the other cell lines, which means it bound better to the cell line it was selected for than the others tested. The greater response when the peptide is intercalated into DNA may be due to the increased valency of the interaction with the cell surface. Other peptides for which MDA-MB-231 has a positive response are DNA**TO1-2** through **4**, DNA**TO1-6** through **8**, **TO1-3**, **TO1-7**, and **TO1-9**. The strongest response of the DNA-containing variables was to DNA**TO1-4**, which contains the RGD-binding motif. MDA-MB-231 expresses αv, β5, and β1 integrins but does not express αvβ3 and does not adhere well to collagen.[41] The presence of β5 integrins may explain why the RGD motif elicited a positive fluorescence response. This effect was only seen in the context of the multivalent receptor, indicating that multiple ligands may be necessary to target integrins. The absence of αvβ3 and low adherence to collagen explains why, with or without DNA, the fluorescence response to the DGR peptide was low. The other DNA-containing receptors that have positive responses were expected to bind other cell lines, which indicates that the peptides are promiscuous. This is also the case for the non-DNA variables **TO1-3** and **TO1-7**. **TO1-9** contains the erbB binding motif and MDA-MB-231 overexpresses EGFR, like many other breast cancers, but does not express erbB-2.[42]

The responses of DNA**TO1-1**, DNA**TO1-3**, DNA**TO1-5**, DNA**TO1-7**, DNA**TO1-9**, **TO1-3**, **TO1-7**, and **TO1-9** determine the position of A-549 on the score plot. A-549 overexpresses CD44,[43] and high CD44 is associated with high HA levels in NSCLCs.[44] **TO1-1** contains a hyaluronan-binding sequence; therefore, this characteristic of A-549 may be why it has a response to DNA**TO1-1** that contributes to its PCA classification. This peptide may only be able to significantly bind to the multivalent HA when presented in the context of a multivalent receptor, which is why **TO1-1** without

DNA did not result in a fluorescence signal enhancement. Although the RGS sequence of **TO1-3** showed strong responses to SK-OV-3 and HCT-15 and weak binding to A-549 in Kolonin's phage display assay, this peptide contributes to the classification of A-549. SK-OV-3, HCT-15, and A-549 are all adenocarcinomas, which are epithelial cancers of glandular tissues. The RGS sequence, in the context of our receptor design, may interact with cell surface features that are common to this type of cancer. The RSS motif of **TO1-7** also shows a fluorescence increase with A-549. Although this motif was selected for MOLT-4/786-O, it differs from the RGS motif by only one amino acid, so it is unsurprising that this peptide binds to A-549 as well. The response to **TO1-5** was positive, although not as strong as that of SK-OV-3. A-549 has been shown to bind well to laminin,[45] which may explain its positive response to the DGR motif in **TO1-5**. **TO1-9** contains the GXRC motif for binding erbB-family receptors and A-549 has been shown to have elevated erbB-2 expression[46] that coincides with elevated, growth-dependent EGFR expression.[47] The **TO1-2** peptide was expected to bind to A-549, but its fluorescence was not enhanced. However, analysis of the loading plot shows that the placement of **TO1-2** without DNA is correlated with the placement of A-549 on the score plot. Interestingly, intercalation of this peptide into DNA seems to remove the expected response to the peptide bearing the cell-line binding motif.

3.2.5 Patterning of Cancer Cell Lines with a 9-Peptide Array

After initial patterning of three cell lines using the array confirmed some of the expected response of peptides to known motifs on the tested cell surfaces, the analytes tested were expanded to include more cell lines of different tissue types.

3.2.5.1 Initial Attempts at Patterning Adherent Cells

In an initial experiment, all cells were plated simultaneously at the same seed density and allowed to grow for the same amount of time, 24 hrs. Each cell line has a different growth rate, therefore the number of cells were counted for each plate and averaged for each cell line. It was observed that some of the cell lines, such as A-549 and U-87 MGΔVIII, had grown more confluent than the others. Many of the cell counts were extremely low, which resulted in decreased signal and increased scatter for the low count and/or low confluency cell types (Table 3.4). This is evident in the PCA score plot (Figure 3.27) which shows the best clustering for U-87 MGΔVIII and A-549, which are the two cell lines which were observed to have retained the most attached cells after the experiment. Interestingly, the LDA score plot is able to reduce this noise significantly, illustrating the power of prior knowledge of classes to improved classification (Figure 3.28). LDA was determined to be the best method of classification as the assumption that the within-class covariance matrices are equal was found to be valid using Box's test. Most of the classification by LDA is explained by the first factor axis, and the cross-validation routine resulted in 87.5% classification, with only A-549, U-87 MG VIII, and MDA-MB-231 correctly classified 100% of the time (Table 3.5).

Cell Line	Average Cell Count	Average Cells Per Well
A-549	109.5	13687.5
U-87 MGΔVIII	20.5	2562.5
DU-145	28.7	3587.5
HCT-15	106.8	13350
SK-MEL-28	12.0	1500
786-O	5.7	2850
SK-OV-3	4.5	562.5
MDA-MB-231	1.3	162.5

Table 3.4 Cell counts for the 8 adherent cell lines. Counts are extremely low for some cell lines.

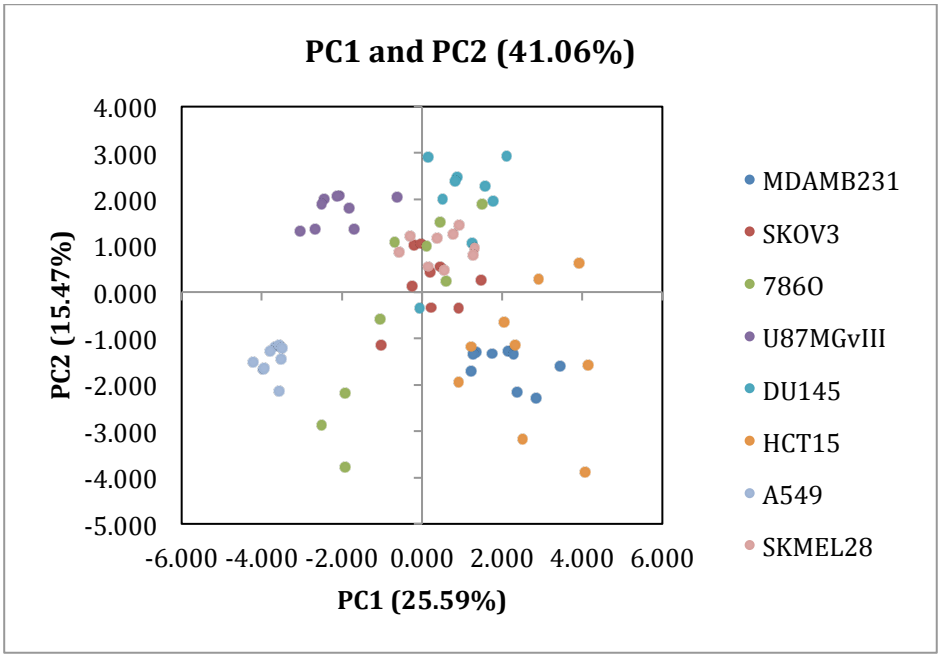


Figure 3.27 PCA on adherent cell response to the 9-peptide array. Cells had become significantly detached over the course of the experiment.

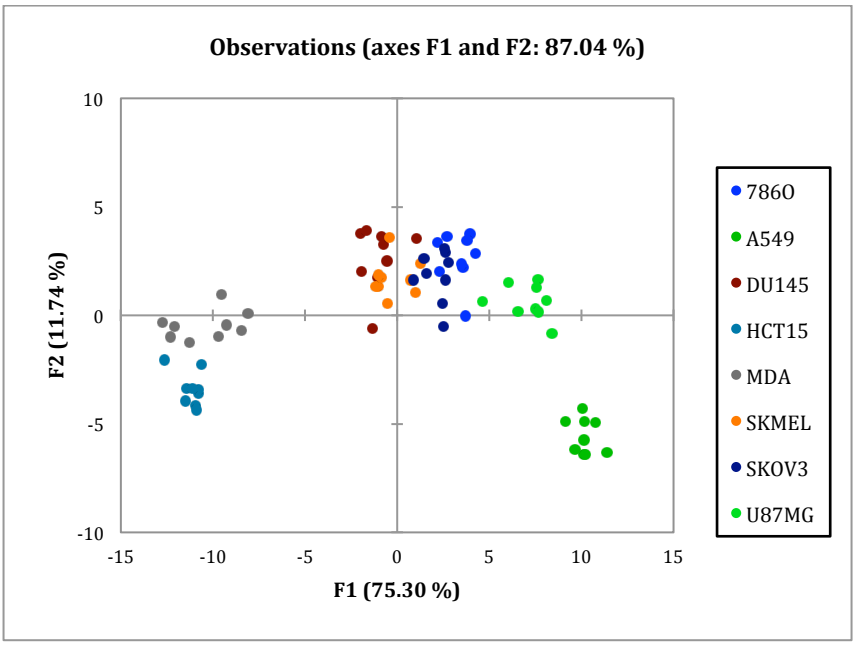


Figure 3.28 Linear Discriminant Analysis (LDA) on highly scattered cell responses to the 9-peptide array.

from \ to	786O	A549	DU 145	HCT 15	MDA	SK MEL	SKO V3	U87 MG	Total	% correct
786O	7	0	0	0	0	0	2	0	9	77.78
A549	0	9	0	0	0	0	0	0	9	100.00
DU145	0	0	7	0	0	1	1	0	9	77.78
HCT15	0	0	0	8	0	1	0	0	9	88.89
MDA	0	0	0	0	9	0	0	0	9	100.00
SKMEL	0	0	0	0	0	8	1	0	9	88.89
SKOV3	2	0	0	0	0	1	6	0	9	66.67
U87MG	0	0	0	0	0	0	0	9	9	100.00
Total	9	9	7	8	9	11	10	9	72	87.50%

Table 3.5 Cross-validation of cell types according to the LDA model built from prior knowledge of the classes.

These results led to the development of experimental conditions where, instead of plating cells simultaneously and allowing them to grow for the same period of time, cells were tested once they had reached confluency. The eight adherent cell lines were each plated at n=6 and allowed to grow until a uniform monolayer was present in each well. Additional wells were grown for cell counts; cells were counted for each plate and averaged over 15,000 cells/well per cell line.

Figure 3.29 shows the 3D PCA plot of the response of each cell line to the 9-peptide array. It is clear that there is a large amount of scatter present for some cell lines, in particular HCT-15 and A-549 and to a lesser extent DU-145. On the first PC, which accounts for 31% of the variance in the data, on average U-87 MGΔVIII and MDA-MB-231 have the highest scores. As seen in the loading plot of the first two PCs (Figure 3.30), the responses of **TO1-2**, **TO1-1**, **TO1-3**, and DNA**TO1-5** are highly positively correlated with this PC. From the loading plot, the response of **TO1-5** is highly negatively correlated with PC1; this peptide has a DGR motif for binding collagen and laminin receptors, but the peptide **TO1-5** may be interacting non-specifically without assembly

on DNA. Although they are not *highly* negatively correlated, DNA**TO1-3** and DNA**TO1-7** are both negatively correlated with PC1 (<-0.45); these two peptides were chosen to bind to the cells with the lowest scores on this PC: HCT-15 and 786-O.

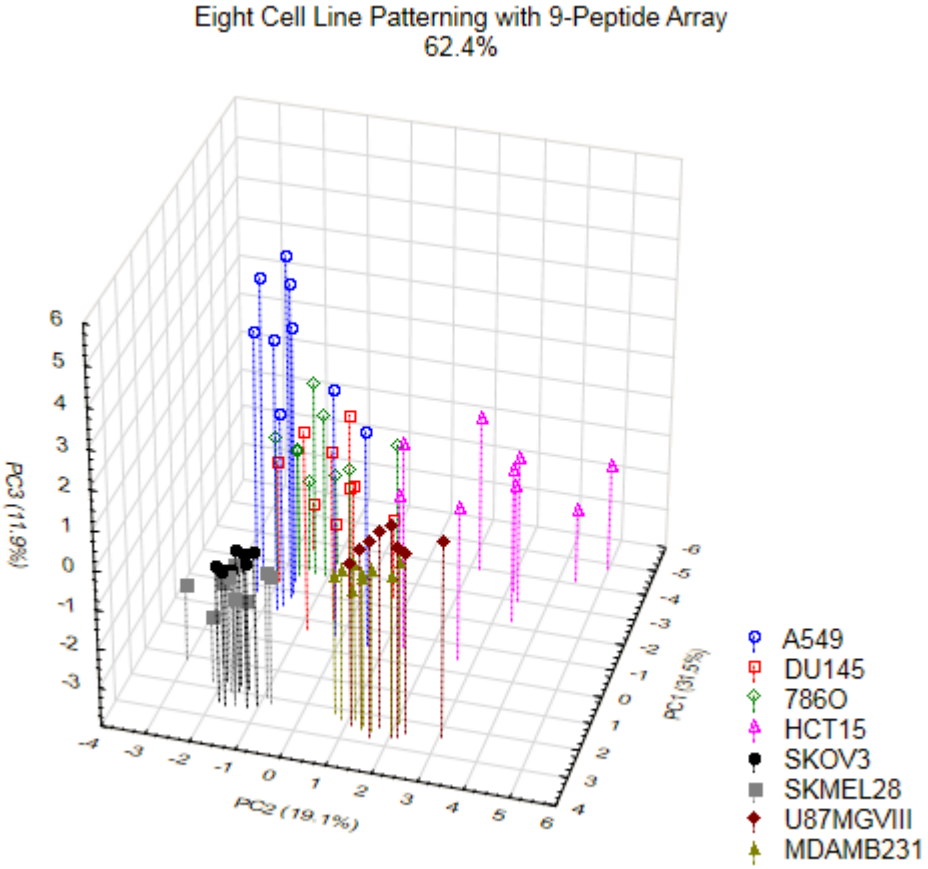


Figure 3.29 Three-dimensional PCA plot of the response of eight adherent cell lines to the 9-peptide array.

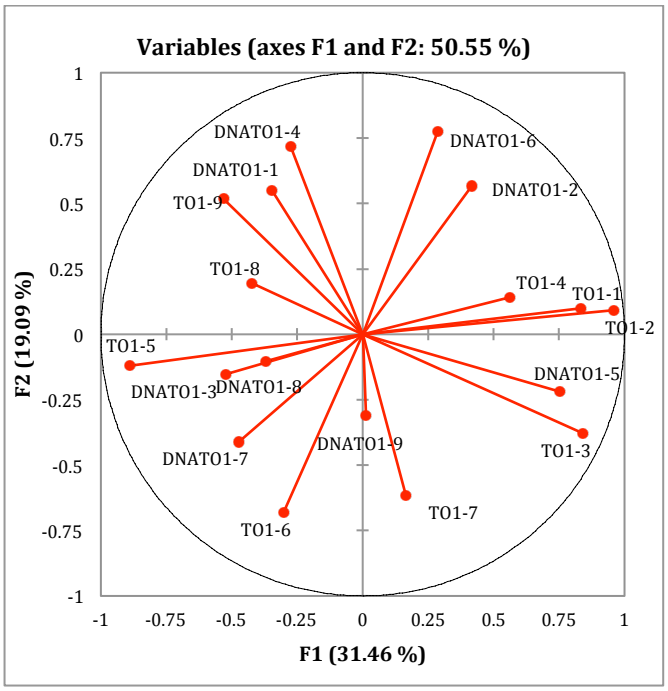


Figure 3.30 Loading plot of the PCA on the eight adherent cell lines grown to confluency and exposed to the 9-peptide array.

Although there are some promising correlations between the responses of some of the peptides with their expected behavior, the negative effects of the experimental procedure on analyte integrity were observable. For both initial eight cell line experiments, cell detachment associated with washing and incubation was observed. Addtionally, when cells were treated with DNA-peptides they detached to a greater extent compared to those that were treated with peptide only or DPBS. This indicates that the peptides are much better at binding to cell surfaces in a way that interferes with their adhesion properties when assembled onto DNA, creating a dynamic multivalent receptor. Unfortunately, such detachment when using the variable set containing DNA results in a lower fluorescence response, increased scatter, and a loss of analyte that cannot be accounted for. In order to eliminate this confounding variable, cells were grown on cell

culture treated plates coated with poly-D-lysine. This polycationic coating enhances the attachment of cells to the growth surface.

3.2.5.2 Patterning of Nine Cell Lines with a 9-Peptide Array

Cells were grown to confluency on poly-D-lysine plates and once again exposed to the 9-peptide array with or without DNA. Classification of the cells according to cell line was achieved using both LDA (Figure 3.31) and PCA (Figure 3.32). For LDA, 75% of the classification occurs on two factors with 61% of the classification on the first factor (F1). In a cross-validation routine, the replicates were correctly classified according to the model 100% of the time.

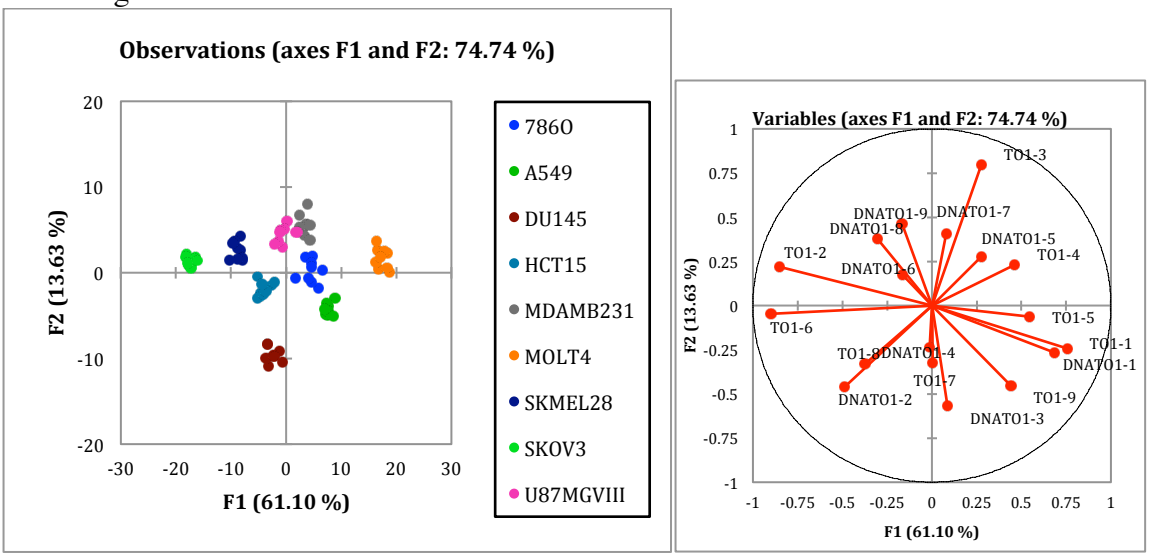


Figure 3.31 LDA score plot (left) and loading plot (right) for classification of nine cell lines using a 9-peptide array.

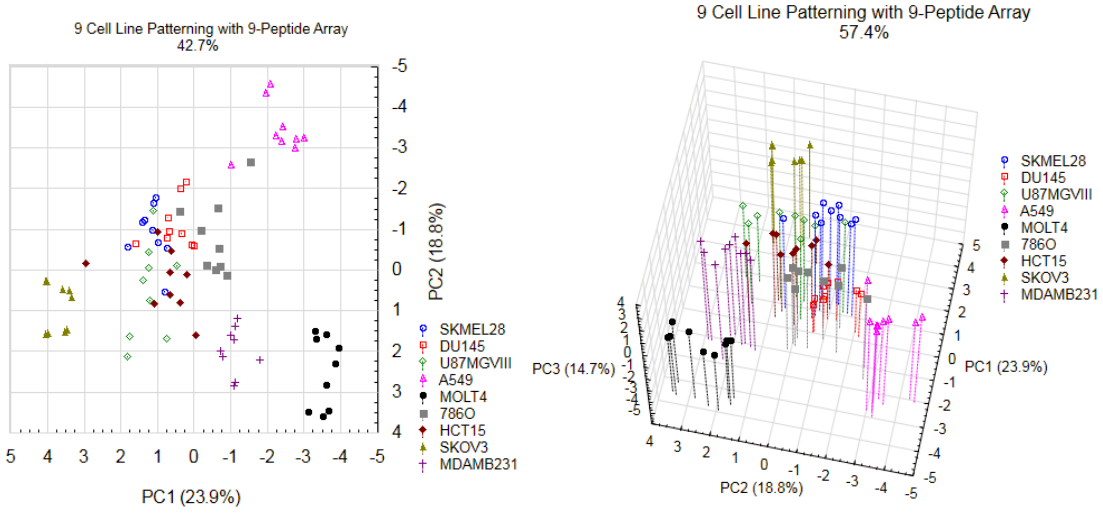


Figure 3.32 PCA plots in two dimensions (left) and three dimensions (right) of nine cancer cell lines exposed to the 9-peptide array. Three dimensions account for 57.4% of the variance in the data set.

Classification was also achieved using the unsupervised method PCA, which indicates that the variance in response of the array is diagnostic for each cell line. Analysis of the PCA loading plots (Figure 3.33) in conjunction with the score plots reveals several contributions to the classification model. From the loading plots it is evident that several variables contribute to each principal component to varying extents and have loadings on multiple components. This means that the array was behaving in a cross-reactive manner. In two dimensions, there is overlap of the responses of SK-MEL-28, DU-145, U-87 MGVIII, and HCT-15 (Figure 3.32). These groups are better classified with the addition of a third principal component. Classification continues on the fourth and fifth PCs; however, the scatter increases as these PCs account for less variance and thus include more relative error (Figure 3.34). Over 70% of the variance in the data is accounted for with the inclusion of the first five components of the model, further supporting the cross-reactivity of the array response.

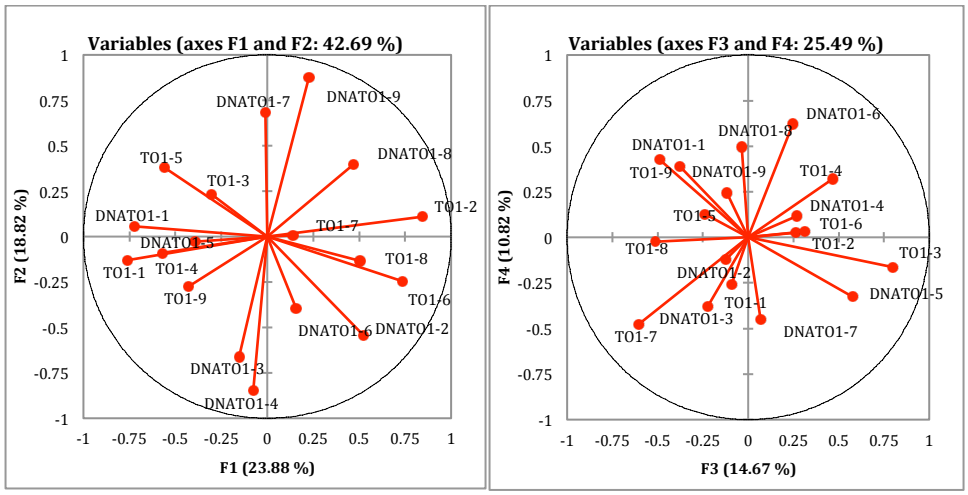


Figure 3.33 Loading plots of array variables on the first four PCs, accounting for 68% of the variance in the data set.

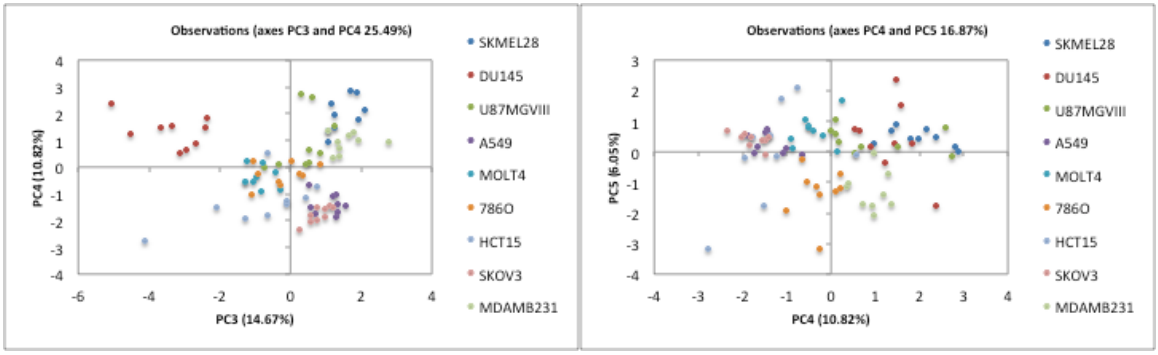


Figure 3.34 PCA score plot of nine cell lines on the third, fourth, and fifth principal components. Addition of PC5 accounts for 74% of variance in the data set.

The variance on each of the retained components contributes to classification. On PC1, **TO1-2** and **TO1-6** have the most positive correlation, while **TO1-1** and DNA**TO1-1** have the most negative correlation. SK-OV-3 has the highest score on this axis, while MOLT-4 and A-549 have the most negative scores. The positions of SK-OV-3 and A-549 on the score plot were previously shown to be influenced by their responses to **TO1-6** and DNA**TO1-1,** respectively. DNA**TO1-2**, **TO1-8**, and DNA**TO1-8** are also positively correlated with PC1, although not as strongly as **TO1-2** and **TO1-6**. The

former peptides were chosen to bind to particular cell lines, A-549/SK-MEL-28 and DU-145/MDA-MB-231, respectively. Although the response of these cell lines to the two peptides does not result in pairing of the cell lines on this axis, both SK-MEL-28 and DU-145 have positive scores. **TO1-6** was selected to act as a mimic of the sugar mannose, but it is also the most hydrophobic of the peptides. Cell lines that score positively on PC1 may display receptors for binding mannose, or may have an abundance of other receptors that provide hydrophobic contacts.

The loading of **TO1-1** onto PC1 indicates that cells that have negative scores on PC1 may display hyaluronan or a similar feature that binds the peptide sequence VVKLK. This sequence also has the most positive charge directed away from the site of intercalation, indicating that cell lines that score negatively on PC1 may have more negatively charged surfaces. **TO1-4** and **TO1-5** are also negatively correlated with PC1; both peptides were designed to bind to integrins. As discussed earlier, A-549 binds laminin and its negative score on PC1 may be due to binding of **TO1-5**. MDA-MB-231 and 786-O display integrins that bind to the RGD motif of **TO1-4**,[48] and part of the invasiveness of these cells comes from the expression of $\alpha 5\beta 1$ integrin.[49]

On PC2, DNA**TO1-9** is highly positively correlated and DNA**TO1-4** is highly negatively correlated. MOLT4 and MDA-MB-231 have the highest scores, while A-549 has the lowest. The **TO1-9** peptide was designed to target erbB receptors. Breast cancers, such as MDA-MB-231, overexpress erbB receptors,[42] which indicates that cell lines that score positively on PC2 may express erbB receptors to a greater extent than those that score negatively. **TO1-4** targets the surface adhesion protein integrins. This PC may be indicative of the adherent properties of the cell lines, as it was observed in cell culture that A-549 was the hardest to enzymatically detach from the growth surface. MOLT-4 is a suspension cell line, and it was observed that MDA-MB-231 frequently detached

during initial experiments. Additionally, MDA-MB-231 is highly metastatic and altered adhesion properties of metastatic cells have been shown to contribute to their motility and invasive potential.

On PC3, MDA-MB-231 and SK-MEL-28 have the highest scores, while DU-145 has the lowest. **TO1-3** is highly positively correlated with this PC, and DNA**TO1-5** is also positively correlated. **TO1-7** and **TO1-8** are negatively correlated. **TO1-8** was chosen to bind to both MDA-MB-231 and DU-145, but as seen on PC1, the response of **TO1-8** significantly contributes to the classification of DU-145 but not MDA-MB-231. This PC also helps to separate the cell lines that are not well classified along the first two PCs; both SK-MEL-28 and U-87 MGVIII have positive scores on PC3, while HCT-15 and DU-145 have negative scores. On this axis, the DNA-peptides and peptides alone generally display the same direction but different magnitudes, with the peptides alone having larger magnitudes.

Despite an increase in scatter, groupings are still observed on PC4 and PC5. DNA**TO1-6**, DNA**TO1-8**, and **TO1-9** are the most positively associated variables with PC4, while **TO1-7**, DNA**TO1-7**, and DNA**TO1-3** are the most negatively associated. On average, SK-MEL-28 has the most positive score, while SK-OV-3 has the most negative score. These cell lines are more closely associated on the first three PCs; the variance described by PC4 is able to better separate them. U-87 MGVIII, DU-145, and MDA-MB-231 also score positively on this PC, and **TO1-8** contains the motif designed to bind the latter two cell lines. Here the response of this peptide contributes to their grouping, unlike on previous axes. The **TO1-3** peptide was designed to bind SK-OV-3 and HCT-15, and these two lines score negatively on this PC. It appears that on axes which describe smaller, yet still significant, amounts of the variance, the hypothesized responses of the cell-targeting peptides contribute to cell classification.

The variables containing **TO1-2**, **TO1-4**, **TO1-5**, and **TO1-6** are highly associated with the PCA classification of SK-MEL-28. **TO1-2** was included as a peptide expected to bind to this cell line. SK-MEL-28 expresses αvβ3, α2β1, and α2β1 integrins that should bind to peptides **TO1-4** and **TO1-5**.[50] The response of mannose mimic **TO1-6** could indicate mannose receptors present on this cell line or other surface features that are capable of interactions with the most hydrophobic peptide in the array.

The classification of DU-145 is highly associated with responses to peptides containing **TO1-8**, **TO1-6**, and to a lesser extent **TO1-4** and **TO1-9**. The motif in **TO1-8** was included to bind to DU-145. This cell line expresses α5β1 integrin and binds to fibronectin, explaining its response to **TO1-4**.[51] ErbB receptor overexpression, especially erbB-2, is highly associated with negative clinical outcome in prostate cancers, like the metastatic, androgen resistant DU-145 that expresses high levels of erbB1-3.[52] Again the response of mannose mimic **TO1-6** contributes to the classification of this cell line, indicating the presence of receptors capable of binding this peptide.

Although no peptides were selected specifically to bind U-87 MG VIII, several peptides are correlated with this cell line's classification, indicating the promiscuity of the array receptors. In particular, **TO1-9** is correlated with classification on PC1, PC2, and PC4, and this peptide binds erbB receptors of the type (αvβ3) that are overexpressed on gliomas. Peptide **TO1-8**, chosen to bind MDA-MB-231, is also correlated with classification of this cell line. Both cell lines are highly metastatic and overexpress erbB, which indicates that the peptide may be targeting erbB or other malignancy-associated receptors on these cell lines.

The classification of A-549 by PCA is correlated with the response of many of the array peptides. On PC1-PC3, the axes with the most variance, **TO1-2**, **TO1-4**, **TO1-5**, **TO1-6**, and **TO1-9** are correlated with classification. As noted in section 3.2.4, A-549

binds to laminin and expresses erbB receptors, explaining the correlation of classification with the loadings of **TO1-4** and **TO1-5**.[46,47] As observed in initial experiments, **TO1-1** and **TO1-3** are correlated with A-549 classification. The peptide chosen to bind A-549, **TO1-2**, also contributes to classification on axes that carry less variance (PC4 and PC5).

A variety of peptides are correlated with the classification of MOLT-4. These include **TO1-1**, **TO1-3**, **TO1-5**, **TO1-7**, and **TO1-9**. **TO1-7** is the peptide chosen to bind MOLT-4. The response of the other peptides to classification indicates that MOLT-4 expresses hyaluronan (**TO1-1**),[53] collagen- and/or laminin-binding integrins (**TO1-5**),[54] surface features similar to SK-OV-3 and HCT-15 (**TO1-3**), or other surface features that are capable of recognizing these sequences.

PCs that classify 786-O have high correlation with **TO1-9**. ErbB receptors are crucial to renal cell development and homeostasis.[55] Additionally, for classification of this cell line, **TO1-3**, **TO1-4**, and **TO1-5** are correlated on PCs that describe higher variance. The latter two peptides likely bind to cell surface integrins, whose assembly is modified in 786-O cells.[48] **TO1-7** and **TO1-8** are correlated with PCs describing lower variance. **TO1-7** was selected to bind 786-O, while **TO1-8** was selected for binding other cell lines.

TO1-2, **TO1-3**, **TO1-5**, **TO1-6**, **TO1-8**, and **TO1-9**, contribute to classification of SK-OV-3. The response of these peptides to classification of this cell line is in agreement with the initial three cell line experiment. **TO1-3** was selected to bind SK-OV-3, **TO1-9** binds erbB-2 that is overexpressed by SK-OV-3, **TO1-5** binds integrin receptors that are expressed to a greater extent on this cell line compared to other ovarian cancers.[38]

The variables that contribute to the classification of MDA-MB-231 are also in agreement with initial experiments. **TO1-8**, the peptide selected to bind MDA-MB-231, is correlated with the classification of this cell on three out of five PCs. As expected, the

integrin binding peptides **TO1-4** and **TO1-5** respond to this cell line, which overexpresses erbB. The rather promiscuous peptide **TO1-3** also contributes to classification of this cell line. This cell line also has a strong response to **TO1-1**, indicating it may express hyaluronan; hyaluronan synthase is associated with breast cancer cell invasion.[56]

Classification of the nine cancer cell lines tested was achieved using the differential array composed of nine peptide conjugates with and without DNA. Classification was observed over five principal components, which described 74% of the variance in the data set. Several peptide receptors were associated with each of the principal components, and the major contributors to each component were variable. Peptide sequences that had been selected to bind particular cell lines and known cell surface receptors generally responded as hypothesized. These sequences also responded to other cell lines that they were not hypothesized to bind, indicating that the receptors behave in a cross-reactive manner.

3.2.5.3 Reproducibility

Eight cell lines were grown from frozen stocks and again passaged to poly-D-lysine plates. The response of each cell line to the 9-peptide array was measured and analyzed using PCA. These results can be compared to PCA performed on the original data set, omitting the ninth cell line, MDA-MB-231. Comparison of the loading plots shows that both models capture similar amounts of variance on the first two PCs (Figure 3.35) Although the variables have different magnitudes, many of the variables that are positively associated with PC1 in the original experiment are also positively associated with PC1 in the reproduction. These are DNA**TO1-6**, **TO1-7**, **TO1-8**, DNA**TO1-8**, and DNA**TO1-9**. This is also true for many variables that are negatively associated with PC1:

TO1-1, DNA**TO1-1**, **TO1-3**, DNA**TO1-3**, and **TO1-9**. Similar congruencies are observed on PC2; in both experiments, **TO1-4**, **TO1-6**, **TO1-8**, and DNA**TO1-5** are positively associated, while DNA**TO1-1**, **TO1-5**, DNA**TO1-8**, and DNA**TO1-9** are negatively associated. It is clear that the models generated from the two data sets are not exactly the same, as the responses of some of the peptides were found to have different contributions to the variance in the two models generated.

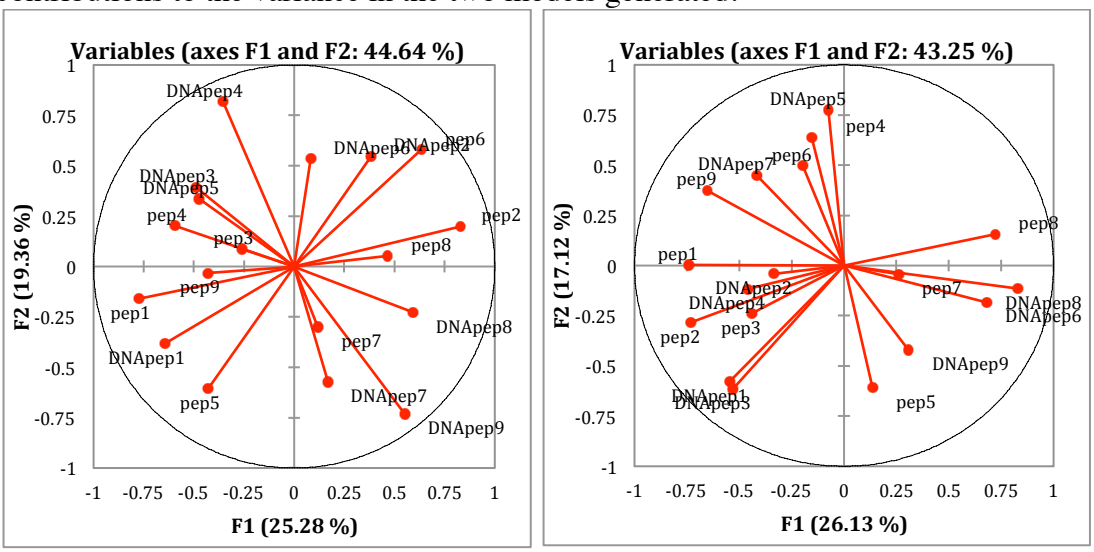


Figure 3.35 Comparison of PCA loading plots of the original experiment (left) and experiment reproduction (right) for eight cell lines.

The score plots of the first two PCs for both experiments are shown in Figure 3.36. It is clear that in both experiments cells can be classified according to cell line. There are major similarities between the reproduction and the original experiment in terms of placement on the score plot; this is likely due to the similarities of variable contributions in the model generated for each data set. This is an indication that, to some extent, the array is responding similarly each time it is exposed to an analyte. For example, in both experiment sets, MOLT-4 scores negatively on the first and second PCs.

Additionally, HCT-15, U87MGVIII, and DU-145 are closely associated in both score plots.

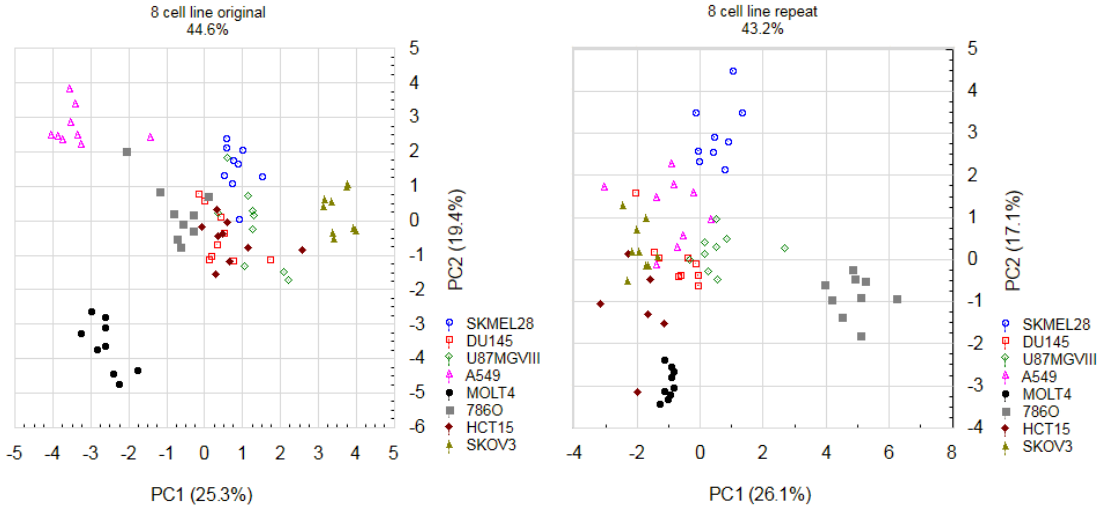


Figure 3.36 Comparison of two dimensional PCA score plots of the original experiment (left) and experiment reproduction (right) for eight cell lines.

As is apparent from the differences in the loading plots (Figure 3.35), there are observable differences in the scores of analytes between the two experiments (Figure 3.36). On the first PC, 786-O and SK-OV-3 have switched responses. Originally, SK-OV-3 had a high positive score and 786-O had a small negative score, while in the reproduction SK-OV-3 has a medium negative score and 786-O has a high positive score. Although both SK-MEL-28 and A-549 both score positively on PC2, the magnitude of their responses is switched between the two experiments

The score plot with a third PC is shown in Figure 3.37. Differences in the scores along PC3 are observed between the two experiments. However, it is shown in Table 3.6 that the sign of the majority of the variables is switched on this axis. The manner in which variables are chosen to positively or negatively correlate with a PC depend on the program algorithms used and are essentially arbitrary. Therefore, it is possible that cell

lines that had a high positive score on PC3 in the original experiment would have a high negative score on PC3 in the reproduction, due to the opposite signs of the correlations of a majority of the variables on this axis. For example, both MOLT-4 and SK-MEL-28 have negative scores on PC3 in the reproduction, but positive scores in the original experiment. DU-145 has a positive score in the reproduction, but a large negative score in the original experiment. The ordering of cell lines along this axis differs between experiments, likely due to the differences in magnitudes of the contribution of each variable to PC3 for the two data sets. Nevertheless, the third PC carries a significant amount of the variance in the data set (>10%) and contributes to classification of the cell lines in both experiments.

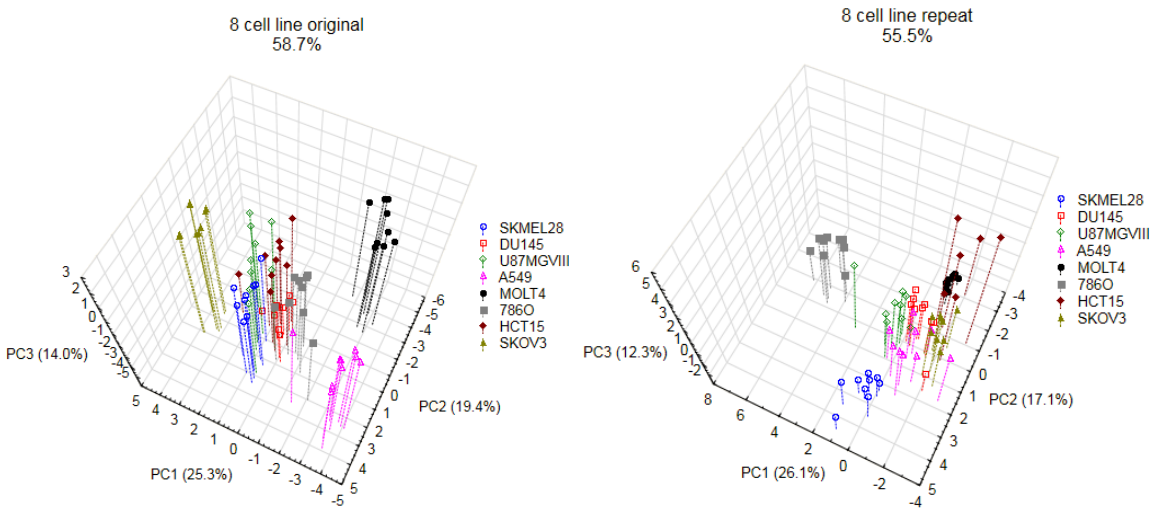


Figure 3.37 Comparison of three-dimensional PCA score plots of the original experiment (left) and experiment reproduction (right) for eight cell lines.

	Original	Repeat
DNApep1	-0.374	0.374409
DNApep2	-0.125	0.655486
DNApep3	-0.222	-0.19253
DNApep4	0.268	-0.53816
DNApep5	0.575	-0.18678
DNApep6	0.247	0.175707
DNApep7	0.069	0.353574
DNApep8	-0.034	0.132327
DNApep9	-0.117	0.326137
pep1	-0.094	0.439931
pep2	0.260	0.324185
pep3	0.799	-0.53936
pep4	0.468	0.100267
pep5	-0.237	-0.35808
pep6	0.313	-0.35334
pep7	-0.604	0.185345
pep8	-0.512	0.317246
pep9	-0.489	0.123915

Table 3.6 Correlations between variables and PC3 in the original and reproduction data sets. Green values are inverted between the two experiments.

3.2.6 Fluorescence and MS Patterning of Cancer Cell Lines Using a Peptide Mixture

3.2.6.1 Titrations of Peptides and Mixture with 2 kb dsDNA

In order to assemble more intercalator peptides along the DNA backbone, a longer dsDNA strand consisting of about 2.5 kilobases was used. Titrations with each peptide were performed to determine the equivalents required for fluorescence saturation. As observed with the 1 kb strand, variable equivalents were needed to reach saturation of the fluorescence signal and different peptide sequences resulted in different emission intensities. A mixture of the nine peptides at the individual saturation concentrations was prepared and titrated with the 2.5 kb strand (Figure 3.38). Maximum fluorescence was seen at about two peptides per DNA base pair, which corresponded to a 1/10

concentration of the nine peptide mixture. For further experiments, such a dilution of the peptide mixture was used.

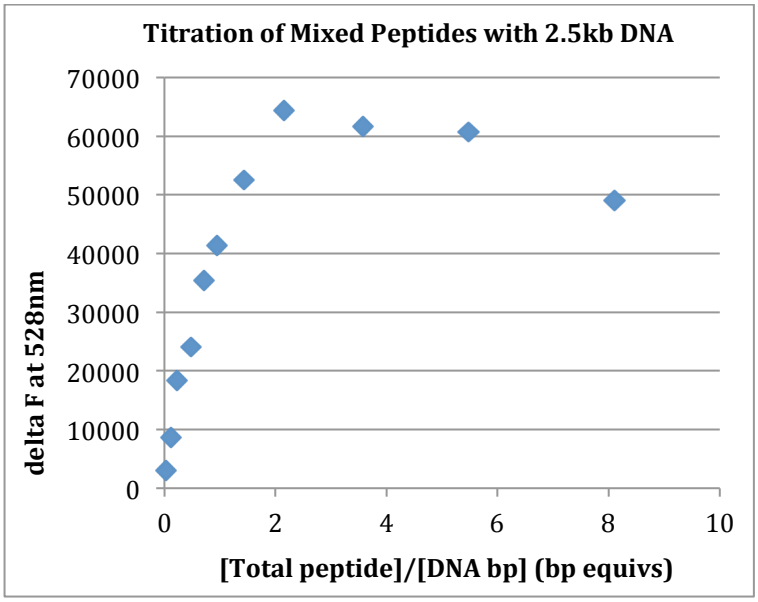


Figure 3.38 Titration of 2.5 kb dsDNA with a nine-peptide mixture at individual peptide concentrations corresponding to fluorescence saturation upon intercalation with DNA.

3.2.6.2 Fluorescence Classification

Cancer cells of different tissue type were grown to confluency on 12-well plates. The peptide mixture was exposed at the same concentration with and without 2.5 kb dsDNA to plates of each cell type and the fluorescence response of 10 replicates was measured. Fluorescence values of the mixture exposed to cells were expressed as ratios over the solution fluorescence. The plot of the fluorescence values with DNA versus those without DNA shows that the mixtures are able to classify the cell line replicates (Figure 3.39). Differences in the fluorescence values for the cell lines arise from either differential binding of sequences, binding of sequences with different emission intensities, or both. The different fluorescence properties of the peptides therefore lead to the differential fluorescence responses that are responsible for the cell line classification.

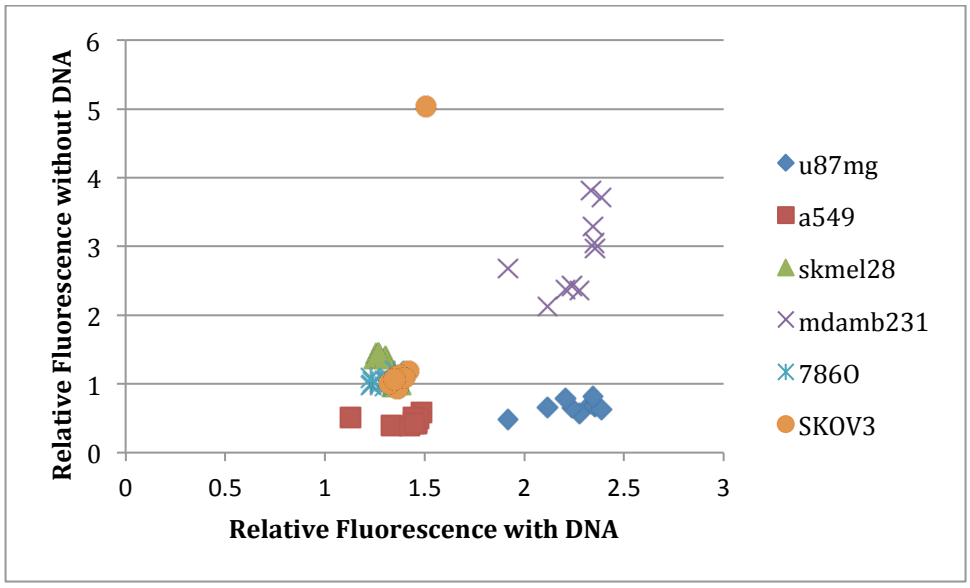


Figure 3.39 Fluorescence responses of peptide mixtures with and without DNA to cells of different tissue origins.

3.2.6.3 Mass Spectrometric Analysis

In order to identify the sequences of the peptides in the mixture that bound to each cell line, cell lines were washed with DPBS, detached, and bound sequences were eluted into water. The supernatant of the elution solution was analyzed using ESI mass spectrometry. Although fluorescence values were measured on concentrated supernatant, masses corresponding to the peptides could not be identified in these solutions. Elimination of the DPBS wash did allow for identification of one peptide bound to DU-145, **TO1-6** (Figure 3.40). This peptide is the most hydrophobic of those in the mixture. However, it was found to contribute to the classification of DU-145 in the nine-peptide array. Further experiments will be conducted to analyze other cell lines under similar low-wash conditions.

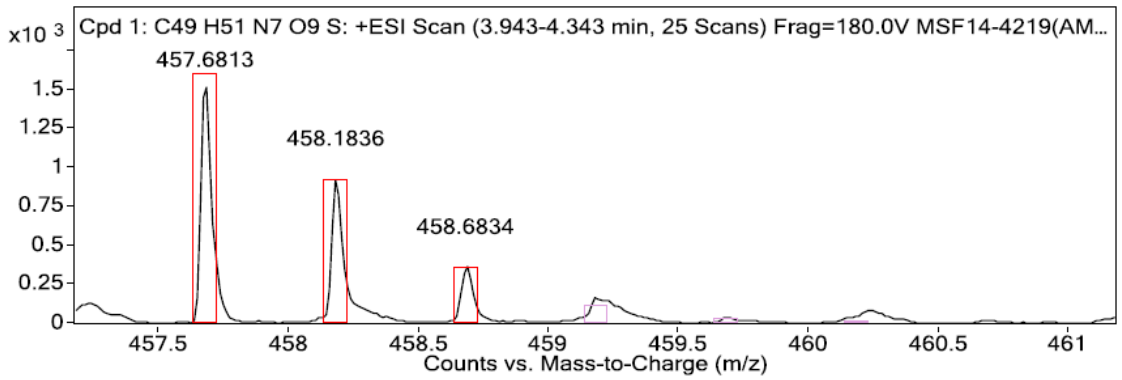


Figure 3.40 Mass spectrum of peptide indentified as present in DU-145 cell supernatant. Expected mass is 457.68070 for $(M+2H)^{2+}$.

3.3 CONCLUSION

A dynamic, DNA-templated array of differential intercalator-peptide receptors was created to pattern multivalent biological analytes. Derivatives of thiazole orange, a fluorescent DNA intercalator, were prepared, and their interactions with dsDNA were studied. Although it was hypothesized that supercoiled DNA would be able to accommodate a larger amount of thiazole orange derivatives, it was concluded that long, linear dsDNA of 1000 bp was the best scaffold for assembling peptide recognition units. Cancer cells of different tissue origin were chosen as exemplar multivalent biological analytes due to their surface complexity and the strength of differential array sensing with complex mixtures. Instead of choosing recognition units of simple functionality, a variety of peptide conjugates of thiazole orange were synthesized. These peptides were chosen to target analytes by incorporating sequences that bind to known cell surface features and sequences that were previously shown to have high association with particular cell lines. These receptors were hypothesized to behave in a cross-reactive manner and bind to many of the cell lines to varying extents because each peptide sequence targets shared

surface features or multiple cell lines. The array was hypothesized to produce a unique pattern for each cell line resulting from a differential response from many peptides binding to each cell line but to different extents. Using chemometric techniques, unique patterns of receptor responses can be used for analyte classification.

Receptor assemblies of dsDNA and nine peptides of variable sequence were prepared and incubated with nine cancer cell lines. The nine peptides were simultaneously exposed to the cancer cell analytes without DNA. The receptor assembly containing DNA displayed the ability to detach the cells from the growth surface, which indicated that the receptors were targeting cell surface features that contribute to adhesion. Detachment of cells occurred to a much greater extent when the peptides were exposed as the DNA-assembled receptors, compared to the free peptides. This indicated that the receptor assembly was better at binding to the cell surface, likely due to multivalent interactions, and therefore behaved as hypothesized.

Fluorescence measurements of the response of the peptide array, with and without DNA, to each cell line were input into chemometric routines for analysis. Both linear discriminant analysis and principal component analysis were able to successfully classify nine replicates of each cell line on score plots by using the fluorescence pattern of the array. Such classification indicated that the fluorescence patterns for each cell line were unique, as hypothesized for a differential array. Analysis using PCA indicated that the classification arose from the data structure, and that the variance in the data set was diagnostic for cell type.

Analysis of the contributions of each peptide to the variance captured and to cell line classification showed that many of the peptides responded as designed. The variable sequences bound to cells that express epitopes for which the peptides were designed to bind and/or showed binding to cell lines for which they were hypothesized to have

affinity. The peptides behaved promiscuously, as each peptide was capable of responding to multiple cell lines. Not all peptides bound to each cell line, but the pattern of receptor response was diagnostic of the cell type. These combined responses allowed for discrimination of cell lines using the differential array.

Investigation into reproducibility of the array classification showed that cell lines were again successfully classified. The array response, as analyzed using PCA, was found to not be exactly reproduced. Many of the peptides responded in a similar manner and had congruent contributions to the variance, as shown in their correlations with the principal components. The scores of cell lines on the score plot, however, were significantly different for some cell lines. This may be due to the difference in the contribution of some variables to the PCs, or due to significant differences in the response of the array to cell lines between the two experiments. Even though the cell lines are successfully classified, the pattern of response for each cell line is subtly different, and any small changes can lead to differences in variable loading and analyte scores, resulting in differences in the appearance of PCA plots. Additionally, the array may be capable of responding to subtle changes in expression of cell surface receptors on the analytes that are associated with growth for different experiments. In such a case, the response may not be exactly the same as a previous pattern, but the surface features would be unique to each cell line and the array response to each analyte would therefore still be unique.

3.4 EXPERIMENTAL

3.4.1 Synthesis

Chemicals were purchased from Sigma-Aldrich, Fisher Scientific, or Acros Organics and used without further purification. All solvents were of at least reagent grade and purchased from Fisher Scientific. NMR spectra were collected on a Varian 400MHz NMR using isotope-enriched solvents from Cambridge Isotopes. Mass spectrometry was performed in water/methanol or water/acetonitrile mixtures on an Agilent instrument.

3.4.1.1 Synthesis of a Carboxylic Acid Thiazole Orange Derivative

*Synthesis of 3-methyl-2-(methylthio)-benzothiazolium tosylate (**3.3**)*

3.7 g (20 mmol) of 2-methylthiobenzothiazole and 4.5 g (24 mmol) of methyl *p*-toluene sulfonate were added to a dry 100 mL round bottom flask and heated under Ar gas to 130°C. After 1 hour, the solution turned yellow and the temperature was reduced to 70°C. Acetone was added to the mixture through the condenser until a yellow solid appeared, and the solid was heated for an additional half-hour then cooled to RT. White crystals were isolated by vacuum filtration and washed with acetone, then dried under high vacuum overnight, yielding 5.71 g of **3.3** (76% yield).

^{1}H-NMR (400 MHz, CD_3OD): δ 2.34 (s, 3H), 3.11 (s, 3H), 4.13 (s, 3H), 7.18 (d, *J*=7.9, 2H), 7.66 (d, *J*=8.2, 2H), 7.71 (td, *J*=1.0, 8.2, 1H), 7.83 (td, *J*=1.2, 8.5, 1H), 8.06 (d, *J*=8.5, 1H), 8.21 (d, *J*=8.2, 1H).

^{13}C-NMR (100 MHz, CD_3OD): δ 18.64, 21.44, 37.00, 116.60, 124.77, 127.03 (2C), 128.63, 129.89, 130.02 (2C), 130.85, 141.68, 143.86, 144.21, 183.35.

*Synthesis of 1-(3-carboxypropyl)-4-methylquinolin-1-ium chloride (**3.8c**)*

1.4 g of lepidine (9.8 mmol) and 2.2 g of 4-bromobutyrate (11.2 mmol) were added to a dry 5 mL microwave tube equipped with a stirbar and microwaved for 40

minutes at 82°C. The microwave tube was rinsed with MeOH, and the resulting purple liquid was poured into 100 mL of Et_2O and stirred until the ether became clear, then the ether was decanted. Upon combining the purple liquid with 50 mL of H_2O in a 100 mL rbf, the solution turned burnt orange. Then 2 mL of 48% HBr was added to the water and the solution was refluxed overnight. The solution volume was reduced by evaporation, and a solution of saturated KPF_6 was added by pipet until the remaining liquid became cloudy. After precipitation in the fridge overnight, an off-white solid, **3.8b**, was obtained from vacuum filtration. The product was dissolved in 3:1 MeCN:H_2O and passed down an Amberlite® 402 Cl ion exchange column and lyophilized. A white fluffy solid was isolated but turned into a taupe hard solid upon exposure to moisture. Yield of **3.8c** was 0.171g (7%).

3.8b: ^{1}H-NMR (400MHz, CD_3CN): δ 2.27 (m, 2H), 2.53 (t, 2H), 3.00 (d, 3H), 4.92 (m, 2H), 7.85 (d, 1H), 8.02 (m, 1H), 8.48 (m, 2H), 8.90 (d, 2H).

^{19}F-NMR (400 MHz, CD_3CN): δ -74.58, -72.02.

^{13}C-NMR (100 MHz, CD_3CN): δ 20.85, 25.89, 30.84, 58.18, 120.28, 124.02, 128.50, 131.06, 131.12, 136.79, 138.78, 149.00, 161.16, 174.50.

MP: 100-105°C dec.

3.8c: 1H-NMR (400MHz, CD3CN): δ 2.57 (t, 3H), 2.99 (s, 3H), 4.95 (2H), 7.86 (d, 1H), 7.99 (t, 1H), 8.21 (t, 1H), 8.47 (d, 1H), 8.54 (d, 1H), 8.98 (d, 1H).
MP: 196-198°C dec.

*Synthesis of **TO1***

The thiazolium tosylate salt **3.3** (0.735 g, 2.0 mmol) and quinolinium chloride **3.8c** (0.534 g, 2.0 mmol) were both weighed into a dry 100 mL rbf. After adding dry CH_3CN (40 mL), the flask was equipped with a condenser, and the mixture heated to 82°C. 600 uL (0.44 g, 4.3 mmol) Et_3N was injected into the mixture, causing the orange

solution to immediately change to a dark red. The solution was heated under Ar for three hours; after forty-five minutes, orange solid was present in the flask, and it emitted a bad odor (methyl sulfide). The flask was cooled and 100 mL Et_2O added into the suspension. The solid was vacuum filtered and rinsed with Et_2O and then recrystallized from MeOH/H_2O. The orangey red solid weighed 0.442g (54% yield).

^{1}H-NMR (400 MHz, DMSO-d_6): δ 1.07 (t, *J*=7.0, 1H), 2.11-2.01 (m, 2H), 2.41 (t, *J*=7.1, 2H), 3.37 (dd, *J*=7.0, 14.0, 3H), 3.98 (s, 3H), 4.55 (t, *J*=7.6, 2H), 6.88 (s, 1H), 7.34 (d, *J*=7.2, 1H), 7.40 (t, *J*=7.6, 1H), 7.59 (t, *J*=7.8, 1H), 7.73 (t, *J*=7.1, 2H), 7.98 (dd, *J*=7.8, 17.2, 2H), 8.16 (d, *J*=8.6, 1H), 8.55 (d, *J*=7.2, 1H), 8.75 (d, *J*=8.2, 1H).

^{13}C-NMR (150 MHz, DMSO-d_6): δ 24.11, 30.30, 33.70, 53.41, 88.03, 107.87, 112.88, 117.83, 122.74, 123.77, 124.11, 124.45, 125.68, 126.74, 128.11, 133.20, 136.96, 140.31, 144.17, 148.50, 159.98, 173.69.

HRMS-ESI (*m/z*): $[M]^+$ calcd for $(C_{22}H_{21}N_2O_2S)^+$ 377.1324; found: 377.1324.

MP: 237-239°C dec.

3.4.1.2 Peptide Coupling

Synthesis of all peptides was undertaken using an automated peptide synthesizer (Protein Technologies) using Fmoc chemistries and Wang resins. All deprotection and coupling steps were repeated twice. Resins and protected amino acids were obtained from NovaBioChem. HPLC-grade solvents were purchased from Sigma-Aldrich and prepared by addition of 0.1% TFA (v/v), filtration through a 0.2 micron filter, and then degassed using sonication. HPLC purification was performed on a Shimadzu with a preparative C-18 column using water and acetonitrile as the mobile phase.

General Procedure

For a 200 umol scale reaction, an Fmoc-protected Wang resin (100-200 mesh) coupled to the C-terminal residue was swelled in DMF. The resin was deprotected with a 20% piperidine solution then washed sequentially with DMF and DCM. A DMF solution of the next peptide in the sequence (100 mM), with appropriate protecting groups, was added to the reaction vessel for coupling between the amine and carboxylic acid. The coupling reagent used was PyBOP (300 mM) with DIPEA (1.2 M) in DMF. After coupling to the final residue in the peptide sequence, the resin was washed with DMF and DCM and removed from the synthesizer. The resin was manually deprotected using 20% piperidine until a positive Kaiser test resulted. At this point, a solution of **TO1** in HOBT/HBTU (300 mM each) with1.2 M DIPEA in DMF was added, and the resin shaken until a negative Kaiser test resulted. The resin turned blood red and retained a red stain after washing with 20 mL DMF, 20 mL DCM, and 20 mL MeOH. The resin was rinsed with 15 mL acetic acid and vacuum dried overnight. Peptides were cleaved from the resin using a TFA cleavage cocktail (95% TFA, 2.5% TIPS, 1.5% EDT, 1% H_2O) and shaking for 4 hours. The TFA solution was separated from the resin by filtration under aspirator pressure, the resin was washed with TFA, and the filtrate collected. TFA was removed from the filtrate on a rotavap. Precipitation of the peptides using cold ether resulted in an orange solid.

HPLC Purification and Isolation

Solid peptides were taken up in a mixture of 50/50 DMSO/H_2O and purified using preparative HPLC. The method was as follows:

0-35 min 5-30% MeCN in H2O, 35-45 min 30% MeCN, 45-60 90% MeCN. The absorbance at 500 nm was monitored and only fractions containing this absorbance were analyzed.

Fractions containing the desired peptide were verified using LC/MS then combined, rotavaped to remove acetonitrile, and lyophilized to remove water. Low resolution mass spectrometry was used for each peptide (Table 3.7)

TO1-peptide	Expected Mass (m/z)			Found (m/z)
	$(M+2H)^{3+}$	$(M+H)^{2+}$	(M^{+})	
TO1-1	315.5	472.7	944.5	315.6, 472.6, 944.1
TO1-2	282.8	423.7	846.4	282.9, 423.7, 846.3
TO1-3	292.8	438.7	876.4	292.9, 438.7, 876.3
TO1-4	310.8	465.7	930.4	310.9, 465.7, 930.3
TO1-5	292.8,	438.7	876.4	292.9, 438.4, 876.3
TO1-6	305.4	457.7	914.3	457.6, 914.0
TO1-7	302.6	453.7	906.4	302.8, 453.6, 907.1
TO1-8	273.1	409.2	817.4	409.2
TO1-9	298.1	446.7	892.4	298.2, 446.6, 892.0

Table 3.7 FIA-ESI + MS (low-res) on HPLC purified peptides

3.4.2 DNA Isolation and Amplification

3.4.2.1 General Procedures

All oligonucleotides were ordered from Integrated DNA Technologies (IDT), were dissolved in nuclease-free water to a concentration of 100 μM, and stored at -20C. Polymerase chain reaction (PCR) was performed under one of two following conditions. Using Taq Polymerase from NEB: 2.5U enzyme per 50ß uL reaction, 1X ThermoPol buffer, 0.2 mM dNTPs, 4 μM primers, 0.1 ng/uL template or 10 ng/uL template

(genomic) were added. Thermal cycling steps were 1) 98 °C 2 min, 2) 98C - 30 sec, 3) Annealing T - 30 sec, 4) 68 °C 1 min/kb, 5) hold at 4 °C, and steps 2-4 were repeated for as many cycles as necessary to yield optimal amplification. Using Q5 high fidelity polymerase from NEB 1X buffer, 0.2 mM dNTPs, 4 μM primers, 1.25 U/25 uL polymerase, and 0.1 ng/uL template were added to 50 or 100 uL reactions. Thermal cycling steps were 1) 98 °C – 30 sec, 2) 98 °C – 10 sec, 3) Annealing T – 30 sec, 4) 72 °C – 25 sec/kb, 5) 72 °C – 2 min, 6) 4 °C – hold, with steps 2-4 repeated as needed.

Concentrations of dsDNA were estimated from absorbance values measured at 260 nm on a nanodrop. Sequences of dsDNA were input into the spreadsheet provided from Tautarov *et al*. in order to estimate the sequence and length-dependent molar absorptivity of each strand at 260 nm.[57] The concentrations of each dsDNA solution were calculated using the absorbance values at 260 nm and the estimated molar absorptivity.

Gel electrophoresis was performed using agarose gels of variable concentration (3-4% for less than 100 bp, 1-2% for 1000 bp or higher) in 1x TAE buffer. Samples were loaded with glycerol loading dyes and run against a Fermentas GeneRuler DNA Ladder Mix or O'RangeRuler 10 bp DNA ladder (<150 bps) (Figure 3.41). Ethidium bromide was added to the gels upon preparation in order to stain the DNA. Gels were run at 15V/cm until appropriate loading dyes ran at least half the length of the gel, and then gels were imaged with UV illumination.

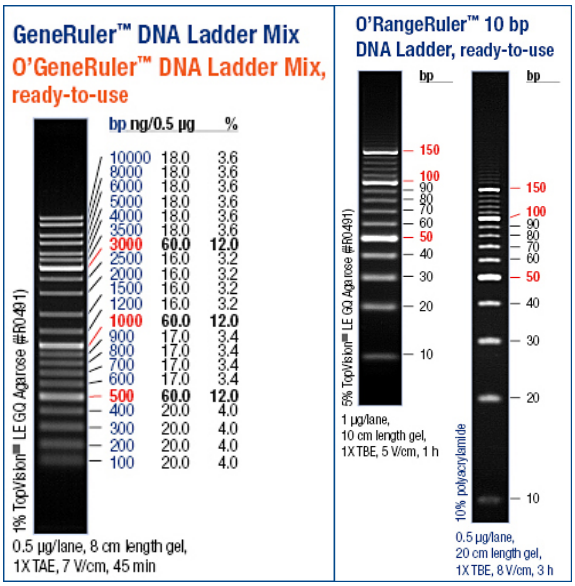


Figure 3.41 DNA ladders used to size PCR products on agarose gels.

3.4.2.2 dsDNA strands

53-mer

An oligonucleotide consisting of 53 base pairs of the sequence:

5'-

TGCTGCAGATGCAACGTCATGTATTATCAGCAGTAACTAGCGACTGGGTCCG-

3'

was designed to have close to 50% GC content and a linear structure. Primers specific to this sequence were also designed using Primer-3. After a temperature gradient and cycle course, the optimal annealing temperature was found to be °C with X cycles of amplification. The absorbance value of purified PCR product at 260nm (A_{260}) of the DNA suspended in DPBS was measured by nanodrop and the concentration calculated from ε = 84,2003 $M^{-1}cm^{-1}$.

152- and 202-mers

These strands of DNA are the double-stranded templates for aptamers, and were donated by Sara Stewart.

1000-mer

Genomic DNA was isolated from buccal cells collected after rinsing the mouth with 0.1 M saline for 2 minutes. The buccal cells were lysed and extracted following the procedures of the Sigma-Aldrich GenElute™ miniprep kit. Isolated genomic DNA was stored at -20°C. Gel electrophoresis showed that the DNA had not been sheared to smaller than 15,000 bps (Figure 3.42). Concentration of the genomic DNA was measured as 58.4 ng/uL on a nanodrop.

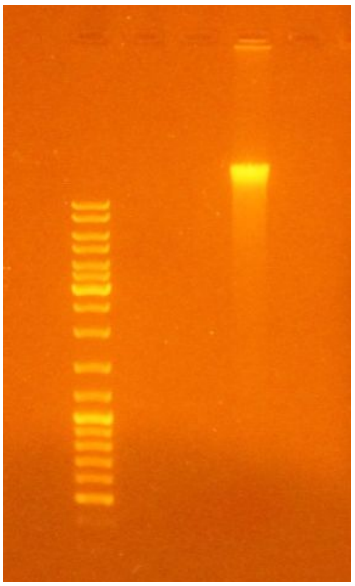


Figure 3.42 Genomic DNA isolated from buccal cells run against GeneRuler ladder on 1% agarose in 1X TAE.

A search through the NIH genome database Nucleotide identified a highly conserved genomic sequence of 2500 bps that codes for the β2 adrenergic receptor. Primers for a 1000 bp product from the gene were developed using Primer-3 and

crosschecked in the Primer-BLAST program from NIH. Primer-BLAST showed that the only product that the primers would give from a genomic template was the intended 1000 bp sequence:

5’CACCAACTACTTCATCACTTCACTGGCCTGTGCTGATCTGGTCATGGGCCTG
GCAGTGGTGCCCTTTGGGGCCGCCCATATTCTTATGAAAATGTGGACTTTTGG
CAACTTCTGGTGCGAGTTTTGGACTTCCATTGATGTGCTGTGCGTCACGGCCA
GCATTGAGACCCTGTGCGTGATCGCAGTGGATCGCTACTTTGCCATTACTTCA
CCTTTCAAGTACCAGAGCCTGCTGACCAAGAATAAGGCCCGGGTGATCATTC
TGATGGTGTGGATTGTGTCAGGCCTTACCTCCTTCTTGCCCATTCAGATGCAC
TGGTACCGGGCCACCCACCAGGAAGCCATCAACTGCTATGCCAATGAGACCT
GCTGTGACTTCTTCACGAACCAAGCCTATGCCATTGCCTCTTCCATCGTGTCC
TTCTACGTTCCCCTGGTGATCATGGTCTTCGTCTACTCCAGGGTCTTTCAGGA
GGCCAAAAGGCAGCTCCAGAAGATTGACAAATCTGAGGGCCGCTTCCATGTC
CAGAACCTTAGCCAGGTGGAGCAGGATGGGCGGACGGGGCATGGACTCCGC
AGATCTTCCAAGTTCTGCTTGAAGGAGCACAAAGCCCTCAAGACGTTAGGCA
TCATCATGGGCACTTTCACCCTCTGCTGGCTGCCCTTCTTCATCGTTAACATTG
TGCATGTGATCCAGGATAACCTCATCCGTAAGGAAGTTTACATCCTCCTAAAT
TGGATAGGCTATGTCAATTCTGGTTTCAATCCCCTTATCTACTGCCGGAGCCC
AGATTTCAGGATTGCCTTCCAGGAGCTTCTGTGCCTGCGCAGGTCTTCTTTGA
AGGCCTATGGGAATGGCTACTCCAGCAACGGCAACACAGGGGAGCAGAGTG
GATATCACGTGGAACAGGAGAAAGAAAATAAACTGCTGTGTGAAGACCTCC
CAGGCACGGAAGACTTTGTGGGCCATCAAGGTACTGTGCCTAGCGATAACAT
TGAT-3’

Primers were ordered from IDT and dissolved in nuclease-free water to a concentration of 100 μM. The conditions for PCR were optimized by running a

temperature gradient to find the ideal annealing temperature of the primers to the genomic sequence, 58.5°C. The product band at 1000 bp was gel-extracted and purified, and the number of cycles for this template optimized for use in further PCR reactions (Figure 3.43). For amplification with NEB Taq, the annealing temperature was 58.5 °C; for NEB Q5, the annealing temperature was the recommended 3 °C above the lower primer T_m. 56 °C. PCR products were purified using a QIAquick PCR purification kit, and the A_{260} of the resulting DNA measured by nanodrop and its concentration calculated from ε=15791142 M^{-1} cm^{-1}.

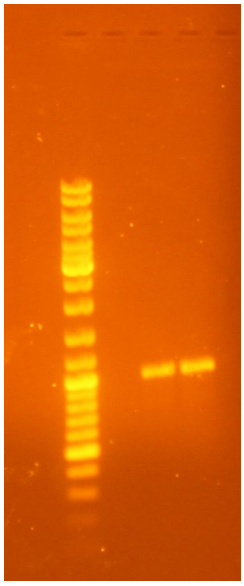


Figure 3.43 PCR amplified 1 kb dsDNA after optimization of annealing temperature and cycles. Samples were run on a 1.5% agarose gel in 1X TAE against GeneRuler ladder.

2.5 kb

The template for this strand was donated from the Ellington lab. It is a mutant of the pKD3 plasmid that consists of 2645 base pairs. This sequence was amplified using Q5 with an annealing temperature of 63 °C and 30 cycles.

3.4.3 Titrations

All fluorescence spectral values were measured on a PTI Fluorimeter with an 814 photomultiplier detection system using a 75W xenon short arc lamp at λ_{ex} = 498 nm and λ_{em}= 508-620 nm with 4 nm slit widths. Absorbance measurements were taken on a Beckmann Coulter DU 800 UV/Vis Spectrophotometer. Well plates were measured on a BioTek Synergy 2 Multimode Microwell-plate reader, with fluorescence filters as follows: excitation – 485/20, emission – 528/20. Well plates for the cell-patterning reproduction were measured on a BioTek Cytation3 Microwell-plate reader, using monochromators set for excitation at 485 nm and emission at 528 nm.

Titrations were undertaken in Dulbecco's phosphate buffered saline without calcium or magnesium (Gibco® Life Technologies). The formulation is 200 mg/L KCl, 200 mg/L KH_2PO_4, 8 g/L NaCl, and 2.16 g/L Na_2HPO_4 □$7H_2O$ which results in a 10 mM phosphate buffer with 150 mM NaCl at pH = 7.4.

3.4.3.1 Initial Titrations

A 1 mM stock solution of **TO1** in DMSO was prepared by weighing 0.0420 g in a 100 mL volumetric flask and diluting to volume with DMSO. This solution was diluted to 100 μM by adding 1 mL by syringe into a 10 mL volumetric half filled with DPBS and then filling to volume with DPBS.

*Titration of **TO1** into DNA*

A solution of **TO1** (typically 5 μM) was prepared by diluting the appropriate volume of a 1 mM stock solution in DMSO in a volumetric flask ¾ filled with DPBS and diluting to volume. The concentration of **TO1** was verified by UV/Vis spectrophotometry using A_{501} = 63,000 M^{-1} cm^{-1}.[14d] A solution of DNA (typically 250 nM) was prepared by adding a DNA stock solution in H_2O into a 1 mL volumetric ¾ filled with DPBS and then diluted to volume. 750 uL of the DNA solution was added into a fluorescence cuvette,

and aliquots of **TO1** were added by microsyringe. For the titration where the concentration of DNA was held constant, an aliquot of DNA stock was added to the **TO1** solution before diluting to volume, such that the concentration was the same as the titrand. The fluorescence was monitored with two averages taken for each titration point.

*Titration of DNA into **TO1***

A solution of **TO1** (typically 250 μM) was prepared by diluting the appropriate amount of a 1 mM stock in DMSO into a volumetric flask ¾ filled with DPBS and diluting to volume. A solution of DNA (typically 1 μM) was prepared by diluting the appropriate amount of a DNA stock in H_2O into a 1 mL volumetric ¾ filled with DPBS, then adding the appropriate volume of a 1 mM stock of **TO1** in DMSO such that its concentration was the same was the titrand. A fluorescence cuvette was filled with 750 uL of **TO1,** and aliquots of DNA containing **TO1** were added by microsyringe. The fluorescence was recorded with two averages taken for each titration point.

Continuous Variation Analysis

4.5 μM solutions of 53 bp DNA and **TO1** were prepared in DPBS by diluting the appropriate volumes of stock solutions in water and DMSO, respectively, into DPBS. Variable amounts of each solution were added into 10 glass vials, such that the total concentration of the two components was equal to 4.5 μM, but the mole fraction of each varied from 0-1. The solutions were diluted to 800uL with DPBS, giving a final total concentration of 700 nM, and transferred to a cuvette and the fluorescence monitored.

Solutions of 4 μM base pair concentration 1 kb DNA (5 nM DNA strand) and 5 μM **TO1** were prepared as above. These solutions were dispensed into a 384-well plate

by micropipette, such that the total volume was 100 uL, the total concentration of the two components held constant at 4 μM, and the mole fraction of each component varied from 0-1. The fluorescence was measured on the well-plate reader with the sensitivity referenced to the wells with the highest fluorescence after an initial reading at the default sensitivity.

ESI-Mass Spectrometry Titrations

Using 14 base pair duplex, d(GGCGTCGGCGTCGC/CCGCAGCCGCAGCG), time dependent intercalation studies were carried out using a 1:1 solution of DNA-to-dye at 5 μM in 30 mM ammonium acetate. Aliquots of the solution were taken every 0.5 hours over the course of two hours and a reaction time of 1 hour was determined to yield sufficient intercalation at detectable levels and was used for all subsequent experiments.

Concentration dependent studies were conducted using the same duplex by varying the molar ratio of duplex-to-dye for a fixed reaction time of 1 hour. The concentration of duplex was maintained at 5 μM in solution, while the dye concentration was varied from 5 – 25 μM.

3.4.3.2 Peptide-Intercalator Titrations

A stock solution of dsDNA was diluted to give a final concentration of 4 nM (1 kb strand) or 2 nM (2.5 kb strand) in appropriate wells of 96-well black, clear bottom plates (Costar® 3631) filled with DPBS. Stock solutions of peptides were added to each well at different volumes in order to provide different concentrations and thus equivalents. Fluorescence emission was measured and subtracted from DPBS blank with dsDNA in order to give a delta emission value for each well.

1 kb dsDNA

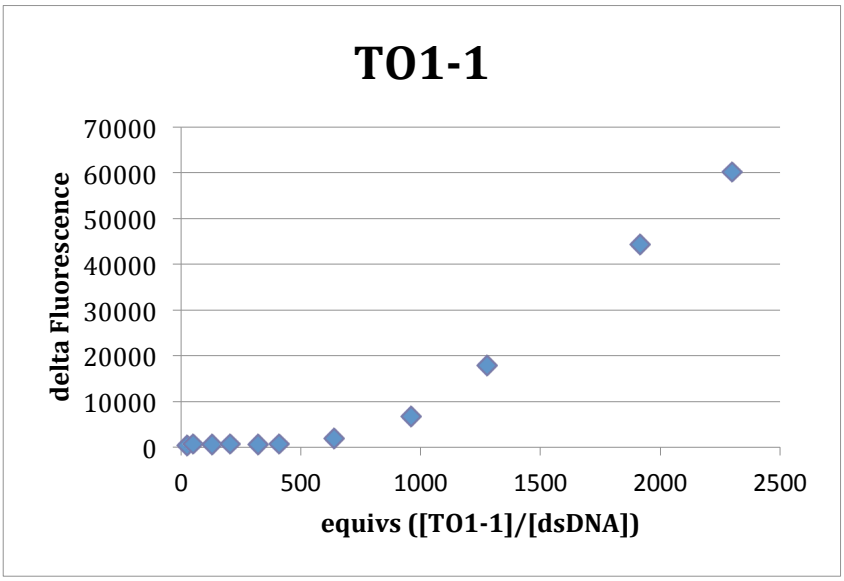


Figure 3.44 Fluorescence emission of titration of 1kb dsDNA (4 nM) with **TO1-1** in DPBS, pH = 7.4. Maximum fluorescence is seen beyond 2300 equivalents, or 9.2 μM, **TO1-1**.

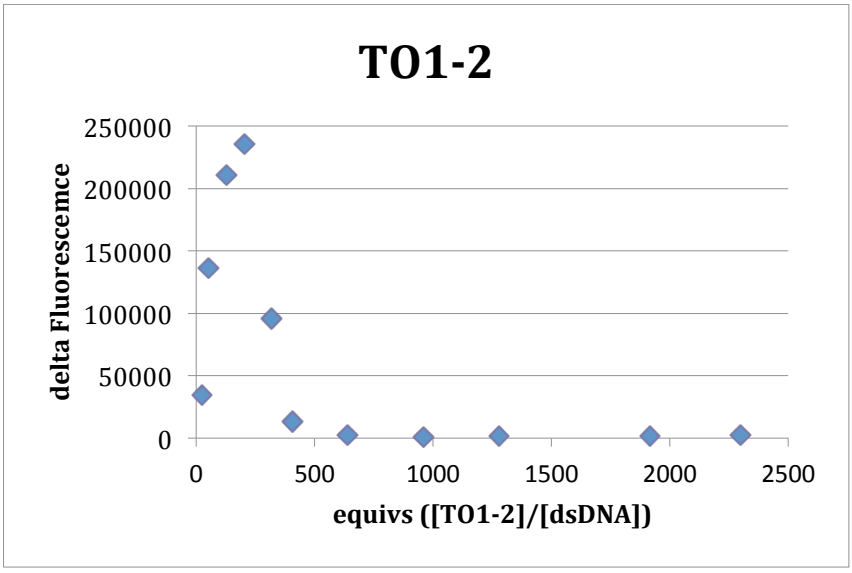


Figure 3.45 Fluorescence emission of titration of 1kb dsDNA (4 nM) with **TO1-2** in DPBS, pH = 7.4. Maximum fluorescence is seen at 200 equivalents, or 0.8 μM, **TO1-2**.

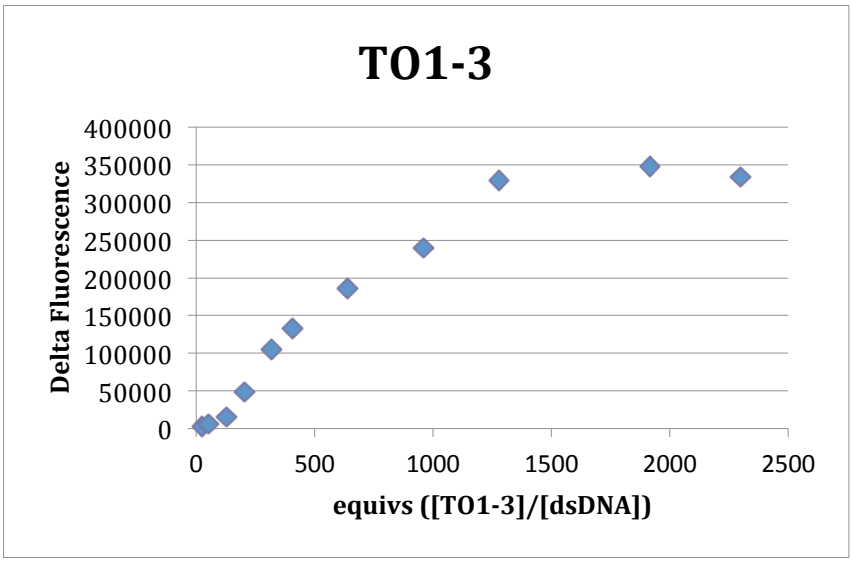


Figure 3.46 Fluorescence emission of titration of 1kb dsDNA (4 nM) with **TO1-3** in DPBS, pH = 7.4. Fluorescence saturation is seen at 1500 equivalents, or 6 μM, **TO1-3**.

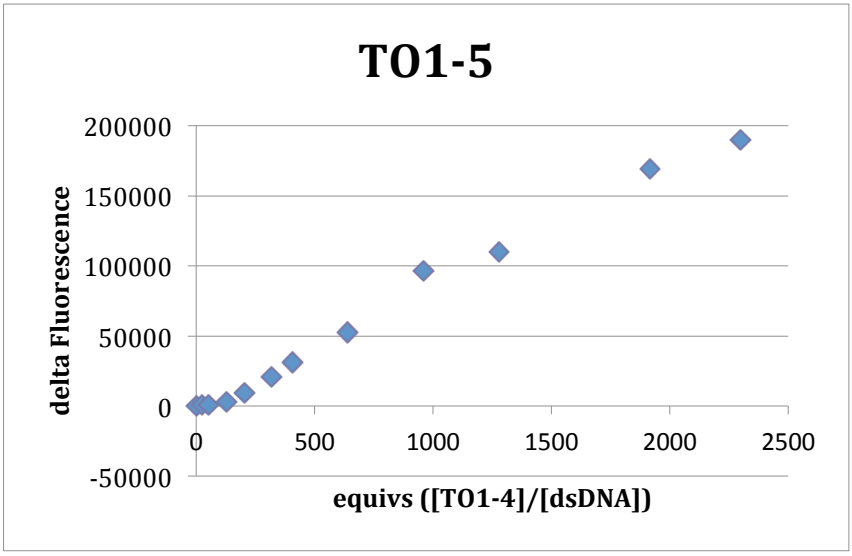


Figure 3.47 Fluorescence emission of titration of 1kb dsDNA (4 nM) with **TO1-5** in DPBS, pH = 7.4. Maximum fluorescence is seen beyond 2300 equivalents, or 9.2 μM, **TO1-5**.

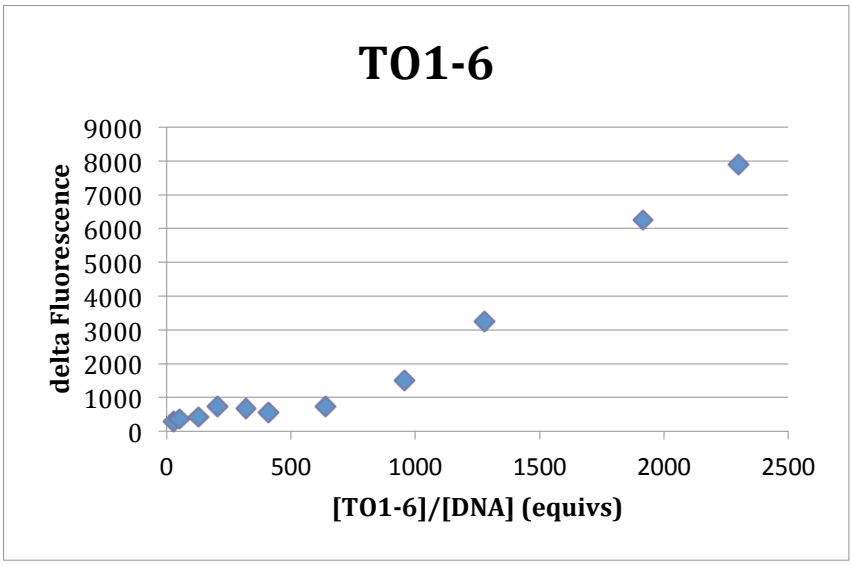


Figure 3.48 Fluorescence emission of titration of 1kb dsDNA (4 nM) with **TO1-6** in DPBS, pH = 7.4. Maximum fluorescence is seen beyond 2300 equivalents, or 9.2 μM, **TO1-6**.

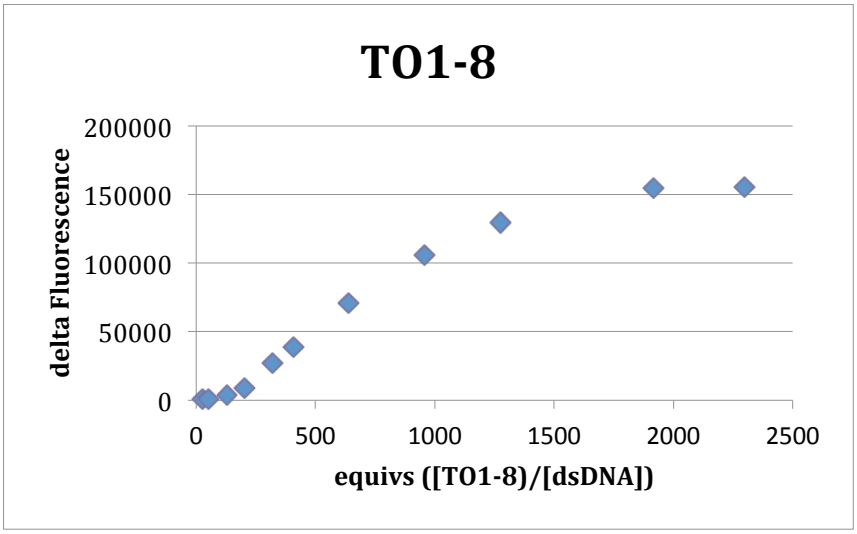


Figure 3.49 Fluorescence emission of titration of 1kb dsDNA (4 nM) with **TO1-8** in DPBS, pH = 7.4. Maximum fluorescence is seen at 1900 equivalents, or 7.6 μM, **TO1-8**.

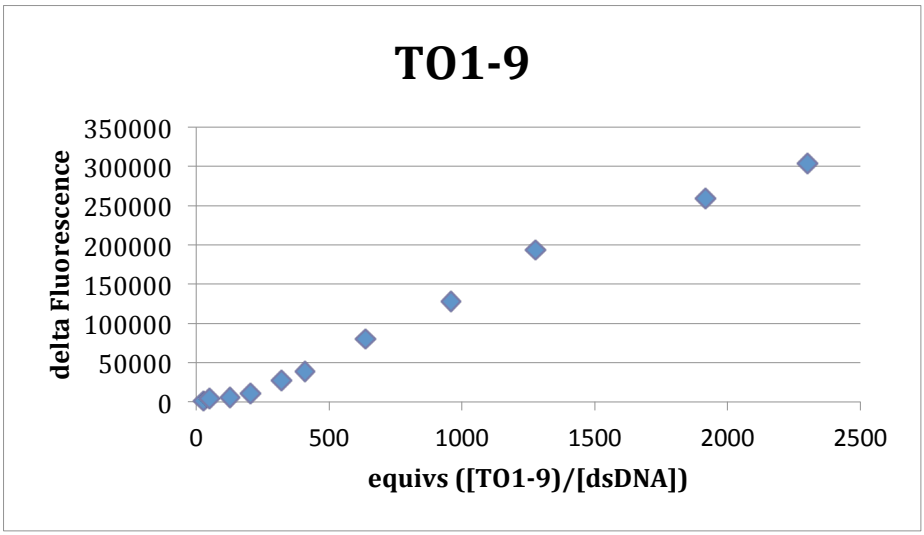


Figure 3.50 Fluorescence emission of titration of 1kb dsDNA (4 nM) with **TO1-9** in DPBS, pH = 7.4. Maximum fluorescence is seen beyond 2300 equivalents, or 9.2 μM, **TO1-9**.

2.5 kb dsDNA

A mixture of nine peptides was titrated into 2.5 kb dsDNA. The relative concentrations of each peptide were determined by the fluorescence response of a peptide alone to the same strand (Table 3.8). 10.44 μM of peptide mixture were needed to reach maximum fluorescence, which is a little lower than average peptide concentration (12.88 μM) (Figure 3.51).

TO1-peptide	Concentration (μM)
TO1-1	30
TO1-2	7.5
TO1-3	6
TO1-4	10
TO1-5	6
TO1-6	15
TO1-7	4
TO1-8	12.5
TO1-9	25
total	116

Table 3.8 Concentrations of peptides used in the peptide mixture, as determined from previous titrations of each peptide with 2 nM 2.5 kb dsDNA in DPBS.

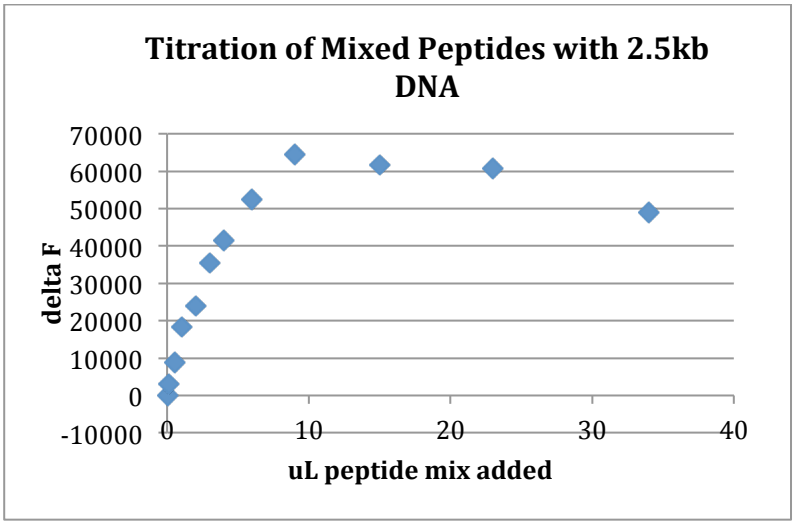


Figure 3.51 Titration of 2 nM 2.5 kb dsDNA with 9-peptide mixture in DPBS. Maximum fluorescence is seen at a dilution of 9 uL into 100 uL total volume.

3.4.4 Cell Culture

Cells were obtained from ATCC (Manassas, VA) and cultured in the recommended growth media supplemented with 10% fetal bovine serum (FBS) and 1x penicillin-streptomycin antibiotic (100 IU/mL and 100 ug/mL, respectively). The Ellington lab provided line U87MGvIII. This line was cultured with an additional antibiotic, G418, which was used in the selection of the cell line transfected with the gene encoding for the EGFRvIII mutant often observed in glioblastomas.

HCT-15, MOLT-4, and 786-O cells were cultured in RPMI-1640 medium, SK-MEL-28, DU-145, and U87MGvIII were cultured in EMEM, A-549 was cultured in F-12K medium, MDA-MB-231 was cultured in DMEM, SK-O-V3 was cultured in McCoy's 5A Medium. Cells were grown in T-25 tissue culture treated flasks (Corning® canted neck 0.2 μM vent cap) until n=4, at which point they were passaged into T-75 tissue culture treated flasks and grown to confluency.

3.4.5 Cancer Cell Patterning

3.4.5.1 Array Generation

Stock solutions of peptides were prepared by dissolving solid peptide into DPBS. For **TO1-6**, a mixture of methanol and DMSO was required to solubilize the peptide. Stocks were diluted to final concentrations of **TO1-peptides** as determined by titrations with 1 kb dsDNA (Table 3.X). These solutions were prepared in DPBS pH = 7.4 with and without 4 nM dsDNA and allowed to incubate at 4 °C overnight.

TO1-peptide	Concentration (μM)
TO1-1	10.0
TO1-2	0.8
TO1-3	5.8
TO1-4	7.8
TO1-5	10.0
TO1-6	10.0
TO1-7	3.7
TO1-8	7.8
TO1-9	10.0

Table 3.9 Concentrations of peptides used in cell patterning experiments, as determined by titration with 4 nM 1 kb dsDNA.

After five passages, cells of one type were passaged from T-75 flasks into three clear-bottom black polystyrene tissue culture treated (costar 3603) or poly-D-lysine treated (Corning® BioCoat) 96-well plates and allowed to grow to confluency for at least 24 hours. Growth medium was removed, and 200 uL DPBS heated to 37°C was used to wash each well. Solutions of nine peptides warmed to 37 °C, with or without DNA, were added to appropriate wells, and three repeats of the same solution on each plate provided nine total repeats over the three plates. Solutions of cells with DPBS and peptides without cells were also prepared on each plate. The well plates were incubated at 37°C with 5% CO_2 for 40 minutes. The peptide solutions were removed and each well with cells was again washed with warm DPBS. Fluorescence emission of each well plate was measured three times and the fluorescence response was averaged over the three measurements.

3.4.5.2 Chemometric Analysis

Fluorescence responses of each peptide were expressed as a percentage compared to the total fluorescence of all of the peptides in the array. This was accomplished by summing the fluorescence values of one repeat of each peptide, dividing one repeat by the sum, and then multiplying by 100. Such values were calculated for the nine repeats of peptides exposed with DNA and the nine repeats of peptides without DNA, as well as for the samples not exposed to cells. The latter percentage values were subtracted from those of the replicates in order to provide a delta % value that showed how each peptide's contribution to the total fluorescence changed before and after exposure to cells.

The delta % values for the nine repeats of each cell line were used as input for chemometric analysis. Using the XLSTAT 2011 program for Microsoft Excel (Addinsoft), Discriminant Analysis and Principal Component Analysis were performed on the data sets. DA was performed using linear functions (assumption that the within class covariance matrices are equal), taking into account the prior probabilities, and using a 5% significance level. Box's test using the Fisher F asymptotic approximation showed that this assumption was valid. PCA was used to reduce the dimensionality of the multivariate data set using the correlation matrix (Pearson's n), and the components that described at least 70% of the variance were retained for further analysis. Three-dimensional score plots were created using the Statistica software package (Dell, formerly StatSoft).

3.5 REFERENCES

1. (a) Bajaj, A.; Miranda, O. R.; Kim, I. B.; Phillips, R. L.; Jerry, D. J.; Bunz, U. H. F.; Rotello, V. M., Detection and differentiation of normal, cancerous, and metastatic cells using nanoparticle-polymer sensor arrays. *P Natl Acad Sci USA* **2009,** *106* (27), 10912-10916; (b) Bajaj, A.; Miranda, O. R.; Phillips, R.; Kim, I. B.; Jerry, D. J.; Bunz, U. H. F.; Rotello, V. M., Array-Based Sensing of Normal,

Cancerous, and Metastatic Cells Using Conjugated Fluorescent Polymers. *J Am Chem Soc* **2010,** *132* (3), 1018-1022.

2. You, C. C.; Agasti, S. S.; Rotello, V. M., Isomeric control of protein recognition with amino acid- and dipeptide-functionalized gold nanoparticles. *Chem-Eur J* **2008,** *14* (1), 143-150.

3. Gumbiner, B. M., Cell Adhesion: The Molecular Basis of Tissue Architecture and Morphogenesis. *Cell* **1996,** *84* (3), 345-357.

4. Shoemaker, R. H., The NCI60 human tumour cell line anticancer drug screen. *Nat Rev Cancer* **2006,** *6* (10), 813-823.

5. Lieber, M.; Todaro, G.; Smith, B.; Szakal, A.; Nelson-Rees, W., A continuous tumor-cell line from a human lung carcinoma with properties of type II alveolar epithelial cells. *International Journal of Cancer* **1976,** *17* (1), 62-70.

6. Friedman, H. S.; Bigner, D. D., Glioblastoma Multiforme and the Epidermal Growth Factor Receptor. *New England Journal of Medicine* **2005,** *353* (19), 1997-1999.

7. Hua, W.; Christianson, T.; Rougeot, C.; Rochefort, H.; Clinton, G. M., SKOV3 ovarian carcinoma cells have functional estrogen receptor but are growth-resistant to estrogen and antiestrogens. *J Steroid Biochem Mol Biol* **1995,** *55* (3-4), 279-89.

8. Shawver, L. K.; Mann, E.; Elliger, S. S.; Dugger, T. C.; Arteaga, C. L., Ligand-like Effects Induced by Anti-c-erbB-2 Antibodies Do Not Correlate with and Are Not Required for Growth Inhibition of Human Carcinoma Cells. *Cancer Research* **1994,** *54* (5), 1367-1373.

9. Bunka, D. H. J.; Stockley, P. G., Aptamers come of age ‚Äì at last. *Nat Rev Micro* **2006,** *4* (8), 588-596.

10. Gresham, D.; Dunham, M. J.; Botstein, D., Comparing whole genomes using DNA microarrays. *Nat Rev Genet* **2008,** *9* (4), 291-302.

11. Hansen, L. F.; Jensen, L. K.; Jacobsen, J. P., Bis-Intercalation of a Homodimeric Thiazole Orange Dye in DNA in Symmetrical Pyrimidine-Pyrimidine-Purine-Purine Oligonucleotides. *Nucleic Acids Research* **1996,** *24* (5), 859-867.

12. Mulder, A.; Huskens, J.; Reinhoudt, D. N., Multivalency in supramolecular chemistry and nanofabrication. *Org Biomol Chem* **2004,** *2* (23), 3409-3424.

13. Nygren, J.; Svanvik, N.; Kubista, M., The interactions between the fluorescent dye thiazole orange and DNA. *Biopolymers* **1998,** *46* (1), 39-51.

14. Yarmoluk, S. M.; Losytskyy, M. Y.; Yashchuk, V. M., Nonradiative deactivation of the electronic excitation energy in cyanine dyes: influence of binding to DNA. *J Photoch Photobio B* **2002,** *67* (1), 57-63.

15. Rye, H. S.; Yue, S.; Wemmer, D. E.; Quesada, M. A.; Haugland, R. P.; Mathies, R. A.; Glazer, A. N., Stable Fluorescent Complexes of Double-Stranded DNA with Bis-Intercalating Asymmetric Cyanine Dyes - Properties and Applications. *Nucleic Acids Res* **1992,** *20* (11), 2803-2812.

16. Svanvik, N.; Westman, G.; Wang, D. Y.; Kubista, M., Light-up probes: Thiazole orange-conjugated peptide nucleic acid for detection of target nucleic acid in homogeneous solution. *Anal Biochem* **2000,** *281* (1), 26-35.

17. Bethge, L.; Jarikote, D. V.; Seitz, O., New cyanine dyes as base surrogates in PNA: Forced intercalation probes (FIT-probes) for homogeneous SNP detection. *Bioorgan Med Chem* **2008,** *16* (1), 114-125.

18. Babendure, J.; Liddell, P. A.; Bash, R.; LoVullo, D.; Schiefer, T. K.; Williams, M.; Daniel, D. C.; Thompson, M.; Taguchi, A. K. W.; Lohr, D.; Woodbury, N. W., Development of a fluorescent probe for the study of nucleosome assembly and dynamics. *Anal Biochem* **2003,** *317* (1), 1-11.

19. Pei, R.; Stojanovic, M. N., Study of thiazole orange in aptamer-based dye-displacement assays. *Anal Bioanal Chem* **2008,** *390* (4), 1093-1099.

20. (a) Thompson, M.; Woodbury, N. W., Fluorescent and photochemical properties of a single zinc finger conjugated to a fluorescent DNA-binding probe. *Biochemistry-Us* **2000,** *39* (15), 4327-4338; (b) Carreon, J. R.; Mahon, K. P.; Kelley, S. O., Thiazole orange-peptide conjugates: Sensitivity of DNA binding to chemical structure. *Org Lett* **2004,** *6* (4), 517-519; (c) Thompson, M., Synthesis, photophysical effects, and DNA targeting properties of oxazole yellow-peptide bioconjugates. *Bioconjugate Chem* **2006,** *17* (2), 507-513; (d) Thompson, M., Spectral properties and DNA targeting features of a thiazole orange-peptide bloconjugate. *Biomacromolecules* **2007,** *8* (11), 3628-3633.

21. (a) Netzel, T. L.; Nafisi, K.; Zhao, M.; Lenhard, J. R.; Johnson, I., Base-Content Dependence of Emission Enhancements, Quantum Yields, and Lifetimes for Cyanine Dyes Bound to Double-Strand DNA: Photophysical Properties of Monomeric and Bichromomphoric DNA Stains. *The Journal of Physical Chemistry* **1995,** *99* (51), 17936-17947; (b) Larsson, A.; Carlsson, C.; Jonsson, M., Characterization of the Binding of Yo to [Poly(Da-Dt)](2) and [Poly(Dg-Dc)](2), and of the Fluorescent Properties of Yo and Yoyo Complexed with the Polynucleotides and Double-Stranded DNA. *Biopolymers* **1995,** *36* (2), 153-167; (c) Armitage, B. A., Cyanine dye-DNA interactions: Intercalation, groove binding, and aggregation. *Top Curr Chem* **2005,** *253*, 55-76; (d) Yarmoluk, S. M.; Lukashov, S. S.; Ogul'chansky, T. Y.; Losytskyy, M. Y.; Kornyushyna, O. S., Interaction of cyanine dyes with nucleic acids. XXI. Arguments for half-intercalation model of interaction. *Biopolymers* **2001,** *62* (4), 219-227; (e) Petty, J. T.; Bordelon, J. A.; Robertson, M. E., Thermodynamic characterization of the

association of cyanine dyes with DNA. *J Phys Chem B* **2000,** *104* (30), 7221-7227.

22. (a) Carlsson, C.; Larsson, A.; Jonsson, M.; Albinsson, B.; Norden, B., Optical and Photophysical Properties of the Oxazole Yellow DNA Probes Yo and Yoyo. *J Phys Chem-Us* **1994,** *98* (40), 10313-10321; (b) Dang, F. Q.; Li, W. H.; Zhang, L. H.; Jabasini, M.; Ishida, T.; Kiwada, H.; Kaji, N.; Tokeshi, M.; Baba, Y., Electrophoretic behavior of plasmid DNA in the presence of various intercalating dyes. *J Chromatogr A* **2006,** *1118* (2), 218-225; (c) Eriksson, M.; Mehmedovic, M.; Westman, G.; Akerman, B., Time-resolved electrophoretic analysis of mobility shifts for dissociating DNA ligands. *Electrophoresis* **2005,** *26* (3), 524-532.

23. Kolonin, M. G.; Bover, L.; Sun, J.; Zurita, A. J.; Do, K.-A.; Lahdenranta, J.; Card√≥-Vila, M.; Giordano, R. J.; Jaalouk, D. E.; Ozawa, M. G.; Moya, C. A.; Souza, G. R.; Staquicini, F. I.; Kunyiasu, A.; Scudiero, D. A.; Holbeck, S. L.; Sausville, E. A.; Arap, W.; Pasqualini, R., Ligand-Directed Surface Profiling of Human Cancer Cells with Combinatorial Peptide Libraries. *Cancer Research* **2006,** *66* (1), 34-40.

24. Yang, B.; Zhang, L.; Turley, E. A., Identification of two hyaluronan-binding domains in the hyaluronan receptor RHAMM. *Journal of Biological Chemistry* **1993,** *268* (12), 8617-8623.

25. (a) Healy, J. M.; Murayama, O.; Maeda, T.; Yoshino, K.; Sekiguchi, K.; Kikuchi, M., Peptide Ligands for Integrin .alpha.v.beta.3 Selected from Random Phage Display Libraries. *Biochemistry* **1995,** *34* (12), 3948-3955; (b) Rahman, S.; Aitken A Fau - Flynn, G.; Flynn G Fau - Formstone, C.; Formstone C Fau - Savidge, G. F.; Savidge, G. F., Modulation of RGD sequence motifs regulates disintegrin recognition of alphaIIb beta3 and alpha5 beta1 integrin complexes. Replacement of elegantin alanine-50 with proline, N-terminal to the RGD sequence, diminishes recognition of the alpha5 beta1 complex with restoration induced by Mn2+ cation. *Biochemical Journal* **1998,** *335*, 247-257.

26. Pramanik, D.; Majeti, B. K.; Mondal, G.; Karmali, P. P.; Sistla, R.; Ramprasad, O. G.; Srinivas, G.; Pande, G.; Chaudhuri, A., Lipopeptide with a RGDK Tetrapeptide Sequence Can Selectively Target Genes to Proangiogenic Œ±5Œ≤1 Integrin Receptor and Mouse Tumor Vasculature. *Journal of Medicinal Chemistry* **2008,** *51* (22), 7298-7302.

27. Yamada, K. M.; Kennedy, D. W., Peptide inhibitors of fibronectin, laminin, and other adhesion molecules: Unique and shared features. *Journal of Cellular Physiology* **1987,** *130* (1), 21-28.

28. Dudak, F. C.; Boyaci, I. H.; Orner, B. P., The Discovery of Small-Molecule Mimicking Peptides through Phage Display. *Molecules* **2011,** *16* (1), 774-789.

29. Jones, J. T.; Akita, R. W.; Sliwkowski, M. X., Binding specificities and affinities of egf domains for ErbB receptors. *FEBS Letters* **1999,** *447* (2,Äì3), 227-231.

30. Ruoslahti, E., RGD AND OTHER RECOGNITION SEQUENCES FOR INTEGRINS. *Annual Review of Cell and Developmental Biology* **1996,** *12* (1), 697-715.

31. Koivunen, E.; Gay, D. A.; Ruoslahti, E., Selection of peptides binding to the alpha 5 beta 1 integrin from phage display library. *Journal of Biological Chemistry* **1993,** *268* (27), 20205-20210.

32. Borron, P. J.; Crouch, E. C.; Lewis, J. F.; Wright, J. R.; Possmayer, F.; Fraher, L. J., Recombinant Rat Surfactant-Associated Protein D Inhibits Human T Lymphocyte Proliferation and IL-2 Production. *The Journal of Immunology* **1998,** *161* (9), 4599-4603.

33. Lotan, R.; Raz, A., Lectins in Cancer Cells. *Annals of the New York Academy of Sciences* **1988,** *551* (1), 385-398.

34. Sironen, R. K.; Tammi, M.; Tammi, R.; Auvinen, P. K.; Anttila, M.; Kosma, V. M., Hyaluronan in human malignancies. *Experimental Cell Research* **2011,** *317* (4), 383-391.

35. Arteaga, C. L.; Engelman, J. A., ERBB Receptors: From Oncogene Discovery to Basic Science to Mechanism-Based Cancer Therapeutics. *Cancer Cell* **2014,** *25* (3), 282-303.

36. Greene, R. S.; Alderfer, J.; Munson, B. R., In vitro effects of Acridine intercalation on RNA polymerase interactions with supercoiled DNA. *International Journal of Biochemistry* **1983,** *15* (10), 1231-1239.

37. (a) Garoff, R. A.; Litzinger, E. A.; Connor, R. E.; Fishman, I.; Armitage, B. A., Helical aggregation of cyanine dyes on DNA templates: Effect of dye structure on formation of homo- and heteroaggregates. *Langmuir* **2002,** *18* (16), 6330-6337; (b) Seifert, J. L.; Connor, R. E.; Kushon, S. A.; Wang, M.; Armitage, B. A., Spontaneous assembly of helical cyanine dye aggregates on DNA nanotemplates. *J Am Chem Soc* **1999,** *121* (13), 2987-2995.

38. Cannistra, S. A.; Ottensmeier, C.; Niloff, J.; Orta, B.; DiCarlo, J., Expression and Function of Œ≤1 and Œ±vŒ≤3 Integrins in Ovarian Cancer. *Gynecologic Oncology* **1995,** *58* (2), 216-225.

39. Casey, R.; Skubitz, A. N., CD44 and Œ≤1 integrins mediate ovarian carcinoma cell migration toward extracellular matrix proteins. *Clin Exp Metastasis* **2000,** *18* (1), 67-75.

40. Hung, M. C.; Zhang, X.; Yan, D. H.; Zhang, H. Z.; He, G. P.; Zhang, T. Q.; Shi, D. R., Aberrant expression of the c-erbB-2/neu protooncogene in ovarian cancer. *Cancer Lett* **1992,** *61* (2), 95-103.

41. Taherian, A.; Li, X.; Liu, Y.; Haas, T. A., Differences in integrin expression and signaling within human breast cancer cells. *BMC Cancer* **2011,** *11*, 293.

42. Subik, K.; Lee, J. F.; Baxter, L.; Strzepek, T.; Costello, D.; Crowley, P.; Xing, L.; Hung, M. C.; Bonfiglio, T.; Hicks, D. G.; Tang, P., The Expression Patterns of ER, PR, HER2, CK5/6, EGFR, Ki-67 and AR by Immunohistochemical Analysis in Breast Cancer Cell Lines. *Breast Cancer (Auckl)* **2010,** *4*, 35-41.

43. Penno, M. B.; August, J. T.; Baylin, S. B.; Mabry, M.; Linnoila, R. I.; Lee, V. S.; Croteau, D.; Yang, X. L.; Rosada, C., Expression of CD44 in Human Lung Tumors. *Cancer Research* **1994,** *54* (5), 1381-1387.

44. (a) Chow, G.; Tauler, J.; Mulshine, J. L., Cytokines and Growth Factors Stimulate Hyaluronan Production: Role of Hyaluronan in Epithelial to Mesenchymal-Like Transition in Non-Small Cell Lung Cancer. *Journal of Biomedicine and Biotechnology* **2010,** *2010*, 11; (b) Pirinen, R.; Tammi, R.; Tammi, M.; Hirvikoski, P.; Parkkinen, J. J.; Johansson, R.; Böhm, J.; Hollmén, S.; Kosma, V.-M., Prognostic value of hyaluronan expression in non-small-cell lung cancer: Increased stromal expression indicates unfavorable outcome in patients with adenocarcinoma. *International Journal of Cancer* **2001,** *95* (1), 12-17.

45. Oyanagi, J.; Ogawa, T.; Sato, H.; Higashi, S.; Miyazaki, K., Epithelial-Mesenchymal Transition Stimulates Human Cancer Cells to Extend Microtubule-based Invasive Protrusions and Suppresses Cell Growth in Collagen Gel. *PLoS ONE* **2012,** *7* (12), e53209.

46. Schneider, P. M.; Hung, M.-C.; Chiocca, S. M.; Manning, J.; Zhao, X.; Fang, K.; Roth, J. A., Differential Expression of the c-erbB-2 Gene in Human Small Cell and Non-Small Cell Lung Cancer. *Cancer Research* **1989,** *49* (18), 4968-4971.

47. Rachwal, W. J.; Bongiorno, P. F.; Orringer, M. B.; Whyte, R. I.; Ethier, S. P.; Beer, D. G., Expression and activation of erbB-2 and epidermal growth factor receptor in lung adenocarcinomas. *Br J Cancer* **1995,** *72* (1), 56-64.

48. Esteban-Barragán, M. A.; Ávila, P.; Álvarez-Tejado, M.; Gutiérrez, M. D.; García-Pardo, Á.; Sánchez-Madrid, F.; Landázuri, M. O., Role of the von Hippel-Lindau Tumor Suppressor Gene in the Formation of β1-Integrin Fibrillar Adhesions. *Cancer Research* **2002,** *62* (10), 2929-2936.

49. Mierke, C. T.; Frey, B.; Fellner, M.; Herrmann, M.; Fabry, B., Integrin α5β1 facilitates cancer cell invasion through enhanced contractile forces. *Journal of Cell Science* **2011,** *124* (3), 369-383.

50. Sanchez, E. E.; Rodriguez-Acosta, A.; Palomar, R.; Lucena, S.; Bashir, S.; Soto, J. G.; Perez, J. C., Colombistatin: a disintegrin isolated from the venom of the South American snake (Bothrops colombiensis) that effectively inhibits platelet aggregation and SK-Mel-28 cell adhesion. *Archives of Toxicology* **2009,** *83* (3), 271-279.

51. Witkowski, C.; Rabinovitz, I.; Nagle, R.; Affinito, K.-S.; Cress, A., Characterization of integrin subunits, cellular adhesion and tumorgenicity of four human prostate cell lines. *Journal of Cancer Research and Clinical Oncology* **1993,** *119* (11), 637-644.

52. El Sheikh, S. S.; Domin, J.; Abel, P.; Stamp, G.; Lalani el, N., Phosphorylation of both EGFR and ErbB2 is a reliable predictor of prostate cancer cell proliferation in response to EGF. *Neoplasia* **2004,** *6* (6), 846-53.

53. Higa, K.; Shimmura, S.; Shimazaki, J.; Tsubota, K., Hyaluronic acid-CD44 interaction mediates the adhesion of lymphocytes by amniotic membrane stroma. *Cornea* **2005,** *24* (2), 206-12.

54. Stoolman, L. M.; Wang, T. L.; Situ, R.; Varani, J., Regulation of fibronectin and laminin binding activity in cultured human lymphoblastic cell lines. *J Cell Physiol* **1993,** *154* (3), 593-600.

55. Melenhorst, W. B. W. H.; Mulder, G. M.; Xi, Q.; Hoenderop, J. G. J.; Kimura, K.; Eguchi, S.; van Goor, H., Epidermal Growth Factor Receptor Signaling in the Kidney: Key Roles in Physiology and Disease. *Hypertension* **2008,** *52* (6), 987-993.

56. Bernert, B.; Porsch, H.; Heldin, P., Hyaluronan synthase 2 (HAS2) promotes breast cancer cell invasion by suppression of tissue metalloproteinase inhibitor 1 (TIMP-1). *J Biol Chem* **2011,** *286* (49), 42349-59.

Chapter 4: Glycoform Discrimination of Cell Surfaces

4.1 Background

The cell surface is covered in a fuzzy, glycoprotein-polysaccharide coating known as the glycocalyx. Carbohydrates on glycolipids and glycoproteins contribute to cell-cell recognition, and the glycocalyx is important for organisms to distinguish healthy endogenous cells from diseased or exogenous cells. Changes in the glycocalyx of cancer cells that result in the presence of bulky glycoproteins help to increase integrin adhesion and signaling.[1] Large glycoproteins are expressed on circulating tumor cells in patients with advanced disease, indicating that such glycocalyx changes are associated with cancer cell mobility and metastasis. The increase in size of tumor cell glycopeptides can be explained in part by the increase in β-1,6 branching of *N*-linked glycans.[2] Other glycan changes that are associated with cancer include altered mucin glycosylation in carcinomas, increase in sialylation, overexpression of sialyl Lewis structures, and high expression of hyaluronan.

Antibodies, lectins, aptamers, and small molecules have been developed as carbohydrate sensors. Small molecule synthetic molecular receptors have been developed that recognize carbohydrates through hydrogen bonding. These interactions are disrupted in aqueous media due to competition from solvent. However, in aqueous media boronic acids form reversible covalent bonds with hydroxyl groups, forming cyclic boronic or boronate esters with 1,2- and 1,3-diols, polyols, catechols, and α-hydroxyacids (Figure 4.1).[3] The covalent nature of the interaction allows for analyte binding in competitive media, and carbohydrates display many hydroxyls that are available for bonding with boronic acids. The binding of a monoboronic acid to sugars in solution is on the order of

10^2-10^3 M^{-1}.[4] The interaction of boronic acids with diols is dependent on the pK_as of the boronic acid and of the diol, the pH of the solution, and sterics.[5] 1,2-diols that form 5-membered rings react preferentially over the 6-membered ring forming 1,3-diols, and *cis*-diols are preferred due to their more favorable complex conformation (see Figure 4.1). Increases in acidity of boronic acids can increase the association with diols. Electron withdrawing groups lower the pK_a of boronic acids and therefore increase the affinity towards diols. The pK_a of phenylboronic acid is 8.8, while the pK_as of 3-aminophenylboronic acid and 3-nitrophenylboronic acid are about 9 and 7, respectively. *ortho*-Aminomethylphenylboronic acid has a pK_a value that is lower than expected based on the electron withdrawing ability of the substituents; this is due to the formation of a zwitterionic, ion paired species. Generally, ligands with higher acidity have more favorable association constants with boronic acids, but solution pH affects affinities of boronic acids for diols variably. The optimal binding does not necessarily occur at a pH above the pK_a of the boronic acid, but instead usually occurs at a pH between the pK_as of the boronic acid and the diol. Lowering the pK_a of the boronic acid therefore should enhance the binding affinity for non-acidic carbohydrates at neutral pH.

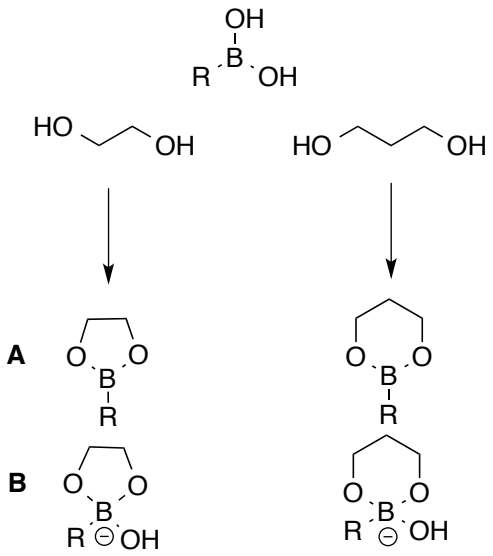


Figure 4.1 Formation of boronic esters (**A**) and boronate esters (**B**) from 1,2- and 1,3-diols.

In general, monoboronic acid receptors display carbohydrate selectivity on the order of fructose > galactose > glucose > mannose.[3] However, sugars are capable of forming a variety of structures in solution; pyranose, furanose, and open chain structures exist in equilibrium in aqueous solution as well as their α- and β-anomers. The preferred mode of boronic acid binding to sugars is not necessarily the preferred conformation in solution. For example, aromatic boronic acids have been shown to bind D-glucose in its furanose form,[6] even though pyranose is largely (>99%) preferred in solution (Figure 4.2). Glucose is able to form two boronate esters with two monoboronic acids in its α-furanose form via bidentate interaction with the *cis*-hydroxyls at the 1,2-positions and tridentate interaction with the 3,5,6-hydroxyls, and this effect is achieved with only two equivalents of boronic acid.

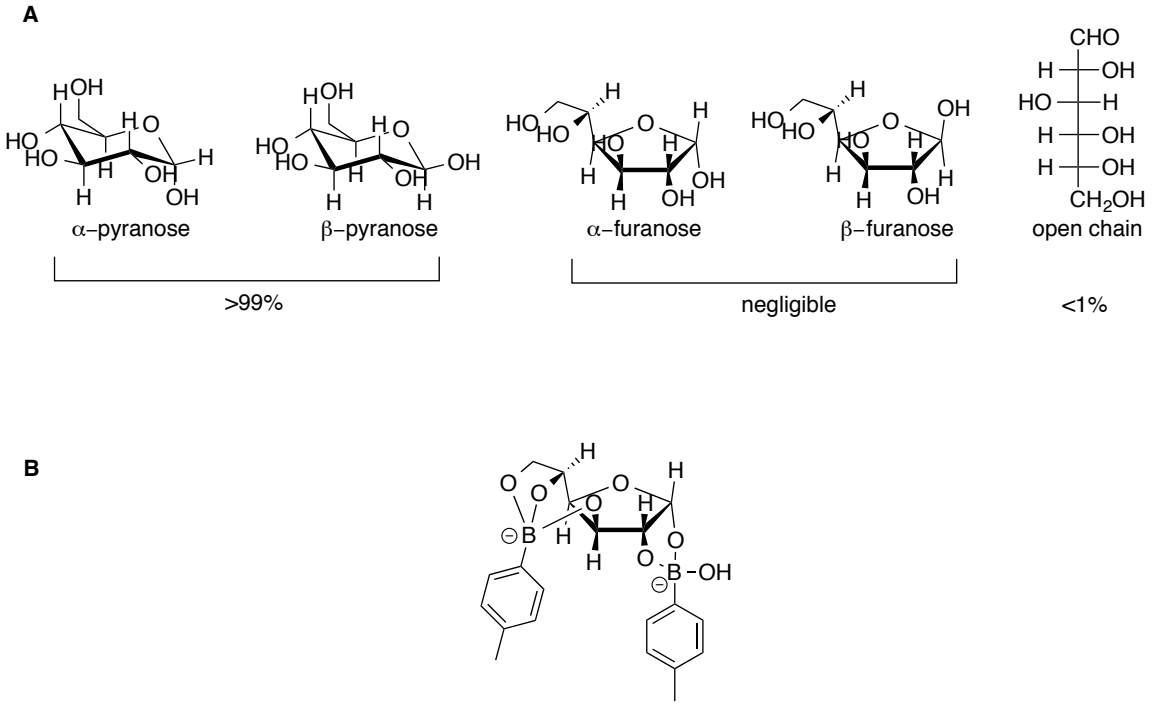


Figure 4.2 The forms of glucose and their preferences at aqueous equilibrium (**A**) and the structure of a boronic acid complex with D-glucose, preferentially formed with the α-furanose form (**B**).[6]

Both aryl boronic acids and alpha-amido boronic acids have been used to bind sugars in aqueous media, though the former sometimes suffer from low water solubility. On cell surfaces, sugars exist primarily in the pyranose form. Nevertheless, boronic acid-based receptors have been developed to target saccharides on cells; the most effective of these are multivalent. The first instance of a such a receptor targeting cell surface sugars was a diboronic acid developed by Hageman *et al*. that was capable of agglutinating erythrocytes and therefore hypothesized to have functionality similar to a lectin.[7] The Wang group reported the development of a different diboronic acid receptor that was selective for cell surface sialyl Lewis^x.[8] Although mannose and galactose display *cis*-vicinal diols ideal for boronic acid complexation, arylboronic acids do not bind well to sugars in their pyranose form, and instead prefer the furanose form. Benzoboroxole-based

repectors have shown affinity for pyranosides in aqueous solutions at physiological pH due to their high Lewis acidity and hydrogen bonding capability.[9]

4.2 Indicator Displacement Assay for Boronic Acid Binding to Cells

An array consisting of boronic acids of varying structure and pH indicators was developed for the discrimination of cancer cell glycocalyces. It was hypothesized that interaction of the cell surface with boronic acids would result in disturbance of the equilibrium of boronic acid-indicator complexes, resulting in a color change indicative of cell binding. The different boronic acids were hypothesized to have variable affinity for cell surfaces based on the boronic acid and glycocalyx structures and respective complementarity. Variable response of each boronic acid-indicator pair would provide a unique fingerprint for each cell type.

Six commercially available boronic acids (**4.1**-**4.6**) displaying a variety of functionalites were chosen for the array (Figure 4.3). These boronic acids were combined with three commercially available indicators that contain catechols for binding boronic acids: Alizarin Red S (ARS), Pyrocatechol Violet (PV), and 4-methylesculetin (4ME). Benzoboroxole **4.5** was chosen because this type of boronic acid has been shown to bind pyranoside carbohydrates, like those on cell surfaces, in aqueous solutions. Other phenyl boronic acids have functional groups that contribute variably to the pKa, such as pK_a elevating electron-donating groups (**4.2**, **4.6**) and pK_a lowering groups (**4.1**, **4.3**). The non-aromatic **4.4** was chosen to provide hydrophobic contacts with the cell surface, despite its lower acidity (pK_a 10.4).

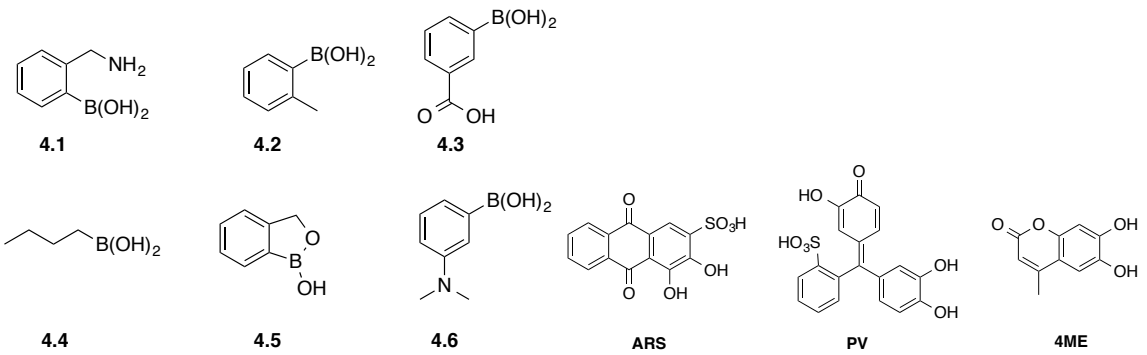


Figure 4.3 The boronic acids and indicators selected for the glycocalyx-patterning array.

4.2.1 Boronic Acid-Indicator Complexes

Titrations of each of the catechol indicators with increasing equivalents of boronic acid were undertaken. The indicator concentrations were chosen such that their absorbance values fell within the range of 0.4 – 2.0 absorbance units in order to provide sufficient signal-to-noise ratios. Increasing equivalents of boronic acid resulted in a decrease in the free indicator peak and a simultaneous increase in the absorbance of the boronic acid-indicator complex. The presence of an isosbestic point was indicative of two absorbing species in solution. For ARS, this manifested as an increase at 460 nm, a decrease at 510 nm, and an isosbestic point at 480 nm (Figure 4.4).

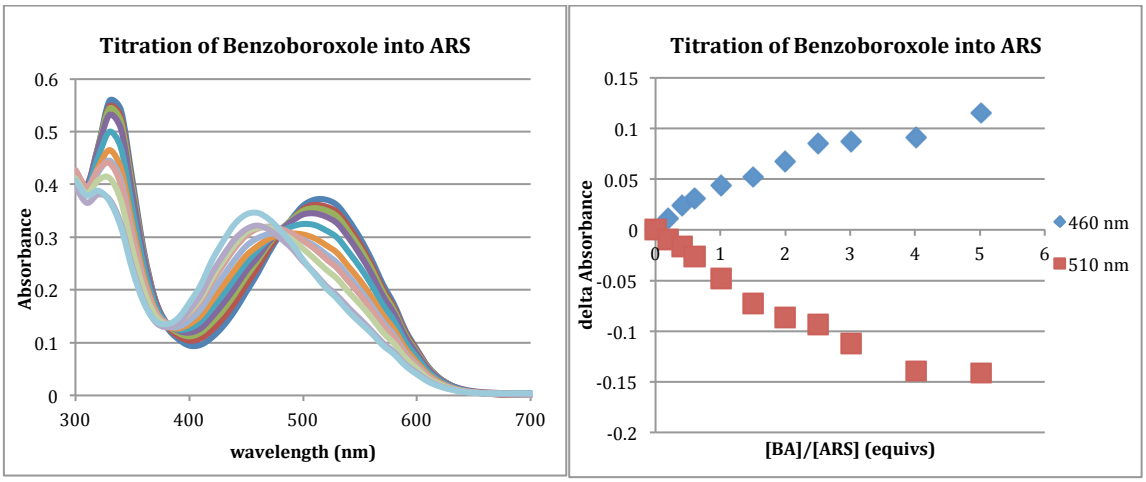


Figure 4.4 Changes in the absorbance spectrum upon titration of **4.5** (0–1 mM) into ARS (200 μM) in 10 mM HEPES 0.1 M NaCl, pH = 7.4 (left). The free indicator and complex absorbances are plotted as a function of equivalents of boronic acid relative to indicator (right).

Fitting of the absorbance increase at 460 nm and the absorbance decrease at 510 nm using a 1:1 binding model lead to an estimation of the association constant between **4.5** and ARS on the order of 2×10^{3} M^{-1}. The magnitude of this binding constant is in agreement with literature values for association of monoboronic acids with diols in neutral aqueous solution.

4.2.2 Indicator Displacement

Increasing numbers of cells were added to boronic acid-indicator complexes in order to investigate the displacement of indicator upon binding of boronic acids to cell surfaces. The absorbance values for the displacement were compared to those of cells added to free indicator. Addition of cells to free indicator resulted in a uniform increase in absorbance associated with the cell density (Figure 4.5). This is due to the light scattering of the cell suspension. The same effect was observed when cells were added to the boronic acid-indicator complex (Figure 4.6).

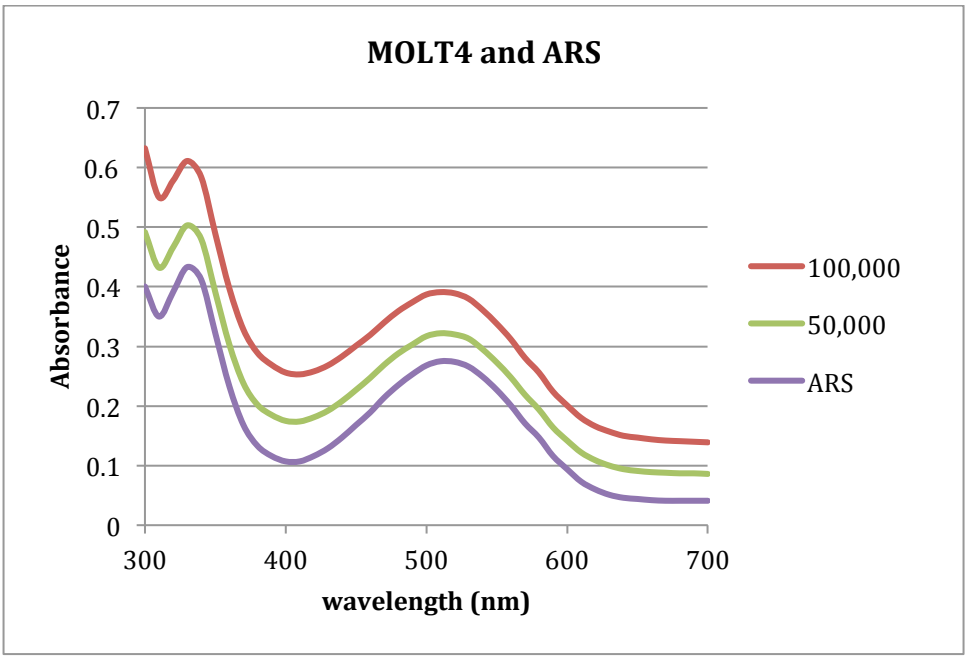


Figure 4.5 Increase in absorbance of free **ARS** (200 μM) upon addition of increasing number of cells in 10 mM HEPES 0.1 M NaCl, pH = 7.4.

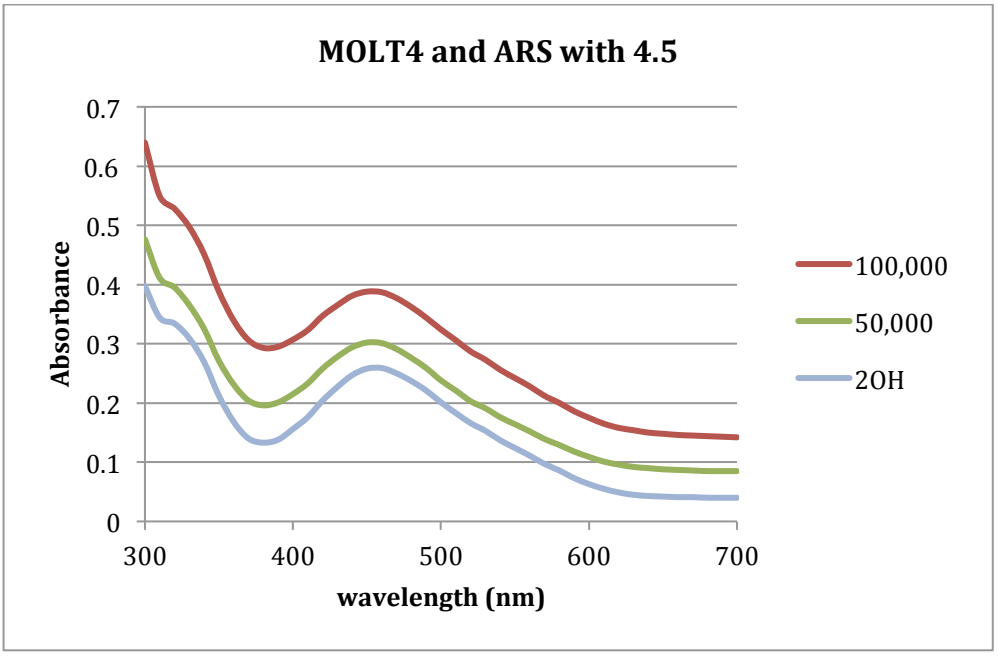


Figure 4.6 Increase in absorbance of boronic acid-indicator complex of **4.5** (800 μM) with ARS (200 μM) upon addition of increasing cells in 10 mM HEPES 0.1 M NaCl, pH = 7.4.

The extent of scatter and absorbance increase differed by cell line for the same number of cells added. This resulted in unique absorbance intensities of indicators for each of the cell lines. The absorbance values of the boronic acid-indicator complexes also displayed intensity differences for each cell line. Although the intensities were perhaps diagnostic, these differences were not due to the expected indicator displacement. The ratio of absorbance intensity of the free indicator and the boronic acid indicator at their respective λ_{max} was plotted for each cell line and compared to the values without cells (Figure 4.7). The intensity of the free indicator peak should increase relative to that of bound indicator if indicator displacement occurs. This would manifest as an increase in the ratio of the λ_{max} intensities. It was clear from Figure 4.7 that displacement of the indicator was not occurring as the ratio of intensities actually decreased across all cell lines. Such an observation may be due to differential changes in molar absorptivites of the free indicator and the indicator-boronic acid complex in the presence of cells, such that the boronic-acid indicator complex shows a larger change in molar absorptivity. It is also likely that indicator displacement does not occur because interaction of the monoboronic acid with cell surface sugars in competitive media is much weaker than indicator complexation.

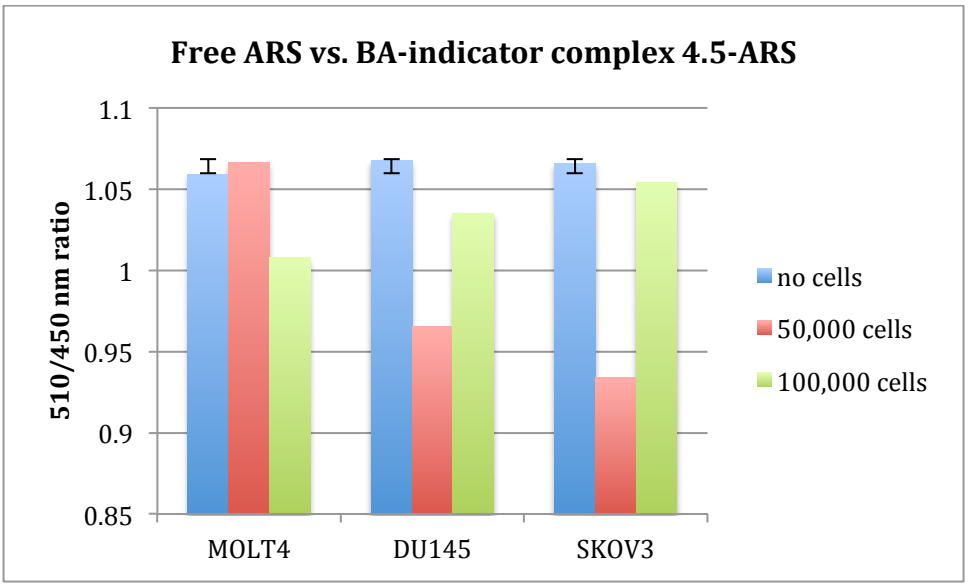


Figure 4.7 Ratio of the λ_{max} intensities of ARS and ARS complexed with **4.5** in 10 mM HEPES 0.1 M NaCl pH= 7.4 with cells of different tissue type added. Error bars show the standard deviation of the ratios of three measurements on cell-free solutions, which are indicative of the errors associated with the absorbance measurements.

It was hypothesized that increasing the concentration of cells relative to the boronic acid indicator complex would make boronic acid binding to the cell surfaces more favorable compared to indicator complexation, despite the differences in affinity. Therefore, an indicator uptake experiment was undertaken at low indicator and boronic acid concentrations. Indicator solutions were added to solutions of boronic acids and cells or free boronic acid. The absorbance intensity changes of these two titrations were compared. It was hypothesized that the uptake of indicator, measured by the appearance of boronic acid-indicator complex, would be different in the presence of cells if boronic acid binding to cell surfaces was occurring. Even using experimental conditions that favored cell binding, there was no appreciable difference in the indicator uptake in the presence of cells, as evidenced by similar absorbance values with and without cells

(Figure 4.8). Indicator uptake did not result in appreciable formation of boronic acid-indicator complex in either case.

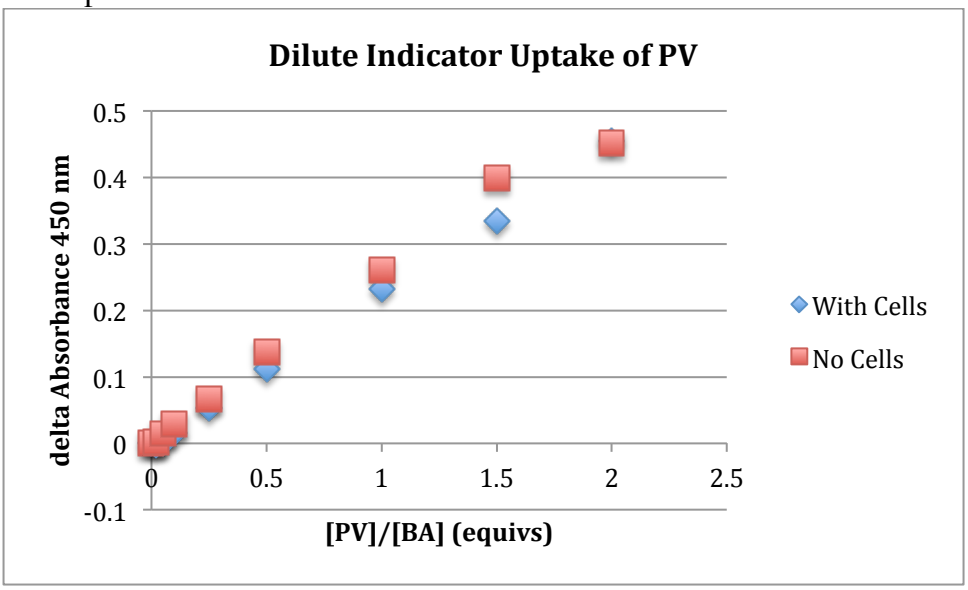


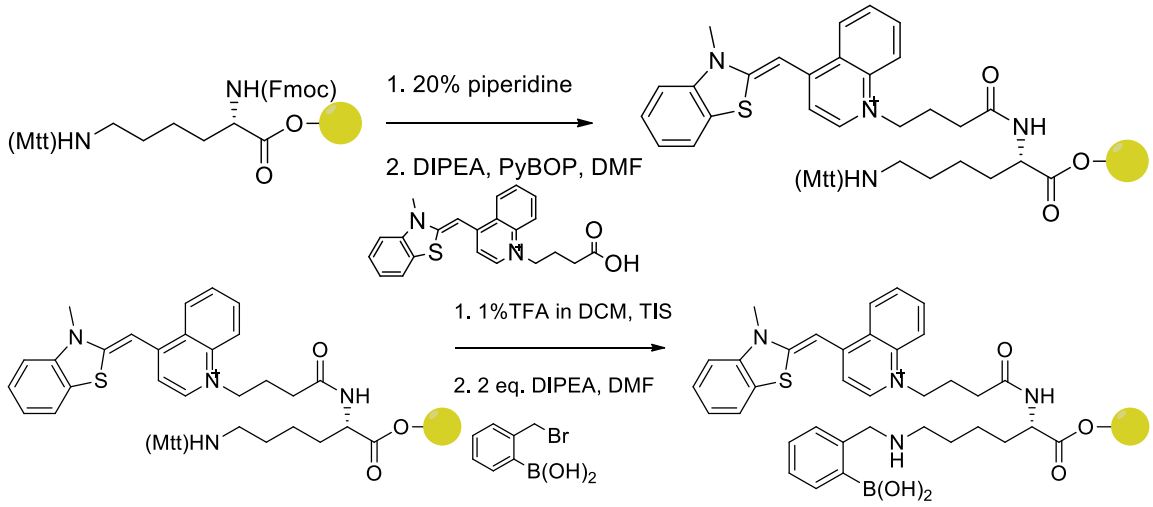
Figure 4.8 Indicator uptake of PV with 100 μM **4.5** in the presence or absence of 50,000 MOLT-4 cells in 10 mM HEPES 0.1 M NaCl, pH = 7.4.

4.3 Fluorescence Response of Cancer Cells to a Lysine-Linked Aromatic Boronic Acid

The synthesis-free array of commercially available boronic acids with indicators developed above could not produce reliable signals of hypothesized boronic acid binding to cell glycocalyces. A more synthetically intense route, where a boronic acid was conjugated to a fluorescent indicator, is presented below. Conjugation of *ortho*-bromomethylboronic acid to a derivative of the DNA intercalator thiazole orange, **TO1**, allows for assembly of the boronic acid onto a DNA scaffold that can present multiple binding units that should increase the association with cell surfaces compared to

monoboronic acids. Additionally, signals can be generated by measuring fluorescence emission upon association of the boronic acid with cell surfaces.

TO1 was conjugated to an Mtt-protected lysine Wang resin, then deprotected and reacted with *ortho*-bromomethylboronic acid (Scheme 4.1). Cleavage from the resin and HPLC purification produced **BALysTO1**. The amino acid conjugate was used in combination with 1 kb dsDNA and incubated with cancer cells of different tissue type at 37°C for 40 minutes. The fluorescence response of the bound boronic acid was recorded for each cell type, and normalized to the starting fluorescence of the free solution (Figure 4.9).



Scheme 4.1 Solid phase synthesis of **BALysTO1** using Fmoc chemistry and a Wang resin.

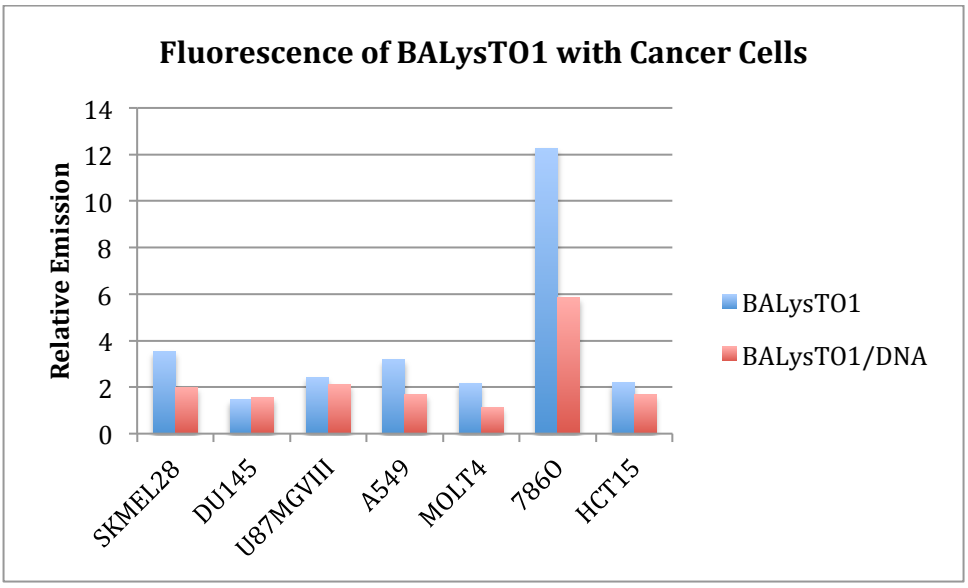


Figure 4.9 The average fluorescence response of **BALysTO1** (30 μM) with and without 1 kb dsDNA (4 nM) to seven different cell lines in DPBS pH=7.4.

The cell line that exhibited the greatest fluorescence response was 786-O, a renal carcinoma. Other cell lines resulted in similar responses to each other. Plotting the relative emission with and without DNA showed some cell classification ability based on response to the boronic acid conjugate (Figure 4.10).

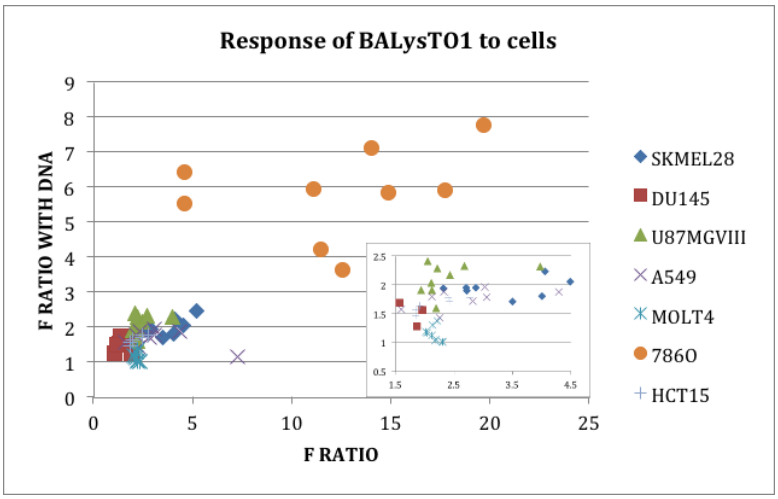


Figure 4.10 Classification of cells from their fluorescence response to **BALysTO1** with or without 2.5 kb dsDNA. *Inset:* Classification of the six cell lines with more similar responses.

4.4 Fluorescence Activated Cell Sorting of Enzyme-Treated Cells

4.4.1 Titrations with 50 base DNA single strands

Single strands of DNA were randomly generated to have 50% GC content, and two fifty-base sequences were chosen that displayed high and low self-complementarity (Figure 4.11). The intercalation of **BALysTO1** into each of these strands was studied. As expected, the fluorescence emission from the high double-strand sequence was higher than that of the less self-complementary sequence, due to greater intercalation (Figure 4.12).

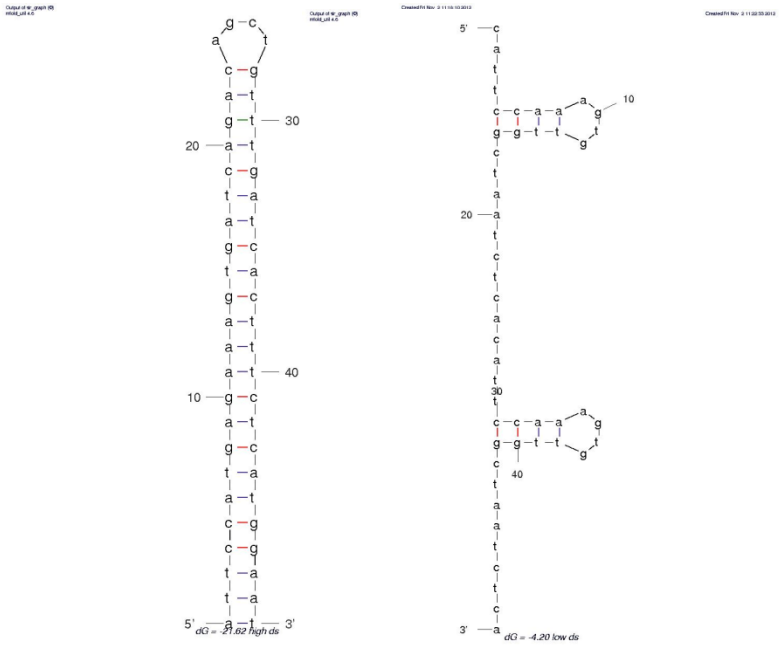


Figure 4.11 High double stranded and low double stranded 50-base ssDNA.

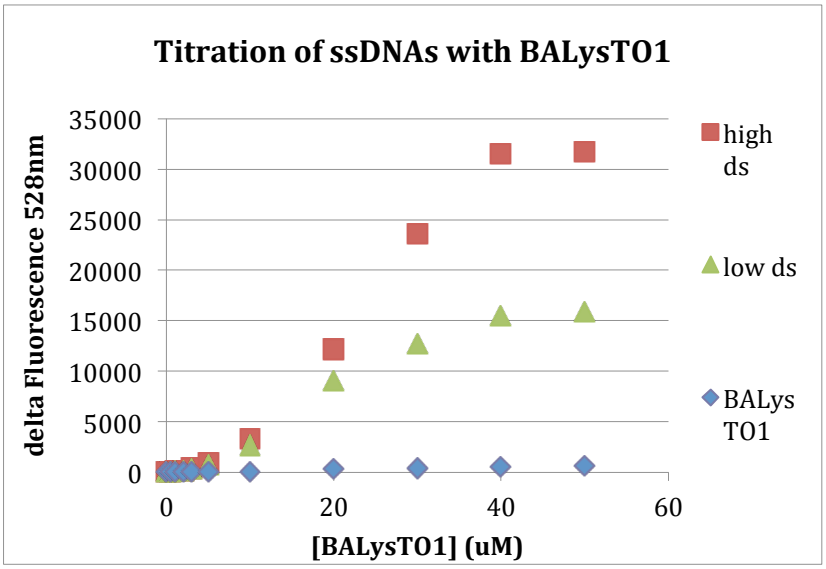


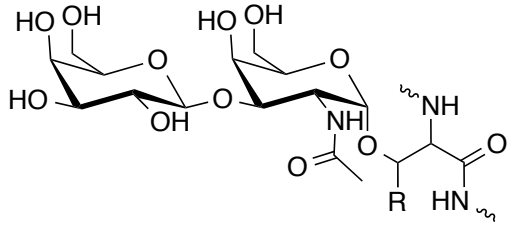
Figure 4.12 Titration of ssDNA sequences of high and low double stranded character with **BALysTO1** in DPBS, pH=7.4.

4.4.2 Glycosidase treatment of cells

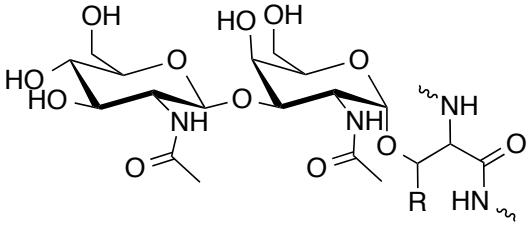
It was hypothesized that boronic acid binding to cell surfaces would be modulated by cell surface sugar expression. Glycosidases are enzymes that remove carbohydrates from glycoproteins. Two types of glycosidases exist that either remove oligosaccharides conjugated to proteins or lipids (endoglycosidases) or remove monosaccharides from a terminal residue or non-reduced end (exoglycosidases). *O*-glycosidases remove sugars that are linked to proteins through the hydroxyl of a threonine or serine residue, while *N*-glycosidases remove sugars that are linked via asparagine.

In order to test the effect of sugar density on boronic acid receptor binding, four enzymes were chosen to remove sugars from cell surface glycans: neuraminidase, *O*-glycosidase, PNGase F, and mannosidase. Neuraminidase is an exoglycosidase that has specificity for hydrolysis of terminal sialic acid residues that have α-linkages with any sugar hydroxyls at the 2-, 6-, or 8-position. Sialic acids are overexpressed on metastatic cancer cell surfaces, resulting in increased negative surface charge. They are the terminal residues on sialyl Lewis^x; this tetrasaccharide is overexpressed on cancer cells and found on both *N*- and *O*-linked glycans. *O*-glycosidase is an endoglycosidase that hydrolyzes core 1 and core 3 *O*-linked disaccharides, detaching polysaccharides from proteins by hydrolyzing the bond to a threonine or serine residue. Core 1 linked saccharides are attached to protein-linked *N*-acetylgalactosamine (GalNAc) via a β-1,3 linkage to galactose (Figure 4.13). Core 3 linked saccharides are attached to GalNAc via a β-1,3 linkage to *N*-acetylglucosamine (GlcNAc). *O*-glycosidase is used in conjuction with neuraminidase because sialic acid residues must be removed for the enzyme to function. PNGase F catalyzes the hydrolysis of *N*-linked glycans by cleaving between GlcNAc and asparagine (Figure 4.14). This enzyme can be used with high mannose, hybrid, and complex oligosaccharides. High mannose oligosaccharides have branching of additional

terminal mannose residues linked from the mannotriose. Hybrid oligosaccharides contain terminal mannose branches as well as further substituted branches from GlcNAc antennae attached to the α-1,3 and α-1,6 mannoses. Complex oligosaccharides do not contain terminal mannose residues but instead display a variety of monosaccharides branched from GlcNAc antennae; these oligosaccharides usually terminate with sialic acid residues. Mannosidase is an exoglycosidase that catalyzes hydrolysis of mannose on glycoproteins. The mannosidase used in these studies is capable of hydrolyzing α-1,2 and α-1,3 linked mannose on high mannose and hybrid oligosaccharides.



Core 1: Gal-β1,3-GalNAc-αS/T



R = H, Serine
R = CH_3, Threonine

Core 3: GlcNAc-β1,3–GalNAc-αS/T

Figure 4.13 Core 1 and core 3 linked residues targeted by *O*-glycosidase.

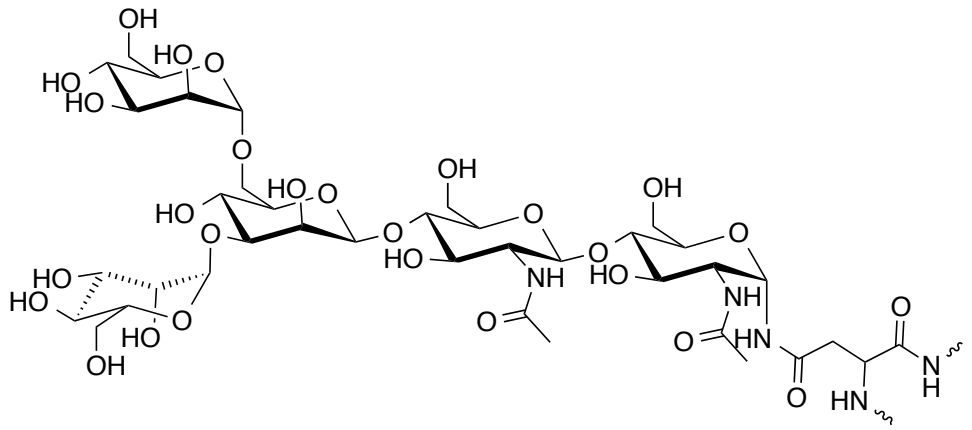


Figure 4.14 *N*-linked glycans contain a core trimannose linked to a GlcNAc dimer.

4.4.3 FACS response of cells to BALysTO1

Cells were treated with enzymes and then incubated with **BALysTO1** intercalated into high double-stranded (high ds) or low double-stranded (low ds) ssDNA. Fluorescence measurements were made on individual cells using flow cytometry and the population fluorescence plotted and compared for each enzyme treatment and each receptor incubation condition. The fluorescence response of the receptors changed the most when incubated with mannosidase-treated cells, when compared to their fluorescence with DPBS-treated cells, resulting in a fluorescence increase (Figure 4.15). The response of low ds ssDNA, high ds ssDNA, and DNA-free **BALysTO1** to the two cell lines was variable. The positively charged free dye exhibited the highest fluorescence, while the presence of DNA lead to decreased binding likely due to the negative charge of the strands. The low ds sequence showed a higher fluorescence than the high ds sequence for MOLT-4 and this response was switched for SK-OV-3. SK-OV-3 also showed a greater fluorescence increase with **BALysTO1**-based receptors than

MOLT-4. SK-OV-3 cells have been shown to exhibit a large amount of high-mannose *N*-glycans, implicating these structures as potential markers for ovarian cancer.[10]

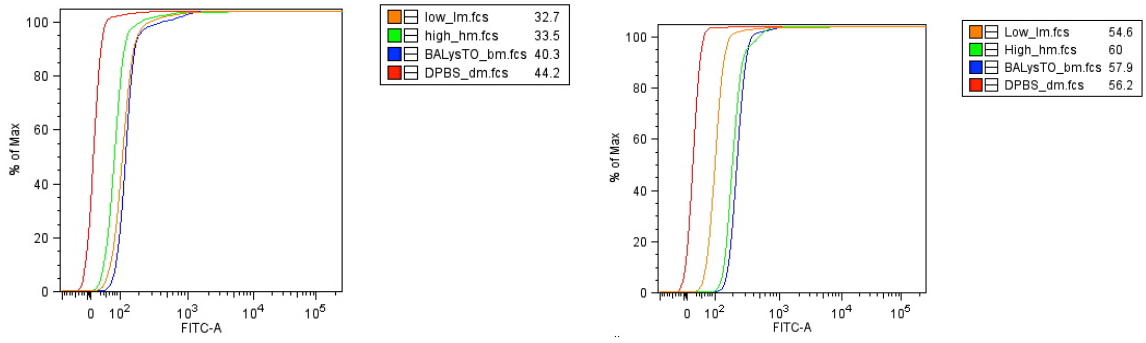


Figure 4.15 Cumulative density function (CDF) of fluorescence response of **BALysTO1** with MOLT-4 (left) and SK-OV-3 (right) cells treated with mannosidase.

Hydrolyzing native proteins on cell surfaces can require increased reaction times and enzyme concentrations. When MOLT-4 cells were treated with higher concentrations of the least responsive enzymes, neuraminidase, *O*-glycosidase, and PNGase F, new fluorescence responses were observed. For the high double-stranded DNA, neuraminidase gave a slight decrease in fluorescence compared to the DPBS treated sample (Figure 4.16) *O*-glycosidase, PNGase F, and mannosidase all resulted in fluorescence increases compared to non-enzyme treated cells. Increase in the concentration of PNGase F resulted in the death of all sample cells, which exhibit large emissions with **TO1** derivatives. The fluorescence increase of live cells was larger for mannosidase than that of *O*-glycosidase.

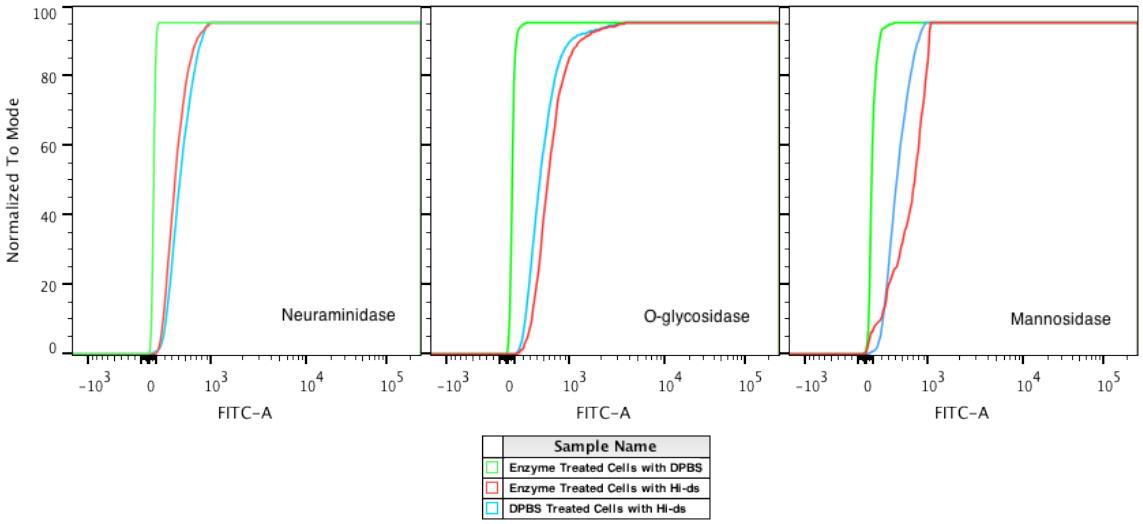


Figure 4.16 CDF plots of MOLT-4 cells treated with different enzymes and incubated with **BALysTO1** and high ds ssDNA.

Treatment of cells with neuraminidase removes sialic acid residues that lead to a negatively charged cell surface. This may be the reason why the positively charged **BALysTO1** exhibits a slight decrease in fluorescence when incubated with neuraminidase treated cells. However, this receptor is intercalated into a negatively charged DNA scaffold. It is therefore more likely that this boronic acid has affinity for sialic acid residues, which contain free 1,2 diols that can bind to boronic acid receptors. Removal of these sugars would result in lower binding and thus lower fluorescence. The negative change in fluorescence with this enzyme treatment group is small, and could therefore be insignificant; however, it was observed both with high ds and low ds ssDNAs.

Treatment with *O*-glycosidase should result in the removal of oligosaccharides to reveal a free serine or threonine hydroxyl. It was hypothesized that removal of sugars would lead to decreased boronic acid binding and decreased fluorescence. The increase in fluorescence when cells are treated with *O*-glycosidase does not support this hypothesis,

but instead indicates that removal of *O*-linked glycans increases the affinity of the receptor for cells. Removal of *O*-linked oligosaccharides could give the boronic acid access to cell surface features for which it has higher affinity.

Removal of α-1,2 and α-1,3 linked mannose from high-mannose and hybrid glycans was also hypothesized to result in decreased affinity of the boronic acid receptor for the cell surface and decreased fluorescence. The fluorescence increase observed could be due to easier access of the receptor to mannose residues as the enzyme should remove branching that may cause hindrance to boronate ester formation.

4.5 CONCLUSION

An array of commercially available boronic acids and indicators was tested for responsiveness to cell surface sugars. Boronic acid indicator complexes did not exhibit absorbance changes in the presence of cells. This is likely due to the weaker association of monoboronic acids with cell surface sugars in competitive media compared to the association with catechol indicators at the concentrations necessary to observe indicator absorbance.

Receptor assemblies of a boronic acid conjugated to a lysine derivative of fluorescent intercalator **TO1**, **BALysTO1**, were better at signaling binding to cells. Intercalation of **BALysTO1** into a 1 kb dsDNA was used to assemble a multivalent receptor for patterning cell surface glycocalyces. The fluorescence response with and without DNA was used to classify cell lines.

The response of two cell lines tested, SK-OV-3 and MOLT-4, to receptor assemblies composed of ssDNA and boronic acid-derived intercalators was shown to be differential in initial cytometry experiments. Although enzyme treatment did not result in

the hypothesized fluorescence responses based on theorized interactions of boronic acid-based receptors with cell surface glycans, the response of the receptors in initial experiments has the potential to be diagnostic of cell type. Using the boronic acid-based receptor assemblies, such glycocalyx patterning can be undertaken on whole cells instead of cell lysates. This gives the advantage of measuring cell surface properties while the cells are still alive. Further cytometry experiments will be undertaken on additional cell lines.

4.6 EXPERIMENTAL

4.6.1 IDA

Cells were grown as detailed in chapter 3. Cells in the growth phase (n=5-15) were detached, if necessary, and counted on a hemacytometer. Appropriate numbers of cells for each experiment were pelleted and their growth media removed. The cells were then washed with DPBS and again pelleted. After DPBS removal, cells were taken up in 10 mM HEPES with 0.1 M NaCl at pH = 7.4.

Boronic acids and indicators were procured from various sources (Sigma-Aldrich, Acros, TCI America) and used without further purification.

Absorbance measurements were made on a Cytation3 Microwell Plate Reader from BioTek measuring 300 – 700 nm at 10 nm intervals. Clear-bottomed black well plates were procured from Grenier (384-well) and Costar (96-well).

4.6.1.1 Boronic Acid and Indicator Titrations

Stock solutions of indicators were prepared in 10 mM HEPES 0.1 M NaCl at a pH = 7.4. These solutions were diluted in the same buffer to a concentration of 500 μM and used for titrations. Stock solutions of boronic acids were prepared at concentrations

greater than 2 mM in the same buffer as indicators, with DMSO added as needed for solubility. These solutions were diluted to 2 mM for use in titrations.

Increasing equivalents of boronic acid were added to indicator solutions at 250 μM (ARS), 200 μM (PV), or 75 μM (4ME) and absorbance values were recorded. Plotting the absorbance changes of the free indicator and the indicator-boronic acid complex allowed for determination of binding equivalents for each boronic acid-indicator pair. These equivalents were used for indicator displacement studies with cells.

Titrations with ARS

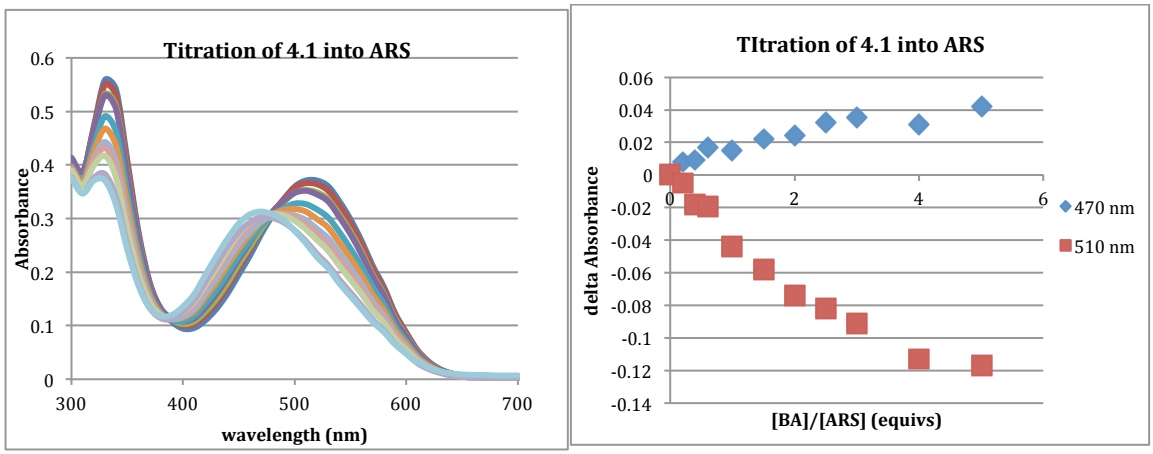


Figure 4.17 Titration of **4.1** into ARS in 10 mM HEPES 0.1 M NaCl pH = 7.4.

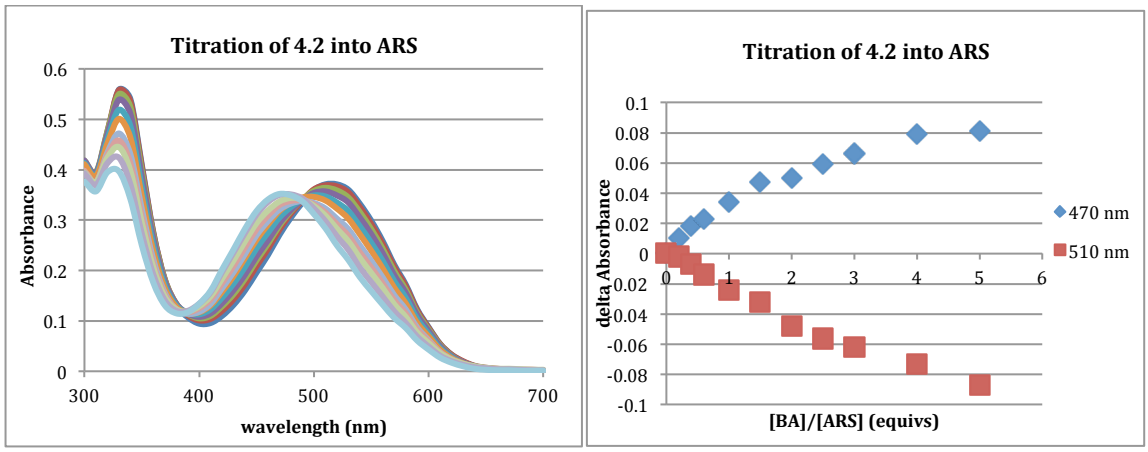


Figure 4.18 Titration of **4.2** into ARS in 10 mM HEPES 0.1 M NaCl pH = 7.4.

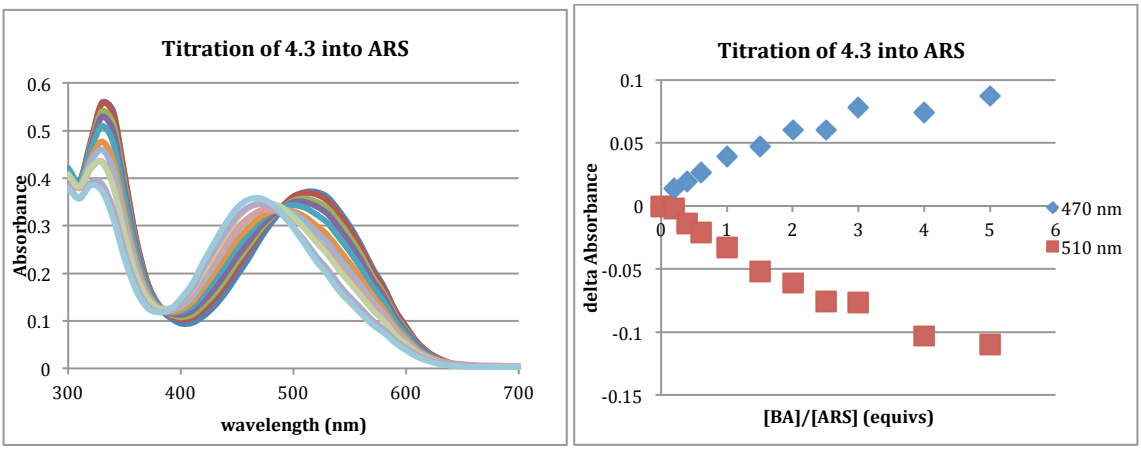


Figure 4.19 Titration of **4.3** into ARS in 10 mM HEPES 0.1 M NaCl pH = 7.4.

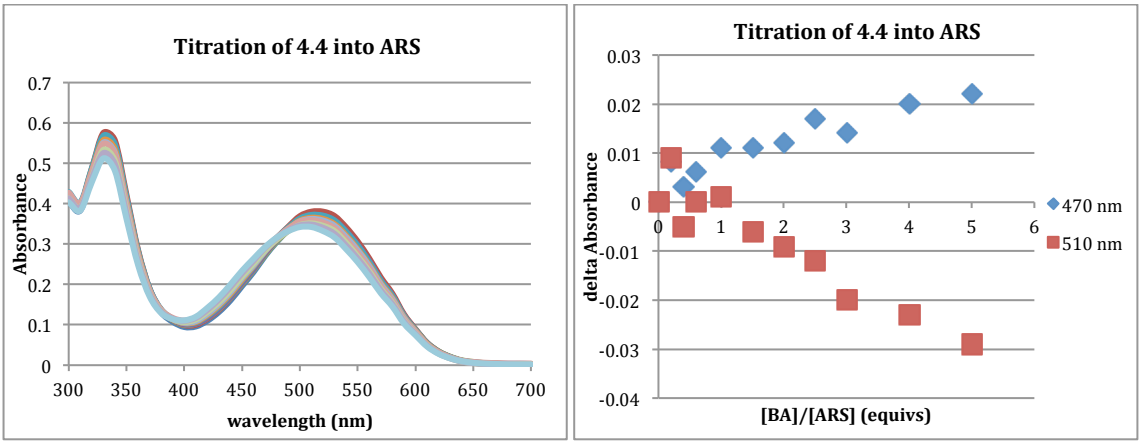


Figure 4.20 Titration of **4.4** into ARS in 10 mM HEPES 0.1 M NaCl pH = 7.4.

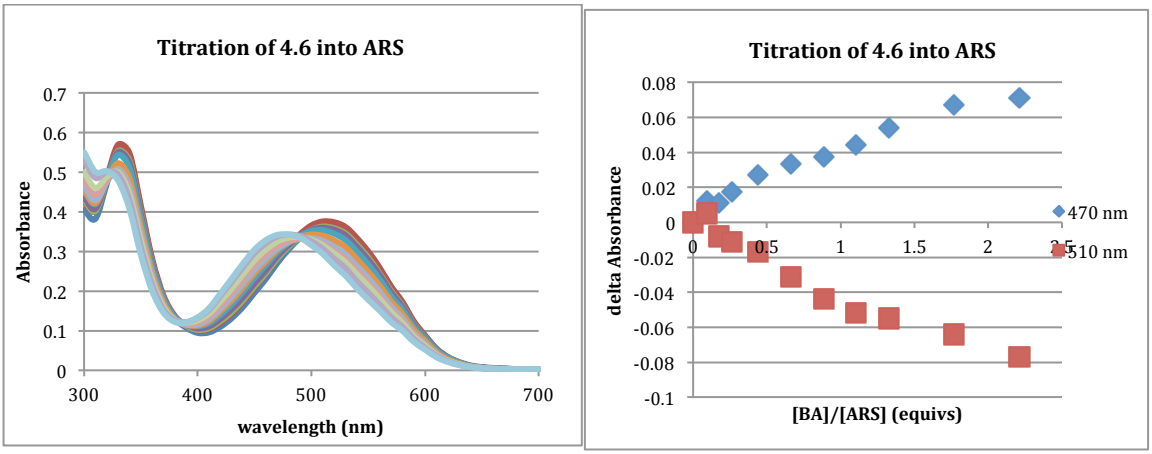


Figure 4.21 Titration of **4.6** into ARS in 10 mM HEPES 0.1 M NaCl pH = 7.4.

Titrations with PV

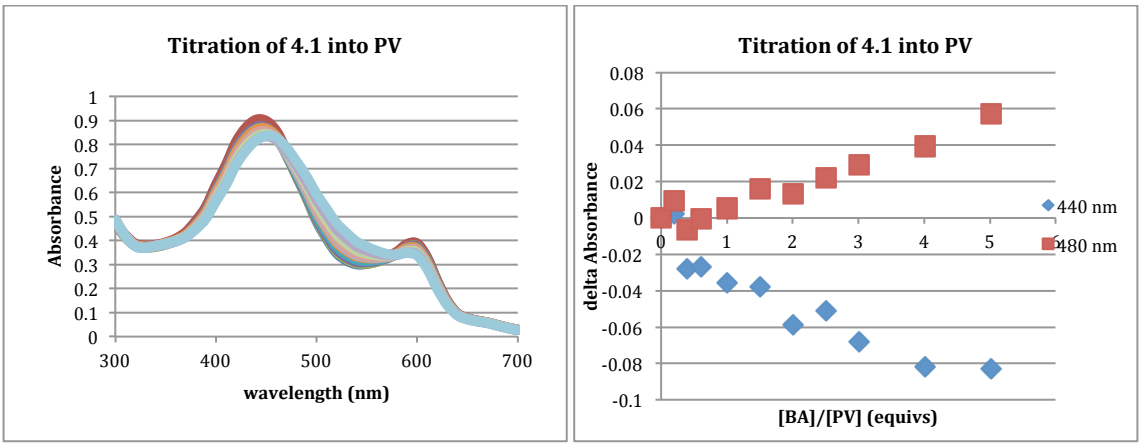


Figure 4.22 Titration of **4.1** into PV in 10 mM HEPES 0.1 M NaCl pH = 7.4.

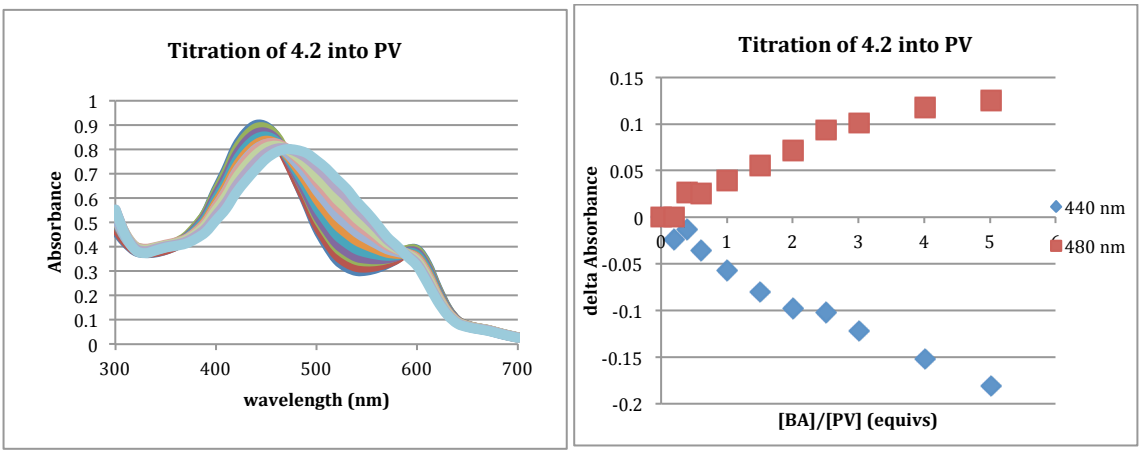


Figure 4.23 Titration of **4.2** into PV in 10 mM HEPES 0.1 M NaCl pH = 7.4.

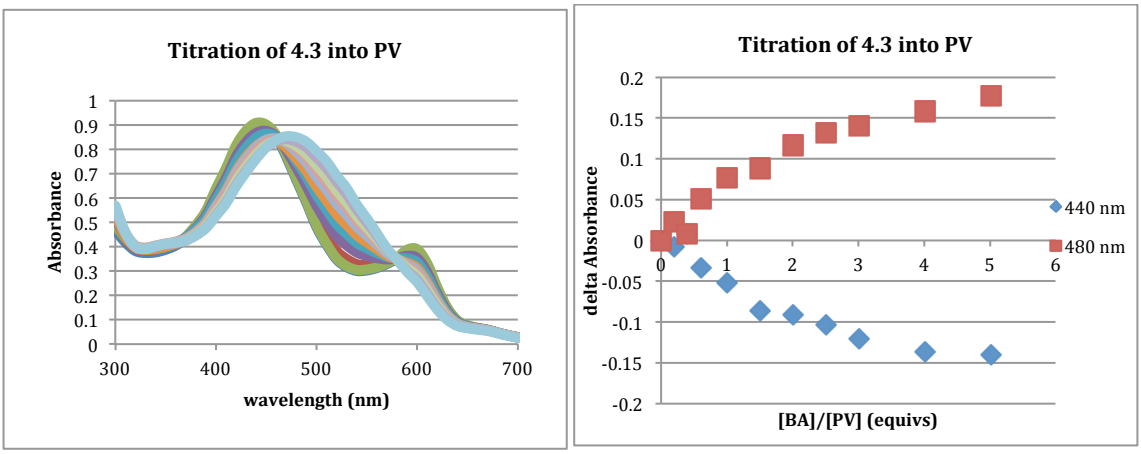


Figure 4.24 Titration of **4.3** into PV in 10 mM HEPES 0.1 M NaCl pH = 7.4.

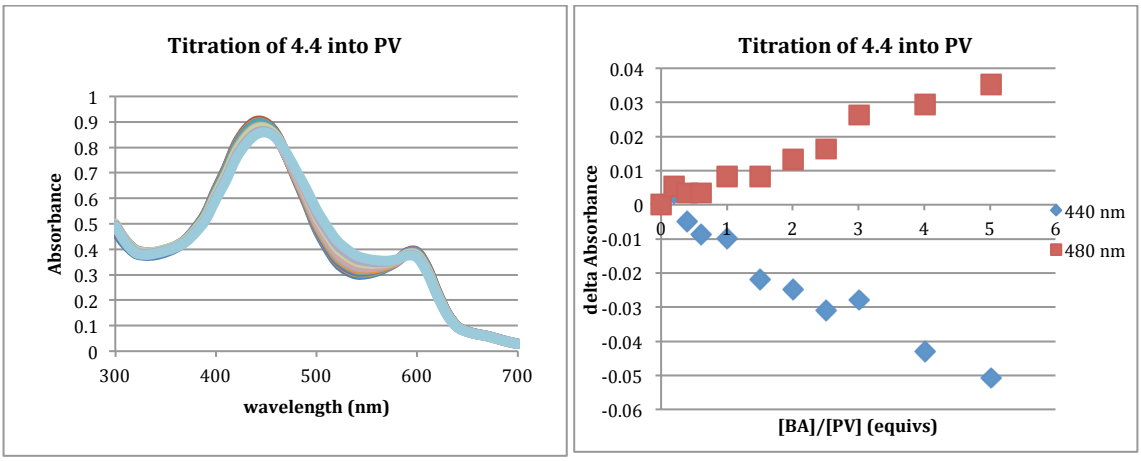


Figure 4.25 Titration of **4.4** into PV in 10 mM HEPES 0.1 M NaCl pH = 7.4.

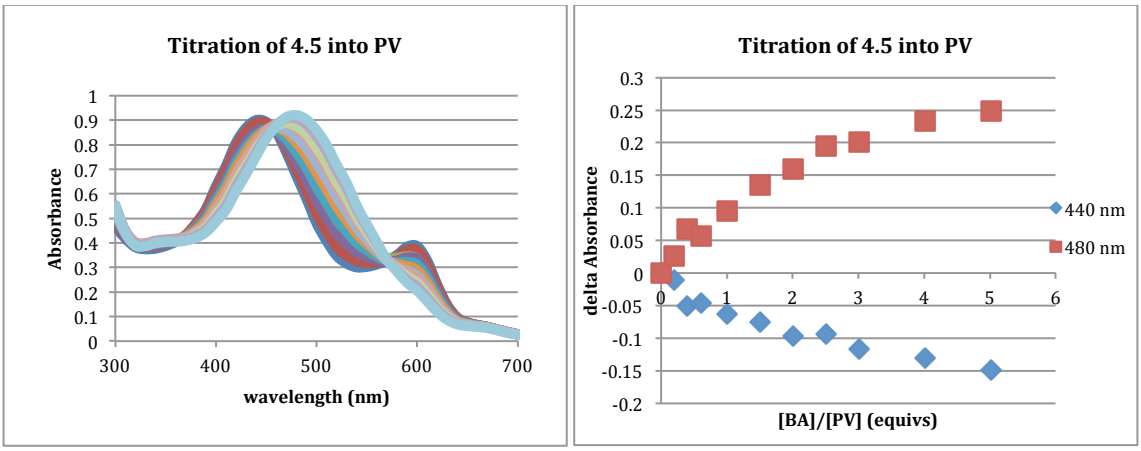


Figure 4.26 Titration of **4.5** into PV in 10 mM HEPES 0.1 M NaCl pH = 7.4.

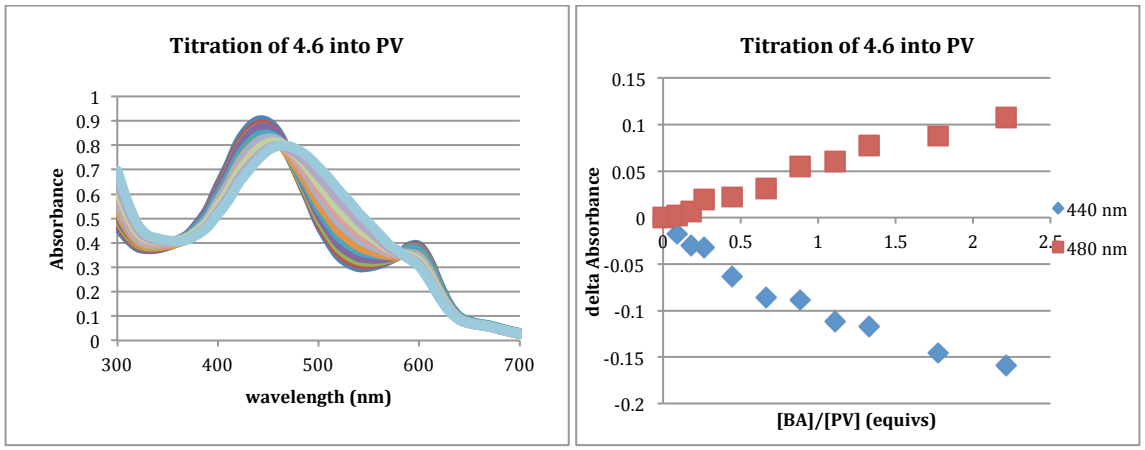


Figure 4.27 Titration of **4.6** into PV in 10 mM HEPES 0.1 M NaCl pH = 7.4.

*Curve Fitting of **4.5** with ARS*

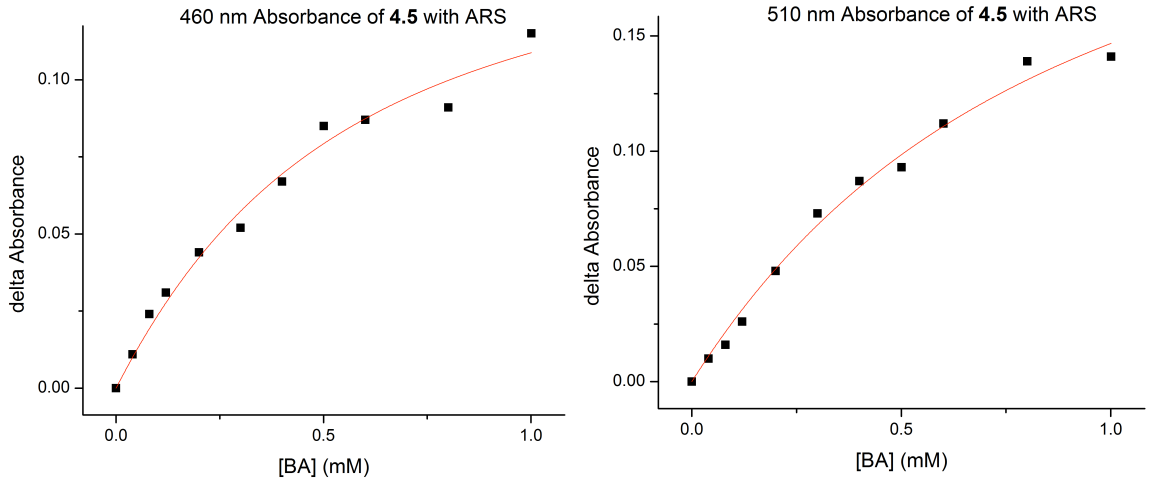
Titration data were expressed as delta absorbance values at 460 nm and 510 nm. The absorbance change at 510 nm was negative, so these values were multiplied by -1 in order to generate positive numbers. These values were plotted against the concentration of boronic acid in mM. Curve fitting was undertaken using OriginLab® 8.1 with a custom parameterized equation derived from 1:1 binding thermodynamics (equation 4.1).[11]

dA = de*(Rt*10^{-3})/(2/(K*(It*10^{-6})-1-K*(Rt*10^{-3})+((1-K*(It*10^{-6})+K*(Rt*10^{-3}))^2+4*K*(It*10^$^{-6}$))^0.5)+1) (equation 4.1)

In this equation, dA are the delta absorbance values (y-values), Rt are the host concentrations at a particular dA (x-values). The parameters that can be fit are the molar absorptivity (de), the guest concentration (It), and the binding constant (K). Any of these parameters can also be fixed, provided the values are known.

The indicator concentration was input in μM and not allowed to vary, while the indicator or indicator-boronic acid complex molar absorptivity and the binding constant

were allowed to freely vary. Convergence was reached when the Chi-squared tolerance was less than 1×10^{-9}. The curve fit and parameter results are shown in Figure 4.28.



	de	K	It	Adj. R-Square
460 nm	803±87	2431±631	200	0.98194
510 nm	1337±152	1367±298	200	0.99059

Figure 4.28 Iterative curve fitting of titration of ARS (200 μM) with **4.5** (0-1 mM) in 10mM HEPES 0.1 M NaCl pH=7.4 at 460 nm and 510 nm using a parameterized equation.

4.6.1.2 Titration of Boronic Acid-Indicator Complexes with Cells

Using the results of the boronic acid-indicator titrations, boronic acid-indicator complexes were prepared at approximately 90% saturation so that the cells and indicators could compete for boronic acid binding. Stock solutions of 400 uM ARS or PV with 1 mM boronic acids **3.1-3.6** were prepared in 10 mM HEPES 0.1 M NaCl pH=7.4. These solutions were added to 384-well plates and increasing volumes of a cell suspension solution (5000 cells/uL) were added. Each well was filled with HEPES buffer to a volume of 50 uL such that the final concentration of dye was 200 uM. Cells were added to boronic acid-free solutions for comparison.

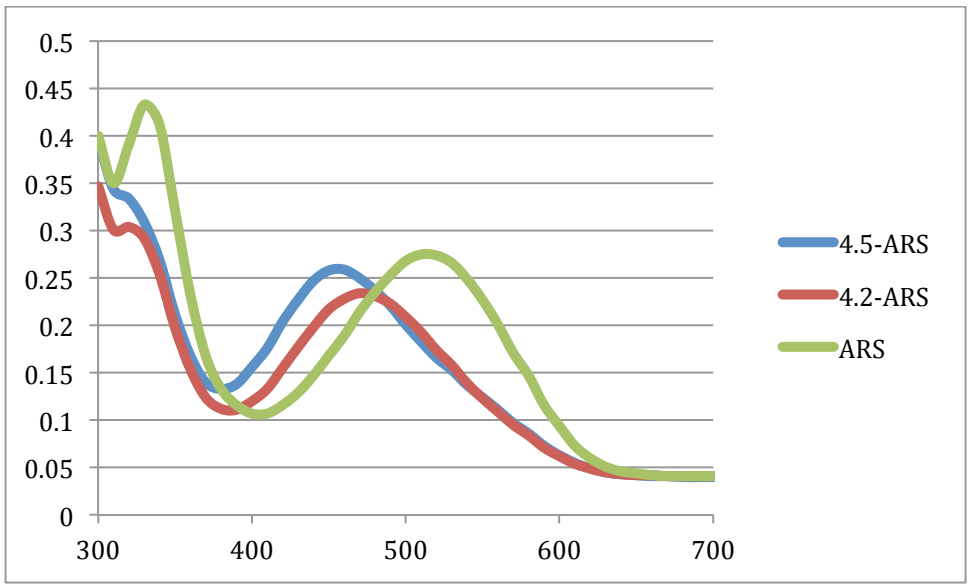


Figure 4.29 ARS and boronic acid-bound ARS without any cells.

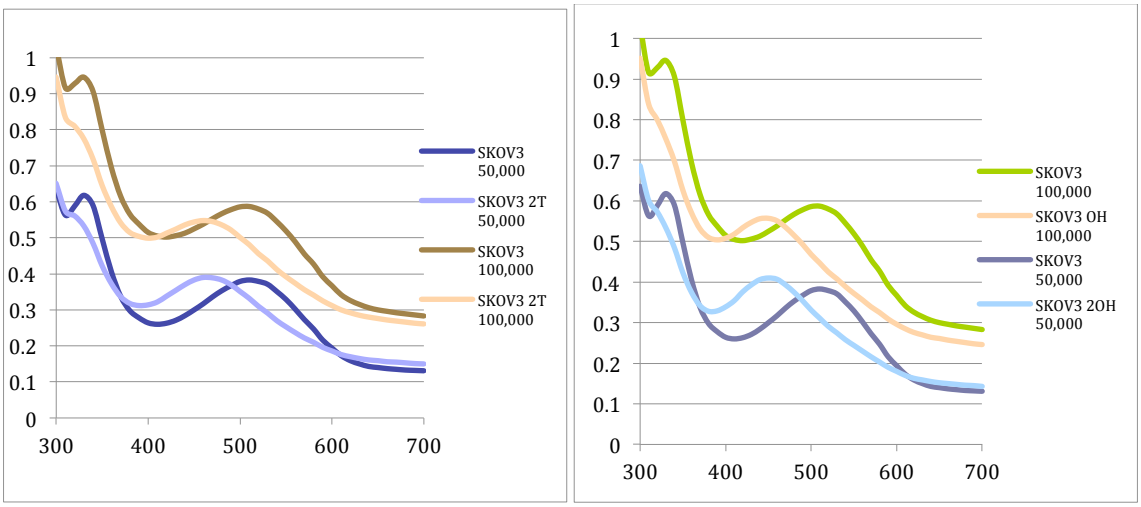


Figure 4.30 Responses of increasing SK-OV-3 to **4.2-ARS** (left) and **4.5-ARS** (right).

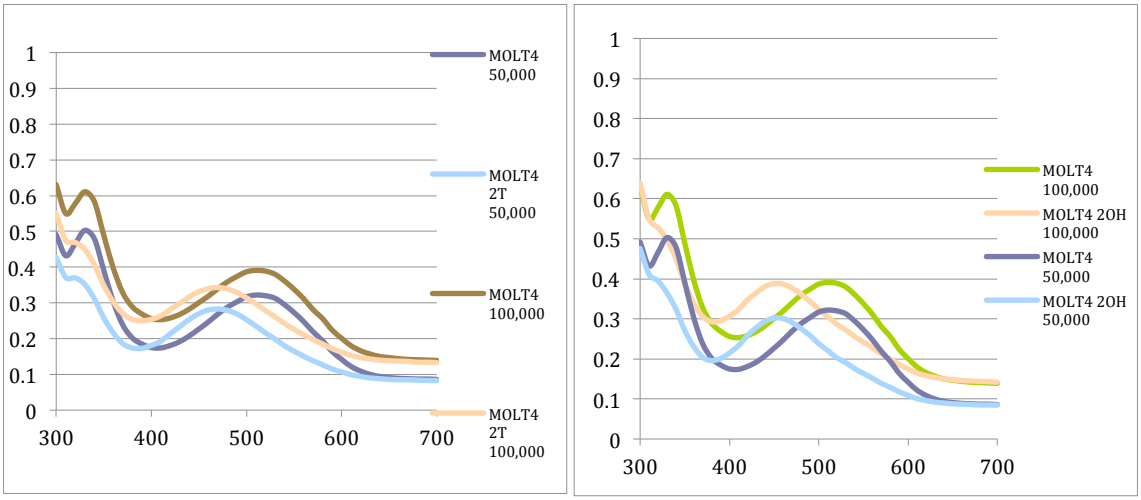


Figure 4.31 Responses of increasing MOLT-4 to **4.2-ARS** (left) and **4.5-ARS** (right).

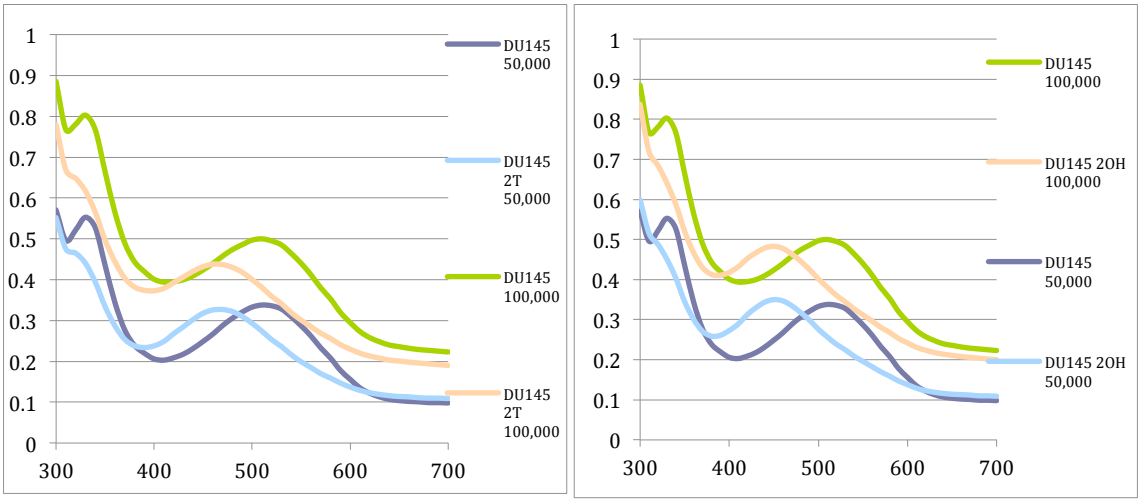


Figure 4.32 Responses of increasing DU-145 to **4.2-ARS** (left) and **4.5-ARS** (right).

4.6.1.3 Indicator Uptake of Cells with Boronic Acids

Initially, cell stocks were mixed with boronic acids and plated at a concentration of 1 mM boronic acid and 50,000 cells per well. A stock solution of indicator was added with increasing volume to each well and diluted to the same total volume (50 uL) with buffer. Indicator concentrations ranged from 0-250 uM. No difference was observed in the indicator-boronic acid peaks between cell-free and cell-added solutions. A dilute

indicator uptake was therefore prepared with 50,000 cells per well at a concentration of 100 uM BA. Indicator concentrations again ranged from 0-250 uM.

4.6.2 Fluorescence of BALysTO1

4.6.2.1 Synthesis

BALysTO1 was synthesized manually using an Fmoc Wang resin of lysine, protected with methyl trityl (Mtt). The resin, 282.8 mg, was swelled in DMF overnight. A 20% piperidine solution in DMF (5 mL) was added to the resin to deprotect the Fmoc. The resin was deprotected for 30 mins, then washed 3x DMF, 2x DCM, 2x MeOH (10 mL each). A coupling solution containing **TO1** (130.85 mg, 2.1 equivs), DIPEA (80.5 mg, 4.1 equivs), and PyBOP (165 mg, 2.1 equivs) in DMF was added to the deprotected resin and allowed to react overnight. The coupling solution was removed, the resin was washed as above, and the lysine Mtt protection group was removed by shaking the resin in a solution of 5% TFA and 5% TIPS (v/v) in DCM. The deprotection solution turned yellow after 10 minutes and was removed. Two more aliquots of deprotection solution were reacted in this manner. The deprotected resin was again washed and a coupling solution containing *o*-bromomethylboronic acid (130.68 mg, 4.1 equivs), DIPEA (78.75 mg, 3.99 equivs) in DMF was added and the resin was shaken overnight. The boronic acid coupling solution was removed, the resin was washed, and then dried overnight under high vacuum. A cleavage cocktail of 95% TFA, 2.5% TIPS, 2.5% H2O was used to cleave **BALysTO1** from the Wang resin by shaking for four hours. The liquid was separated from the resin, rotavaped to remove TFA, and **BALysTO1** was precipitated by adding cold ether.

Prep HPLC was performed on a Shimadzu instrument on a C-18 column. BALysTO1 was dissolved in MeOH and separated using a gradient of acetonitrile and

water as follows: 0-30 min, 5-75% MeCN in H_2O; 30-40 min 95% MeCN. The peak with absorbance at 500 nm was collected and analyzed using LC/MS. ES-APCI + showed a masses at m/z 621.8, 320.0, and 311.2. These correspond to expected m/z 621.3 $(M-H_2O)^+$, 320.1 $(M+H)^{2+}$, and 311.1 $(M-OH)^{2+}$.

4.6.2.2 Titration of BALysTO1 into 1 kb dsDNA

1 kb dsDNA was prepared as outlined in section 3.4. Increasing volumes of a stock solution of BALysTO1 (50 uM in DPBS) were added to 1 kb dsDNA and diluted with DPBS to a final volume of 100 uL in each well such that the dsDNA concentration was 4 nM. The BALysTO1 concentration varied from 0 – 30 uM.

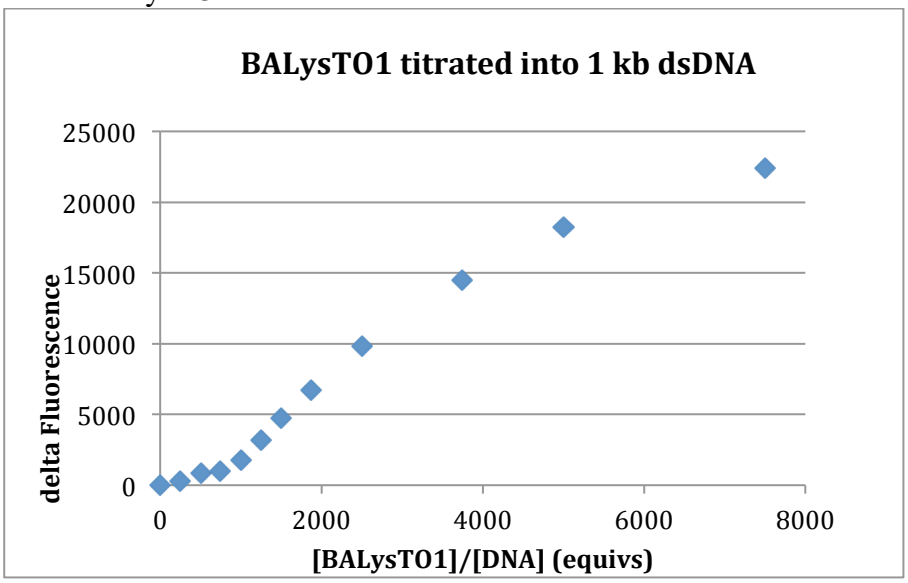


Figure 4.33 Titration of **BALysTO1** into 1 kb dsDNA (4 nM) in DPBS at pH = 7.4

4.6.3 FACS

4.6.3.1 Solution Preparation and Enzyme Digestion

Stock solutions of ssDNA were prepared by diluting oligonucleotides ordered from IDT to a concentration of 100 µM in ddH_2O. A stock solution of 357.85 µM BALysTO1 was prepared by diluting the solid amino acid conjugate with 2 mL DPBS.

Solutions of ssDNA and BALysTO1 were prepared using the stocks at 1 μM each in DPBS and incubated at 4°C overnight.

An aliquot of suspension cells was removed from the culture flask after aspiration and counted on a hemacytometer. Cells were stained with 2X trypan blue to identify viability. The hemacytometer provides four replicates of measurements in a 10 uL sample of cells by measuring the cells in 1 mm^2 and a volume of 0.1 uL. The amount of cells in 1 cm^3 (1 mL) is calculated after averaging the cell counts, multiplying for the dilution factor, and multiplying by 10,000. For adherent cells, T-75's grown to confluency were trypsinized with 2 mL 0.25% trypsin/EDTA. 3 mL appropriate growth media was added, the cell suspension aspirated, and an aliquot of cells counted on a hemacytometer. Appropriate volumes of cells were removed from suspension such that each enzyme treatment group would have 1×10^6 cells. Cells were pelleted at 175 × g and then taken into DPBS and split into 2 mL DNA LoBind eppendorf tubes such that each tube contained 1×10^6 cells. Cells were again pelleted, DPBS removed, and the appropriate amount of recommended enzyme buffer added to the solution, along with the following amounts of enzymes (Table 4.1):

Enzyme	**Low (U)**	**High (U)**
Neuraminidase	50	400
O-glycosidase	40,000	400,000
PNGase F	500	3,000
α-1-2,3-mannosidase	160	128

Table 4.1 Enzyme concentrations per 1×10^6 cells used in low enzyme and high enzyme experiments.

Solutions of cells and enzymes, along with a DPBS control solution, were incubated at 37°C for four hours. After incubation, cells were pelleted, taken up into DPBS and split into equal volumes for incubation with receptor. These aliquots were pelleted and 100 uL of **BALysTO1** receptor solutions were added. The receptor was incubated with cells for 30 min at 37°C. Each enzyme treatment group was also incubated with a DPBS control and a dead cell stain (LIVE/DEAD® Fixable Red Dead Cell Stain, Life Technologies). After incubation cells were pelleted and taken up into 500 uL DPBS, transferred to FACS tubes (BD Falcon 352052), and fluorescence values measured.

4.6.3.2 Flow Cytometry

Flow cytometry measurements were made using a BD LSRFortessa™ cell analyzer with laser excitation at 488 nm and emission recorded at 519 nm (FITC) and 695 nm (PerPE Cy 5.5) for **TO1** and dead cell stain, respectively. Flow cytometry solutions were prepared in DPBS without calcium and magnesium. A minimum of 1,000 cells was recorded per experiment, but most experiments used 5,000-10,000 cells per sample. Voltages for FSC, SSC, FITC, and PerPE Cy 5.5 were adjusted to place the sample population appropriately on the dot plot. FSC and SSC values had to be adjusted for each cell line. FITC was adjusted so that the DPBS-only treated sample had an emission at 0, while PerPE Cy 5.5 voltage was adjusted so that both dead and live cell populations appeared on the fluorescence dot plot.

4.7 REFERENCES

1. Paszek, M. J.; DuFort, C. C.; Rossier, O.; Bainer, R.; Mouw, J. K.; Godula, K.; Hudak, J. E.; Lakins, J. N.; Wijekoon, A. C.; Cassereau, L.; Rubashkin, M. G.; Magbanua, M. J.; Thorn, K. S.; Davidson, M. W.; Rugo, H. S.; Park, J. W.;

Hammer, D. A.; Giannone, G.; Bertozzi, C. R.; Weaver, V. M., *Nature* **2014,** *advance online publication*.

2. Varki, A.; Kannagi, R.; Toole, B. P., Glycosylation Changes in Cancer. . In *Essentials of Glycobiology*. , 2nd ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY: 2009.
3. James, T. D.; Sandanayake, K. R. A. S.; Shinkai, S., *Angewandte Chemie International Edition in English* **1996,** *35*, 1910-1922.
4. Sandanayake, K. R. A. S.; James, T. D.; Shinkai, S., *Chemistry Letters* **1995,** *24*, 503-504.
5. Springsteen, G.; Wang, B., *Tetrahedron* **2002,** *58*, 5291-5300.
6. Norrild, J. C.; Eggert, H., *Journal of the American Chemical Society* **1995,** *117*, 1479-1484.
7. Burnett, T. J.; Peebles, H. C.; Hageman, J. H., *Biochemical and Biophysical Research Communications* **1980,** *96*, 157-162.
8. Yang, W.; Gao, S.; Gao, X.; Karnati, V. V. R.; Ni, W.; Wang, B.; Hooks, W. B.; Carson, J.; Weston, B., *Bioorganic & Medicinal Chemistry Letters* **2002,** *12*, 2175-2177.
9. Dowlut, M.; Hall, D. G., *Journal of the American Chemical Society* **2006,** *128*, 4226-4227.
10. Machado, E.; Kandzia, S.; Carilho, R.; Altevogt, P.; Conradt, H. S.; Costa, J. l., *Glycobiology* **2011,** *21*, 376-386.
11. Hargrove, A. E.; Zhong, Z.; Sessler, J. L.; Anslyn, E. V., *New J Chem* **2010,** *34*, 348-354.

Bibliography

Nat Biotech **2000,** *18*, IT50 - IT52.

Introduction. In *Introduction to Multivariate Statistical Analysis in Chemometrics*, CRC Press: 2009.

Principal Component Analysis. In *Introduction to Multivariate Statistical Analysis in Chemometrics*, CRC Press: 2009.

Classification. In *Introduction to Multivariate Statistical Analysis in Chemometrics*, CRC Press: 2009.

Albert, K. J.; Lewis, N. S.; Schauer, C. L.; Sotzing, G. A.; Stitzel, S. E.; Vaid, T. P.; Walt, D. R., *Chemical Reviews* **2000,** *100*, 2595-2626.

Alfonso, I.; Dietrich, B.; Rebolledo, F.; Gotor, V.; Lehn, J. M., *Helv Chim Acta* **2001,** *84*, 280-295.

Ambrosi, G.; Ciattini, S.; Formica, M.; Fusi, V.; Giorgi, L.; Macedi, E.; Micheloni, M.; Paoli, P.; Rossi, P.; Zappia, G., *Chemical Communications* **2009**, 7039-7041.

Anslyn, E. V.; Dougherty, D. A., *Modern Physical Organic Chemistry*. 2nd ed.; University Science: Sausalito, CA, 2006.

Arteaga, C. L.; Engelman, J. A., *Cancer Cell* **2014,** *25*, 282-303.

Arter, J. A.; Diaz, J. E.; Donavan, K. C.; Yuan, T.; Penner, R. M.; Weiss, G. A., *Analytical Chemistry* **2012,** *84*, 2776-2783.

Bajaj, A.; Miranda, O. R.; Kim, I. B.; Phillips, R. L.; Jerry, D. J.; Bunz, U. H. F.; Rotello, V. M., *P Natl Acad Sci USA* **2009,** *106*, 10912-10916.

Bajaj, A.; Miranda, O. R.; Phillips, R.; Kim, I. B.; Jerry, D. J.; Bunz, U. H. F.; Rotello, V. M., *Journal of the American Chemical Society* **2010,** *132*, 1018-1022.

Bajaj, A.; Rana, S.; Miranda, O. R.; Yawe, J. C.; Jerry, D. J.; Bunz, U. H. F.; Rotello, V. M., *Chemical Science* **2010,** *1*, 134-138.

Beer, P. D.; Gale, P. A., *Angewandte Chemie International Edition* **2001,** *40*, 486-516.

Berger, M.; Schmidtchen, F. P., *Journal of the American Chemical Society* **1996,** *118*, 8947-8948.

Berger, M.; Schmidtchen, F. P., *Angew Chem Int Edit* **1998,** *37*, 2694-2696.

Bernert, B.; Porsch, H.; Heldin, P., *J Biol Chem* **2011,** *286*, 42349-59.

Blondeau, P.; Segura, M.; Perez-Fernandez, R.; de Mendoza, J., *Chem Soc Rev* **2007,** *36*, 198-210.

Bonnet, C. S.; Devocelle, M.; Gunnlaugsson, T., *Organic & Biomolecular Chemistry* **2012,** *10*, 126-133.

Borron, P. J.; Crouch, E. C.; Lewis, J. F.; Wright, J. R.; Possmayer, F.; Fraher, L. J., *The Journal of Immunology* **1998,** *161*, 4599-4603.

Bridgham, J. T.; Carroll, S. M.; Thornton, J. W., *Science* **2006,** *312*, 97-101.

Bruzzoniti, M. C.; Sarzanini, C.; Mentasti, E., *Journal of Chromatography A* **2000,** *902*, 289-309.

Bunka, D. H. J.; Stockley, P. G., *Nat Rev Micro* **2006,** *4*, 588-596.

Burnett, T. J.; Peebles, H. C.; Hageman, J. H., *Biochemical and Biophysical Research Communications* **1980,** *96*, 157-162.

Butler, C.; Goetz, S.; Fitchett, C. M.; Kruger, P. E.; Gunnlaugsson, T., *Inorganic Chemistry* **2011,** *50*, 2723-2725.

Cahill, D. J., *Journal of Immunological Methods* **2001,** *250*, 81-91.

Caltagirone, C.; Gale, P. A., *Chemical Society Reviews* **2009,** *38*, 520-563.

Cannistra, S. A.; Ottensmeier, C.; Niloff, J.; Orta, B.; DiCarlo, J., *Gynecologic Oncology* **1995,** *58*, 216-225.

Casey, R.; Skubitz, A. N., *Clin Exp Metastasis* **2000,** *18*, 67-75.

Chen, Z. H.; He, Y. B.; Hu, C. G.; Huang, X. H., *Tetrahedron-Asymmetr* **2008,** *19*, 2051-2057.

Chow, G.; Tauler, J.; Mulshine, J. L., *Journal of Biomedicine and Biotechnology* **2010,** *2010*, 11.

Collins, B. E. A kinetic investigation of boronic acid/diol interactions and pattern-based recognition of [alpha]-chiral carboxylates. PhD. Dissertation, The University of Texas at Austin, 2010.

De, M.; Rana, S.; Akpinar, H.; Miranda, O. R.; Arvizo, R. R.; Bunz, U. H. F.; Rotello, V. M., *Nat Chem* **2009,** *1*, 461-465.

Dixon, S. J.; Brereton, R. G., *Chemometrics and Intelligent Laboratory Systems* **2009,** *95*, 1-17.

dos Santos, C. M. G.; Gunnlaugsson, T., *Dalton Transactions* **2009**, 4712-4721.

dos Santos, C. M. G.; Gunnlaugsson, T., *Supramol Chem* **2009,** *21*, 173-180.

Dowlut, M.; Hall, D. G., *Journal of the American Chemical Society* **2006,** *128*, 4226-4227.

Dudak, F. C.; Boyaci, I. H.; Orner, B. P., *Molecules* **2011,** *16*, 774-789.

Echavarren, A.; Galan, A.; Lehn, J. M.; Demendoza, J., *Journal of the American Chemical Society* **1989,** *111*, 4994-4995.

El Sheikh, S. S.; Domin, J.; Abel, P.; Stamp, G.; Lalani el, N., *Neoplasia* **2004,** *6*, 846-53.

Ellington, A. D.; Szostak, J. W., *Nature* **1990,** *346*, 818-22.

Esipenko, N. A.; Koutnik, P.; Minami, T.; Mosca, L.; Lynch, V. M.; Zyryanov, G. V.; Anzenbacher, P., *Chemical Science* **2013,** *4*, 3617-3623.

Esteban-Barragán, M. A.; Ávila, P.; Álvarez-Tejado, M.; Gutiérrez, M. D.; García-Pardo, Á.; Sánchez-Madrid, F.; Landázuri, M. O., *Cancer Research* **2002,** *62*, 2929-2936.

Fabbrizzi, L.; Poggi, A., *Chemical Society Reviews* **1995,** *24*, 197-202.

Folmer-Andersen, J. F.; Kitamura, M.; Anslyn, E. V., *Journal of the American Chemical Society* **2006,** *128*, 5652-5653.

Folmer-Andersen, J. F.; Lynch, V. M.; Anslyn, E. V., *Journal of the American Chemical Society* **2005,** *127*, 7986-7987.

Friedman, H. S.; Bigner, D. D., *New England Journal of Medicine* **2005,** *353*, 1997-1999.

Galan, A.; Andreu, D.; Echavarren, A. M.; Prados, P.; Demendoza, J., *Journal of the American Chemical Society* **1992,** *114*, 1511-1512.

Gale, P. A.; Sessler, J. L.; Kral, V., *Chemical Communications* **1998**, 1-8.

Gale, P. A. a. H., Cally J. E., Anion Receptors Containing Heterocyclic Rings. In *Supramolecular Chemistry: from Molecules to Nanomaterials*, First ed.; Steed, P. A. G. a. J. W., Ed. John Wiley & Sons, Ltd: West Sussex, 2012; Vol. 3, pp 1125-1151.

Garrett, R.; Grisham, C. M., *Biochemistry*. 2nd ed.; Saunders College Pub.: Fort Worth, 1999.

Garribba, E.; Micera, G.; Sanna, D.; Strinna-Erre, L., The Cu(II)-2,2'-bipyridine system revisited. *Inorg Chim Acta* **2000,** *299*, 253-261.

Geysen, H. M.; Meloen, R. H.; Barteling, S. J., *Proceedings of the National Academy of Sciences* **1984,** *81*, 3998-4002.

Gleich, A.; Schmidtchen, F. P.; Mikulcik, P.; Muller, G., *J Chem Soc Chem Comm* **1990**, 55-58.

Gresham, D.; Dunham, M. J.; Botstein, D., *Nat Rev Genet* **2008,** *9*, 291-302.

Gulgas, C. G.; Reineke, T. M., *Inorganic Chemistry* **2008,** *47*, 1548-1559.

Gumbiner, B. M., *Cell* **1996,** *84*, 345-357.

Hamula, C. L.; Zhang, H.; Guan, L. L.; Li, X. F.; Le, X. C., *Anal Chem* **2008,** *80*, 7812-9.

Hansen, L. F.; Jensen, L. K.; Jacobsen, J. P., *Nucleic Acids Research* **1996,** *24*, 859-867.

Hargrove, A. E.; Zhong, Z.; Sessler, J. L.; Anslyn, E. V., *New J Chem* **2010,** *34*, 348-354.

Healy, J. M.; Murayama, O.; Maeda, T.; Yoshino, K.; Sekiguchi, K.; Kikuchi, M., *Biochemistry* **1995,** *34*, 3948-3955.

Higa, K.; Shimmura, S.; Shimazaki, J.; Tsubota, K., *Cornea* **2005,** *24*, 206-12.

Hoffmann, E. K., *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes* **1986,** *864*, 1-31.

Houghten, R. A., *Proceedings of the National Academy of Sciences* **1985,** *82*, 5131-5135.

Hua, W.; Christianson, T.; Rougeot, C.; Rochefort, H.; Clinton, G. M., *J Steroid Biochem Mol Biol* **1995,** *55*, 279-89.

Hummer, G.; Pratt, L. R.; García, A. E., *The Journal of Physical Chemistry* **1996,** *100*, 1206-1215.

Hung, M. C.; Zhang, X.; Yan, D. H.; Zhang, H. Z.; He, G. P.; Zhang, T. Q.; Shi, D. R., *Cancer Lett* **1992,** *61*, 95-103.

Itano, N.; Sawai, T.; Miyaishi, O.; Kimata, K., *Cancer Research* **1999,** *59*, 2499-2504.

Iyer, S.; Gaikwad, R. M.; Subba-Rao, V.; Woodworth, C. D.; Sokolov, I., *Nat Nanotechnol* **2009,** *4*, 389-393.

Jadhav, V. D.; Herdtweck, E.; Schmidtchen, F. P., *Chemistry* **2008,** *14*, 6098-107.

Jadhav, V. D.; Schmidtchen, F. P., *Org Lett* **2005,** *7*, 3311-4.

Jadhav, V. D.; Schmidtchen, F. P., *Org Lett* **2006,** *8*, 2329-32.

Jadhav, V. D.; Schmidtchen, F. P., *J Org Chem* **2008,** *73*, 1077-1087.

James, T. D.; Sandanayake, K. R. A. S.; Shinkai, S., *Angewandte Chemie International Edition in English* **1996,** *35*, 1910-1922.

Jones, J. T.; Akita, R. W.; Sliwkowski, M. X., *FEBS Letters* **1999,** *447*, 227-231.

Joyce, L. A.; Maynor, M. S.; Dragna, J. M.; da Cruz, G. M.; Lynch, V. M.; Canary, J. W.; Anslyn, E. V., *Journal of the American Chemical Society* **2011,** *133*, 13746-13752.

Kane, R. S., *Langmuir* **2010,** *26*, 8636-8640.

Kano, K.; Kamo, H.; Negi, S.; Kitae, T.; Takaoka, R.; Yamaguchi, M.; Okubo, H.; Hirama, M., *Journal of the Chemical Society, Perkin Transactions 2* **1999**, 15-22.

Kay, B. K.; Kurakin, A. V.; Hyde-deruyscher, R., *Drug Discovery Today* **1998,** *3*, 370-378.

Kobiro, K.; Inoue, Y., *Journal of the American Chemical Society* **2003,** *125*, 421-427.

Koivunen, E.; Gay, D. A.; Ruoslahti, E., *Journal of Biological Chemistry* **1993,** *268*, 20205-20210.

Kolonin, M. G.; Bover, L.; Sun, J.; Zurita, A. J.; Do, K.-A.; Lahdenranta, J.; Card√≥-Vila, M.; Giordano, R. J.; Jaalouk, D. E.; Ozawa, M. G.; Moya, C. A.; Souza, G. R.; Staquicini, F. I.; Kunyiasu, A.; Scudiero, D. A.; Holbeck, S. L.; Sausville, E. A.; Arap, W.; Pasqualini, R., *Cancer Research* **2006,** *66*, 34-40.

Koltai, H.; Weingarten-Baror, C., *Nucleic Acids Research* **2008,** *36*, 2395-2405.

Kotova, O.; Kitchen, J. A.; Lincheneau, C.; Peacock, R. D.; Gunnlaugsson, T., *Chemistry – A European Journal* **2013,** *19*, 16181-16186.

Krishnamurthy, V. M.; Estroff, L. A.; Whitesides, G. M., Multivalency in Ligand Design. In *Fragment-based Approaches in Drug Discovery*, Wiley-VCH Verlag GmbH & Co. KGaA: 2006; pp 11-53.

Lavigne, J. J.; Anslyn, E. V., *Angewandte Chemie International Edition* **2001,** *40*, 3118-3130.

Leung, D.; Folmer-Andersen, J. F.; Lynch, V. M.; Anslyn, E. V., *Journal of the American Chemical Society* **2008,** *130*, 12318-12327.

Lichtenthaler, F. W., *Angewandte Chemie International Edition in English* **1995,** *33*, 2364-2374.

Lieber, M.; Todaro, G.; Smith, B.; Szakal, A.; Nelson-Rees, W., *International Journal of Cancer* **1976,** *17*, 62-70.

Lincheneau, C.; Duke, R. M.; Gunnlaugsson, T., *Organic & Biomolecular Chemistry* **2012,** *10*, 6069-6073.

Lincheneau, C.; Leonard, J. P.; McCabe, T.; Gunnlaugsson, T., *Chemical Communications* **2011,** *47*, 7119-7121.

Lorsch, J. R.; Szostak, J. W., *Biochemistry* **1994,** *33*, 973-982.

Lotan, R.; Raz, A., *Annals of the New York Academy of Sciences* **1988,** *551*, 385-398.

Machado, E.; Kandzia, S.; Carilho, R.; Altevogt, P.; Conradt, H. S.; Costa, J. l., *Glycobiology* **2011,** *21*, 376-386.

Mammen, M.; Choi, S. K.; Whitesides, G. M., *Angew Chem Int Edit* **1998,** *37*, 2755-2794.

Md. Alamgir Hossain, R. A. B., Victor W. Day, Kristin Bowman-James, Amide and Urea-Based Receptors. In *Supramolecular Chemistry: from Molecules to Nanomaterials*, First ed.; Steed, P. A. G. a. J. W., Ed. John Wiley & Sons, Ltd: West Sussex, 2012; Vol. 3, pp 1153-1178.

Melenhorst, W. B. W. H.; Mulder, G. M.; Xi, Q.; Hoenderop, J. G. J.; Kimura, K.; Eguchi, S.; van Goor, H., *Hypertension* **2008,** *52*, 987-993.

Metzger, A.; Anslyn, E. V., *Angewandte Chemie International Edition* **1998,** *37*, 649-652.

Mierke, C. T.; Frey, B.; Fellner, M.; Herrmann, M.; Fabry, B., *Journal of Cell Science* **2011,** *124*, 369-383.

Miranda, O. R.; Chen, H. T.; You, C. C.; Mortenson, D. E.; Yang, X. C.; Bunz, U. H. F.; Rotello, V. M., *Journal of the American Chemical Society* **2010,** *132*, 5285-5289.

Miyaji, H.; Hong, S.-J.; Jeong, S.-D.; Yoon, D.-W.; Na, H.-K.; Hong, J.; Ham, S.; Sessler, J. L.; Lee, C.-H., *Angewandte Chemie International Edition* **2007,** *46*, 2508-2511.

Mulder, A.; Huskens, J.; Reinhoudt, D. N., *Organic & Biomolecular Chemistry* **2004,** *2*, 3409-3424.

Müller, G.; Riede, J.; Schmidtchen, F. P., *Angewandte Chemie International Edition in English* **1988,** *27*, 1516-1518.

Muller, G.; Riehl, J., *J Fluoresc* **2005,** *15*, 553-558.

Nelson, G.; Chandrashekar, J.; Hoon, M. A.; Feng, L.; Zhao, G.; Ryba, N. J. P.; Zuker, C. S., *Nature* **2002,** *416*, 199-202.

Ng, E. W. M.; Shima, D. T.; Calias, P.; Cunningham, E. T.; Guyer, D. R.; Adamis, A. P., *Nat Rev Drug Discov* **2006,** *5*, 123-132.

Norrild, J. C.; Eggert, H., *Journal of the American Chemical Society* **1995,** *117*, 1479-1484.

Osborne, S. E.; Ellington, A. D., *Chemical Reviews* **1997,** *97*, 349-370.

Oyanagi, J.; Ogawa, T.; Sato, H.; Higashi, S.; Miyazaki, K., *PLoS ONE* **2012,** *7*, e53209.

Palacios, M. A.; Nishiyabu, R.; Marquez, M.; Anzenbacher, P., *Journal of the American Chemical Society* **2007,** *129*, 7538-7544.

Paszek, M. J.; DuFort, C. C.; Rossier, O.; Bainer, R.; Mouw, J. K.; Godula, K.; Hudak, J. E.; Lakins, J. N.; Wijekoon, A. C.; Cassereau, L.; Rubashkin, M. G.; Magbanua, M. J.; Thorn, K. S.; Davidson, M. W.; Rugo, H. S.; Park, J. W.; Hammer, D. A.; Giannone, G.; Bertozzi, C. R.; Weaver, V. M., *Nature* **2014,** *advance online publication*.

Penno, M. B.; August, J. T.; Baylin, S. B.; Mabry, M.; Linnoila, R. I.; Lee, V. S.; Croteau, D.; Yang, X. L.; Rosada, C., *Cancer Research* **1994,** *54*, 1381-1387.

Phillips, R. L.; Miranda, O. R.; You, C. C.; Rotello, V. M.; Bunz, U. H. F., *Angew Chem Int Edit* **2008,** *47*, 2590-2594.

Pirinen, R.; Tammi, R.; Tammi, M.; Hirvikoski, P.; Parkkinen, J. J.; Johansson, R.; Böhm, J.; Hollmén, S.; Kosma, V.-M., *International Journal of Cancer* **2001,** *95*, 12-17.

Polak, E. H.; Fombon, A. M.; Tilquin, C.; Punter, P. H., *Behavioural Brain Research* **1989,** *31*, 199-206.

Pramanik, D.; Majeti, B. K.; Mondal, G.; Karmali, P. P.; Sistla, R.; Ramprasad, O. G.; Srinivas, G.; Pande, G.; Chaudhuri, A., *Journal of Medicinal Chemistry* **2008,** *51*, 7298-7302.

Qing, G. Y.; He, Y. B.; Wang, F.; Qin, H. J.; Hu, C. G.; Yang, X., *Eur J Org Chem* **2007**, 1768-1778.

Rachwal, W. J.; Bongiorno, P. F.; Orringer, M. B.; Whyte, R. I.; Ethier, S. P.; Beer, D. G., *Br J Cancer* **1995,** *72*, 56-64.

Rahman, S.; Aitken A Fau - Flynn, G.; Flynn G Fau - Formstone, C.; Formstone C Fau - Savidge, G. F.; Savidge, G. F., *Biochemical Journal* **1998,** *335*, 247-257.

Richardson, F. S., *Chemical Reviews* **1982,** *82*, 541-552.

RoyChowdhury, A.; Ghosh, P.; Saha, S. K.; Mitra, P.; Banerjee, P., *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2014,** *124*, 492-499.

Ruoslahti, E., *Annual Review of Cell and Developmental Biology* **1996,** *12*, 697-715.

Sanchez, E. E.; Rodriguez-Acosta, A.; Palomar, R.; Lucena, S.; Bashir, S.; Soto, J. G.; Perez, J. C., *Arch Toxicol* **2009,** *83*, 271-279.

Sandanayake, K. R. A. S.; James, T. D.; Shinkai, S., *Chemistry Letters* **1995,** *24*, 503-504.

Schmuck, C., *Chem-Eur J* **2000,** *6*, 709-718.

Schneider, P. M.; Hung, M.-C.; Chiocca, S. M.; Manning, J.; Zhao, X.; Fang, K.; Roth, J. A., *Cancer Research* **1989,** *49*, 4968-4971.

Sefah, K.; Shangguan, D.; Xiong, X.; O'Donoghue, M. B.; Tan, W., *Nat. Protocols* **2010,** *5*, 1169-1185.

Shabbir, S. H.; Joyce, L. A.; da Cruz, G. M.; Lynch, V. M.; Sorey, S.; Anslyn, E. V., *Journal of the American Chemical Society* **2009,** *131*, 13125-13131.

Shawver, L. K.; Mann, E.; Elliger, S. S.; Dugger, T. C.; Arteaga, C. L., *Cancer Research* **1994,** *54*, 1367-1373.

Shoemaker, R. H., *Nat Rev Cancer* **2006,** *6*, 813-823.

Sironen, R. K.; Tammi, M.; Tammi, R.; Auvinen, P. K.; Anttila, M.; Kosma, V. M., *Experimental Cell Research* **2011,** *317*, 383-391.

Smith, G. P.; Petrenko, V. A., *Chemical Reviews* **1997,** *97*, 391-410.

Springsteen, G.; Wang, B., *Tetrahedron* **2002,** *58*, 5291-5300.

StatSoft, I., Electronic Statistics Textbook. StatSoft: 2013.

Stewart, S.; Ivy, M. A.; Anslyn, E. V., *Chemical Society Reviews* **2014,** *43*, 70-84.

Stibor, I.; Holakovsky, R.; Mustafina, A. R.; Lhotak, P., *Collect Czech Chem C* **2004,** *69*, 365-383.

Stibor, I.; Zlatukov, P., Chiral Recognition of Anions. In *Anion Sensing*, Stibor, I., Ed. Springer Berlin Heidelberg: 2005; Vol. 255, pp 31-63.

Stoolman, L. M.; Wang, T. L.; Situ, R.; Varani, J., *J Cell Physiol* **1993,** *154*, 593-600.

Subik, K.; Lee, J. F.; Baxter, L.; Strzepek, T.; Costello, D.; Crowley, P.; Xing, L.; Hung, M. C.; Bonfiglio, T.; Hicks, D. G.; Tang, P., *Breast Cancer (Auckl)* **2010,** *4*, 35-41.

Taherian, A.; Li, X.; Liu, Y.; Haas, T. A., *BMC Cancer* **2011,** *11*, 293.

Theodore J. Leitereg, D. G. G., Jean Harris, Thomas R. Mon, Roy Teranishi, *Journal of Agricultural and Food Chemistry* **1971,** *19*, 785-787.

Tuerk, C.; Gold, L., *Science* **1990,** *249*, 505-10.

Uphoff, K. W.; Bell, S. D.; Ellington, A. D., *Current Opinion in Structural Biology* **1996,** *6*, 281-288.

Varki, A.; Kannagi, R.; Toole, B. P., Glycosylation Changes in Cancer. . In *Essentials of Glycobiology*. , 2nd ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY: 2009.

Vogler, A.; Kunkely, H., *Inorganica Chimica Acta* **2006,** *359*, 4130-4138.

Wigler, M.; Navin, N.; Kendall, J.; Troge, J.; Andrews, P.; Rodgers, L.; McIndoo, J.; Cook, K.; Stepansky, A.; Levy, D.; Esposito, D.; Muthuswamy, L.; Krasnitz, A.; McCombie, W. R.; Hicks, J., *Nature* **2011,** *472*, 90-U119.

Willener, Y.; Joly, K. A.; Moody, C. J.; Tucker, J. H. R., *J Org Chem* **2008,** *73*, 1225-1233.

Wingren, C.; Borrebaeck, C. A. K., *Current Opinion in Biotechnology* **2008,** *19*, 55-61.

Winter, G.; Griffiths, A. D.; Hawkins, R. E.; Hoogenboom, H. R., *Annual Review of Immunology* **1994,** *12*, 433-455.

Wiskur, S. L.; Ait-Haddou, H.; Lavigne, J. J.; Anslyn, E. V., *Accounts of Chemical Research* **2001,** *34*, 963-972.

Wiskur, S. L.; Anslyn, E. V., *Journal of the American Chemical Society* **2001,** *123*, 10109-10110.

Witkowski, C.; Rabinovitz, I.; Nagle, R.; Affinito, K.-S.; Cress, A., *J Cancer Res Clin Oncol* **1993,** *119*, 637-644.

Wold, S.; Esbensen, K.; Geladi, P., *Chemometrics and Intelligent Laboratory Systems* **1987,** *2*, 37-52.

Wright, A. T.; Anslyn, E. V., *Chemical Society Reviews* **2006,** *35*, 14-28.

Wright, A. T.; Griffin, M. J.; Zhong, Z. L.; McCleskey, S. C.; Anslyn, E. V.; McDevitt, J. T., *Angew Chem Int Edit* **2005,** *44*, 6375-6378.

Xu, K. X.; Wang, Y. X.; Jiao, S. Y.; Zhao, J.; Wang, C. J., *Can J Chem* **2010,** *88*, 367-374.

Yamada, K. M.; Kennedy, D. W., *Journal of Cellular Physiology* **1987,** *130*, 21-28.

Yamada, T.; Shinoda, S.; Tsukube, H., *Chemical Communications* **2002**, 1218-1219.

Yang, B.; Zhang, L.; Turley, E. A., *Journal of Biological Chemistry* **1993,** *268*, 8617-8623.

Yang, D.; Li, X.; Fan, Y. F.; Zhang, D. W., *Journal of the American Chemical Society* **2005,** *127*, 7996-7997.

Yang, W.; Gao, S.; Gao, X.; Karnati, V. V. R.; Ni, W.; Wang, B.; Hooks, W. B.; Carson, J.; Weston, B., *Bioorganic & Medicinal Chemistry Letters* **2002,** *12*, 2175-2177.

You, C.-C.; Agasti, S. S.; Rotello, V. M., *Chemistry – A European Journal* **2008,** *14*, 143-150.

You, C. C.; Miranda, O. R.; Gider, B.; Ghosh, P. S.; Kim, I. B.; Erdogan, B.; Krovi, S. A.; Bunz, U. H. F.; Rotello, V. M., *Nat Nanotechnol* **2007,** *2*, 318-323.

Yuan, J.; Wang, G., *J Fluoresc* **2005,** *15*, 559-568.

Zhou, H. C.; Baldini, L.; Hong, J.; Wilson, A. J.; Hamilton, A. D., *Journal of the American Chemical Society* **2006,** *128*, 2421-2425.

Zhou, H. Y.; Jiao, P. F.; Yang, L.; Li, X.; Yan, B., *Journal of the American Chemical Society* **2011,** *133*, 680-682.

Zhu, L.; Anslyn, E. V., *Journal of the American Chemical Society* **2004,** *126*, 3676-3677.

Zhu, L.; Zhong, Z. L.; Anslyn, E. V., *Journal of the American Chemical Society* **2005,** *127*, 4260-4269.

Vita

Alex Gade received her BS in Chemistry with high honors from the College of William and Mary in May 2008. After a year of continued chemistry research and instruction in Oslo, Norway, Alex moved to Austin, TX to begin her PhD in Organic Chemistry. In October 2009, she joined the Anslyn group and began research into differential array sensing under the supervision of Dr. Eric V. Anslyn.

Permanent email: alexandragade@gmail.com

This dissertation was typed by Alexandra Moore Gade