Title

Site-Specific Surface Immobilization of Proteins and Development of Synthetic Fluorescent Sensors for Studying the Cell Biology of Metals

Author

Lee, Jessica Pui Yan

Publication Date

2016

Peer reviewed|Thesis/dissertation

Site-Specific Surface Immobilization of Proteins

and

Development of Synthetic Fluorescent Sensors for Studying the Cell Biology of Metals

By
Jessica Pui Yan Lee

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy
in
Chemistry
in the
Graduate Division
of the
University of California, Berkeley

Committee in charge:

Professor Matthew B. Francis, Co-Chair
Professor Sanjay Kumar, Co-Chair
Professor Evan W. Miller
Professor Danielle Tullman-Ercek

Spring 2016

Site-Specific Surface Immobilization of Proteins
and
Development of Synthetic Fluorescent Sensors for Studying the Cell Biology of Metals

By: Jessica Pui Yan Lee

Abstract

Site-Specific Surface Immobilization of Proteins
and
Development of Synthetic Fluorescent Sensors for Studying the Cell Biology of Metals

by

Jessica Pui Yan Lee

Doctor of Philosophy in Chemistry
University of California, Berkeley
Professor Matthew B. Francis, Co-Chair
Professor Sanjay Kumar, Co-Chair

As the major building block of life, proteins have an extraordinary range of structures and functions, which can be harnessed through their controlled immobilization. The covalent attachment of proteins onto surfaces has broad applications in biotechnology and materials science, but current strategies rely on nonspecific conjugation reactions. These reactions often affect protein properties and function. To address this problem, we have developed two site-specific immobilization strategies that target only the N-termini of proteins: one for the functionalization of hydrogels and one for the photopatterning of proteins on glass substrates. We have successfully applied the novel hydrogel material for studying substrate stiffness-dependent cell behavior. We envision that this site-specific immobilization strategy will be applicable to other hydrogel substrates and widely adopted in the field of mechanobiology.

Metal homeostasis is vital for the survival and health of all living organisms. Monitoring metal ion accumulation, efflux, trafficking and organelle distribution is essential for the understanding of roles of sodium and copper ions in cell biology. Synthetic fluorescent sensors have been used to investigate such molecular processes in high spatial and temporal resolution. We have developed novel sodium sensors, which show significant improvements over the commerically-available sensors in their sensitivities and affinities. They are able to reliably report sodium ion levels in live cells. We have also expanded the molecular imaging tool box with a series of photostable copper(I) sensors, which are able to report cellular copper(I) levels in real time.

Dedicated to my friends and family

Table of Contents

Acknowledgments

I am truly grateful for all the wonderful people who have supported me through the ups and downs of my 6-year journey at Berkeley. I would not have been able to do it without you all. I am honored that I was able to work with and have gotten to know many brilliant scientists and friends.

I could not have asked for more supportive mentors than Matt and Sanjay. They have made me the scientist I am today. I am often amazed by their infinite wisdom. I really appreciate Matt's optimistic yet realistic advice (both science and non-science related). He has also been there for me through my toughest times in graduate school. I have to thank Sanjay for taking the risk to accept me, his first chemistry student, into his lab. His positivity, scientific insight and rigor have shaped how I approach and do science.

Being a joint student in the Kumar and Francis labs, I was able to learn very different skill sets from all my labmates. The two labs are filled with great people and I have shared joy and tears with them inside and outside of lab. This thesis would not have been possible without the help and expertise of my labmates. The Kumar lab, especially Elena, taught me all about cell culture and biology in general. Mani and Nithya showed me the ins and outs of the IDP brush project. I will definitely miss the OB chicken nights and bottles of Rosé with the Kumar lab.

I learned about all aspects of bioconjugation chemistry in the Francis lab. It was a lot of fun working with the surface immobilization team, Kanwal and Matt S., who helped me get started in the lab. Jim and Jake have always been helpful and provided me with insightful discussions. Even though we started the IDP delivery project only several months ago, I truly enjoyed working with Alex, Ioana and Jenna. We were such a great team and I will never forget laughing about the floating eppe tubes in the bucket.

I have to thank my college chemistry professors, Matt Sazinsky and Dan O'Leary, who ignited my curiosity in chemistry and scientific research. I appreciate that they have continued to be my mentors even when I am at Berkeley.

Annie, Vivian, Sara and our luscious feasts helped me through the long days and nights in the first few years of graduate school. Alex and Michelle have also become my good friends even after they have left Berkeley to be professors.

I have to thank the Croucher Foundation and the De Benedictis Fellowship for funding my graduate studies.

Last but not least, I am very fortunate to have a supportive family in Hong Kong and my Bay Area friends, who have been my family away from home. Judy, of course, is my rock.

Chapter 1

N-terminal Specific Conjugation of Extracellular Matrix Proteins to 2-Pyridinecarboxaldehyde Functionalized Polyacrylamide Hydrogels

Abstract

Polyacrylamide hydrogels have been used extensively to study cell responses to the mechanical and biochemical properties of their substrates. A key step in fabricating these substrates is the conjugation of cell adhesion proteins to the polyacrylamide surfaces, which typically involves nonspecifically anchoring these proteins via side-chain functional groups. This can result in a loss of presentation control and altered bioactivity. Here, we describe a new functionalization strategy in which we anchor full-length extracellular matrix proteins to polyacrylamide substrates using 2-pyridinecarboxaldehyde, which can be co-polymerized into polyacrylamide gels and used to immobilize proteins by their N-termini. This one-step reaction proceeds under mild aqueous conditions and does not require additional reagents. We demonstrate that these substrates can readily conjugate to various extracellular matrix proteins, as well as promote cell adhesion and spreading. Notably, this chemistry supports the assembly and cellular remodeling of large collagen fibers, which is not observed using conventional side-chain amine-conjugation chemistry.

The studies described in this chapter were conducted in collaboration with Elena Kassianidou and James I. MacDonald, and will be a part of a peer-reviewed publication.

1.1. Introduction

Cells are sensitive to the mechanical properties of their environment. In particular, the stiffness of the extracellular matrix (ECM) substrate to which cells adhere can affect cell morphology, adhesion, migration and stem cell differentiation.[1-6] To study this behavior in culture, polyacrylamide (PAAm) hydrogels are most commonly used because the compliance of this material can easily be tuned to mimic stiffness values of soft tissues.[2-4,7-12] Additionally, PAAm does not promote protein adsorption or cell adhesion, enabling improved control over substrate functionalization with cell-adhesive ligands.[8,10,11]

The functionalization of PAAm with the ECM proteins required for cell adhesion has been achieved using pendant *N*-hydroxysuccinimide (NHS) esters that acylate lysine residues,[13] hydrazides that couple to periodate-oxidized glycans on proteins,[14] non-covalent adsorption,[15] and biotin–streptavidin interactions.[16] Perhaps the most common strategy uses a heterobifunctional crosslinker, sulfo-SANPAH (**1**, Figure 1.1a-c), which inserts into the PAAm backbone following photolysis of the aryl azide moiety.[7,8] The resulting polymer-linked NHS ester is then used to acylate amine groups on proteins. Although widely used, this process can create significant batch-to-batch variability, and the immobilized NHS esters are subject to competitive hydrolysis under protein attachment conditions. Moreover, NHS-esters can react with any amines on the proteins, resulting in nonspecific tethering at multiple sites with uncontrollable and unpredictable adhesive ligand presentation (Figure 1.1b). Furthermore, protein structure and activity may be affected if functionally and structurally important lysine residues are engaged with the surface.

This conjugation approach has also contributed to a major controversy in the field, questioning whether the density of these lateral tethers directly affects cell behavior and function.[17-19] It has been hypothesized that due to smaller pore size, stiffer substrates have shorter distances between surface anchorage points. This reduces the local deformability of these proteins relative to softer substrates, which have comparatively greater surface porosity and longer tethering distances, implying that local tether density controls cell behavior.[17] In contrast, a subsequent study showed that systematically varying PAAm porosity without altering stiffness does not significantly influence protein tethering, substrate deformation, or stem cell differentiation, implying that cells respond to bulk substrate stiffness rather than the degree of protein tethering.[18] Nevertheless, anchoring density was still observed to increase with increasing sulfo-SANPAH crosslinker concentration, complicating data interpretation.[18] Such studies have stimulated strong interest to develop ECM-hydrogel conjugation strategies with predictable and well-controlled attachment chemistry and ligand presentation.

This problem can be addressed by developing immobilization strategies that only anchor the ECM proteins to the hydrogel surface once, thereby ensuring that each protein is tethered predictably and with minimal perturbation of protein structure and function. Site-selective protein immobilization strategies using native cysteine residues, small molecule or peptide fusion tags have been developed for substrates other than PAAm, but they are generally incompatible with commercially-available tissue-purified ECM proteins.[20-22] As a result, there remains a significant need for site-specific strategies to immobilize ECM proteins to hydrogel surfaces for cell culture applications.

We report herein a novel and well-defined conjugation strategy in which native

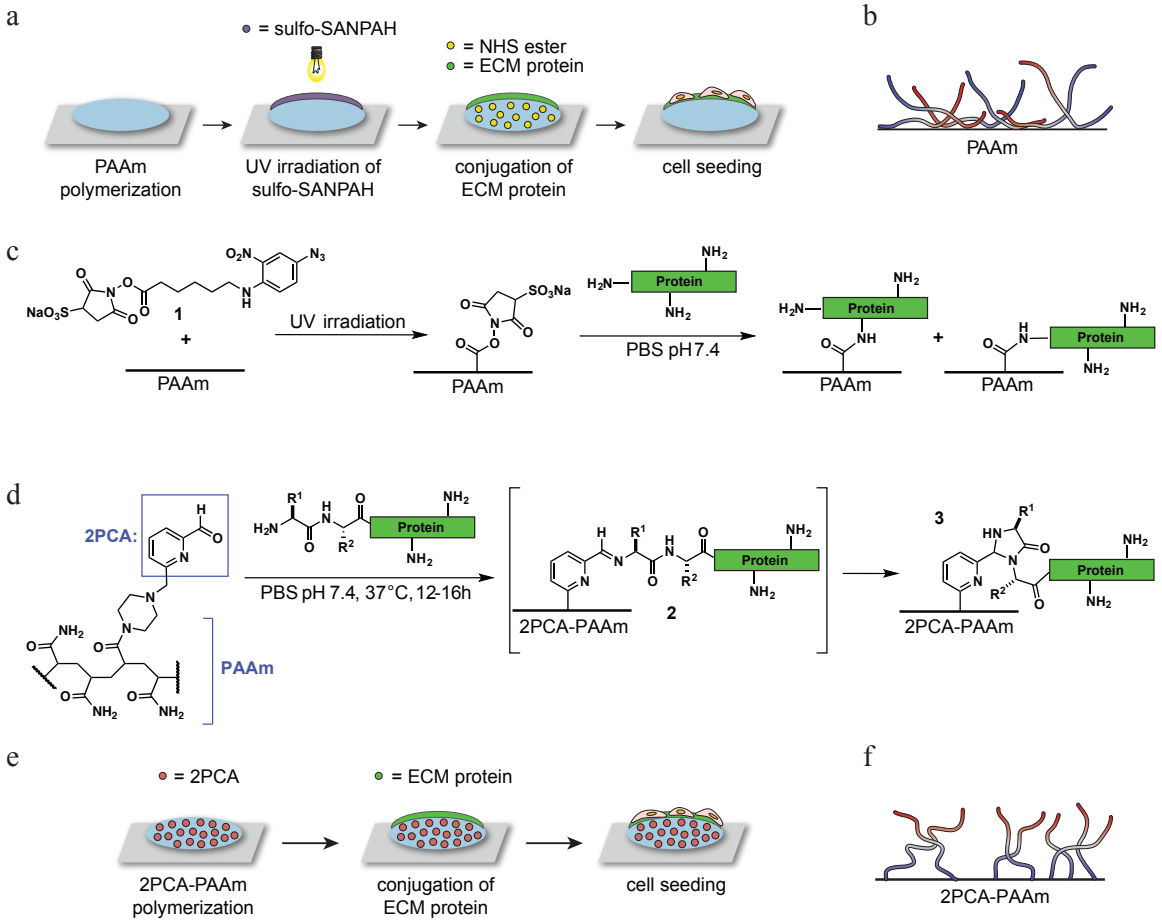


Figure 1.1. ECM protein conjugation on polyacrylamide (PAAm) substrates. (a-c) A scheme is shown for ECM protein conjugation using sulfo-SANPAH (**1**). This reaction targets all available amines, resulting in heterogeneous protein attachment. (d-f) In this work, protein N-termini are coupled to 2-pyridinecarboxaldehyde (2PCA)-modified PAAm gels, resulting in a more uniform presentation. Protein N-termini are shown in blue in (b) and (f).

proteins are conjugated to PAAm hydrogels specifically through their N-termini using 2-pyridinecarboxyaldehyde (2PCA) groups (Figure 1.1d-f). We demonstrate that this immobilization strategy is applicable to multiple widely used ECM proteins. Cells readily adhere and spread on these substrates, and despite the presence of only a single protein-hydrogel tethering point, cells recapitulate stiffness-dependent behavior. Moreover, 2PCA-conjugated surfaces remarkably and uniquely support the assembly of attached collagen chains into large fibers, which can be remodeled and bundled by attached cells in a manner reminiscent of collagen remodeling in tissue.

1.2. Results and Discussion

1.2.1. Conjugation of 2PCA to ECM Proteins

Our approach for site-specifically immobilizing ECM proteins to PAAm surfaces uses a recently reported one-step N-terminal specific modification of native proteins with 2-pyridinecarboxyaldehyde (2PCA) derivatives (Figure 1.1d-f).[23] This reaction involves the

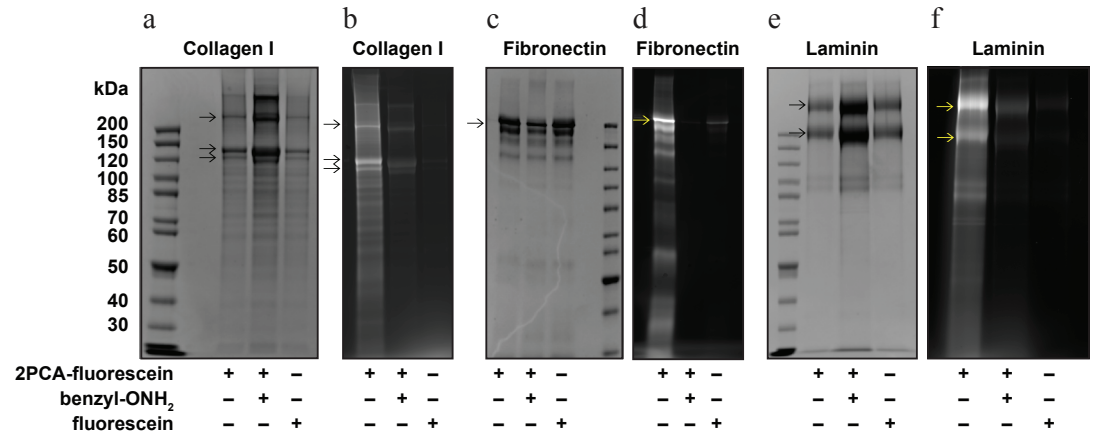


Figure 1.2. 2PCA conjugation of ECM proteins in solution. The N-terminal modification of type I collagen (a, b), fibronectin (c, d) and laminin (e, f) with 2PCA-fluorescein was monitored by SDS-PAGE. The SDS-PAGE gels were fluorescently imaged (b, d, f) to visualize conjugated products and Coomassie stained (a, c, e) to show total protein content. Benzylalkoxyamine (benzyl-ONH_2) and fluorescein were used as controls, which show minimal fluorescence. Arrows indicate corresponding ECM protein subunits.

formation of cyclic imidazolidinone product **3** through the addition of an immediately adjacent amide N-H group to imine intermediate **2** (Figure 1.1d). Importantly, this cyclization cannot occur when 2PCA imines are formed with lysine side chain amines, confining the reaction to the N-terminal position. In previous work, proteolytic digest and MS/MS analyses have confirmed the N-terminal specificity of the reaction.[23] This site-selective reaction is expected to minimize disruption of biologically active motifs and to improve presentation (Figure 1.1f). This is in contrast to conventional conjugation chemistry using sulfo-SANPAH crosslinker, which conjugates matrix proteins to PAAm surfaces via side-chain lysines.

Prior to using 2PCA for ECM protein-PAAm conjugation, we first assessed the ability of 2PCA to conjugate to the N termini of commonly used ECM proteins by incubating fluorescein labeled 2PCA (2PCA-fluorescein) with type I collagen, fibronectin and laminin. We verified by SDS-PAGE and fluorescent imaging that 2PCA-fluorescein is indeed covalently conjugated to all three ECM proteins (Figure 1.2). To test whether the conjugation occurs through the 2PCA moiety, benzylalkoxyamine was used to quench the aldehyde of 2PCA through oxime formation. ECM proteins show greatly reduced fluorescence when excess benzylalkoxyamine was added. Incubation with non-functionalized fluorescein also showed minimal fluorescence, suggesting no non-specific binding of fluorescein (Figure 1.2b,d,f, indicated by arrows). Together with our earlier work,[23] these results support the use of 2PCA to functionalize these ECM proteins specifically at their N-terminal positions.

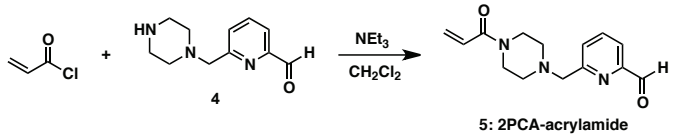


Figure 1.3. Synthesis of 2PCA-acrylamide (**4**) by coupling acryloyl chloride and piperazine-2PCA derivative **S5**.

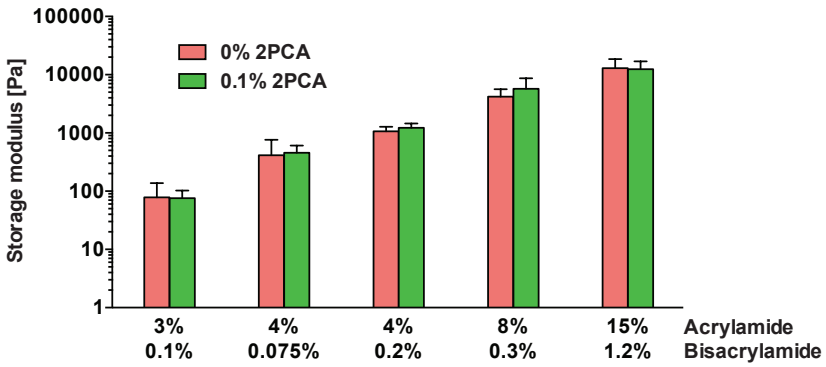


Figure 1.4. Storage modulus of PAAm substrates with varying acrylamide and bisacrylamide content. For 0.1% 2PCA substrates, 0.1% 2PCA-acrylamide was added relative to acrylamide monomer content. Bars represent mean ± 95% confidence intervals for $N \geq 3$ gels per condition. Student's unpaired t tests of gels with the same acrylamide and bisacrylamide content show no significant difference in storage moduli between PAAm substrates containing 0% and 0.1% 2PCA-acrylamide.

1.2.2. Mechanical Properties of 2PCA-PAAm Hydrogels

Having demonstrated that 2PCA can modify the N termini of ECM proteins, we next asked whether we could use this technique for ECM-PAAm immobilization. We decided to directly incorporate 2PCA into the growing PAAm chain during hydrogel polymerization and crosslinking (Figure 1.1a) to provide precise control over total 2PCA density, with even surface coverage. We therefore synthesized 2PCA-acrylamide **5** by coupling piperazine-2PCA derivative **4** and acryloyl chloride (Figure 1.3). We then added a small amount of the 2PCA-acrylamide monomer into the usual acrylamide/bisacrylamide mixture before polymerization (at 0.1% of the acrylamide monomer content). From a solvent screen, we identified that a minimal amount (<0.1% v/v) of acetone could solubilize 2PCA-acrylamide while keeping the PAAm substrate clear and uniform. Rheological measurements revealed that PAAm hydrogels polymerized with and without 0.1% 2PCA-acrylamide have statistically indistinguishable storage moduli, demonstrating that the presence of 0.1% 2PCA does not alter the bulk stiffness. Additionally, the storage moduli of PAAm hydrogels increase with acrylamide and/or bisacrylamide content whether or not 2PCA is incorporated (Figure 1.4).

1.2.3. Cell Response to ECM Protein Density on 2PCA-PAAm Substrates

We used U2OS human osteosarcoma cells as a model system to characterize our substrates, as their adhesion and motility are highly sensitive to substrate stiffness and adhesivity.[24] Since the extent of U2OS cell spreading correlates with adhesion protein density, we asked whether we could recapitulate this relationship with 2PCA-based immobilization (Figure 1.5). We studied changes in U2OS morphology on collagen-, fibronectin- and laminin-conjugated PAAm substrates as a function of 2PCA-acrylamide content (which is expected to control the amount of immobilized protein) at a constant stiffness of 5.8 kPa. In the absence of 2PCA, cells are poorly adherent and rounded (Figure 1.5, red bars and PAAm images), as PAAm substrates do not support passive adsorption of adhesion proteins.[8,10,11] We next attempted to conjugate the three ECM proteins to 2PCA-PAAm gel surfaces over a range of 2PCA concentrations. For all three proteins, U2OS spread area increases with 2PCA content, which suggests that higher 2PCA content leads to higher density of immobilized ECM proteins (Figure 1.5). Because U2OS cells spread well on 0.01% 2PCA-PAAm

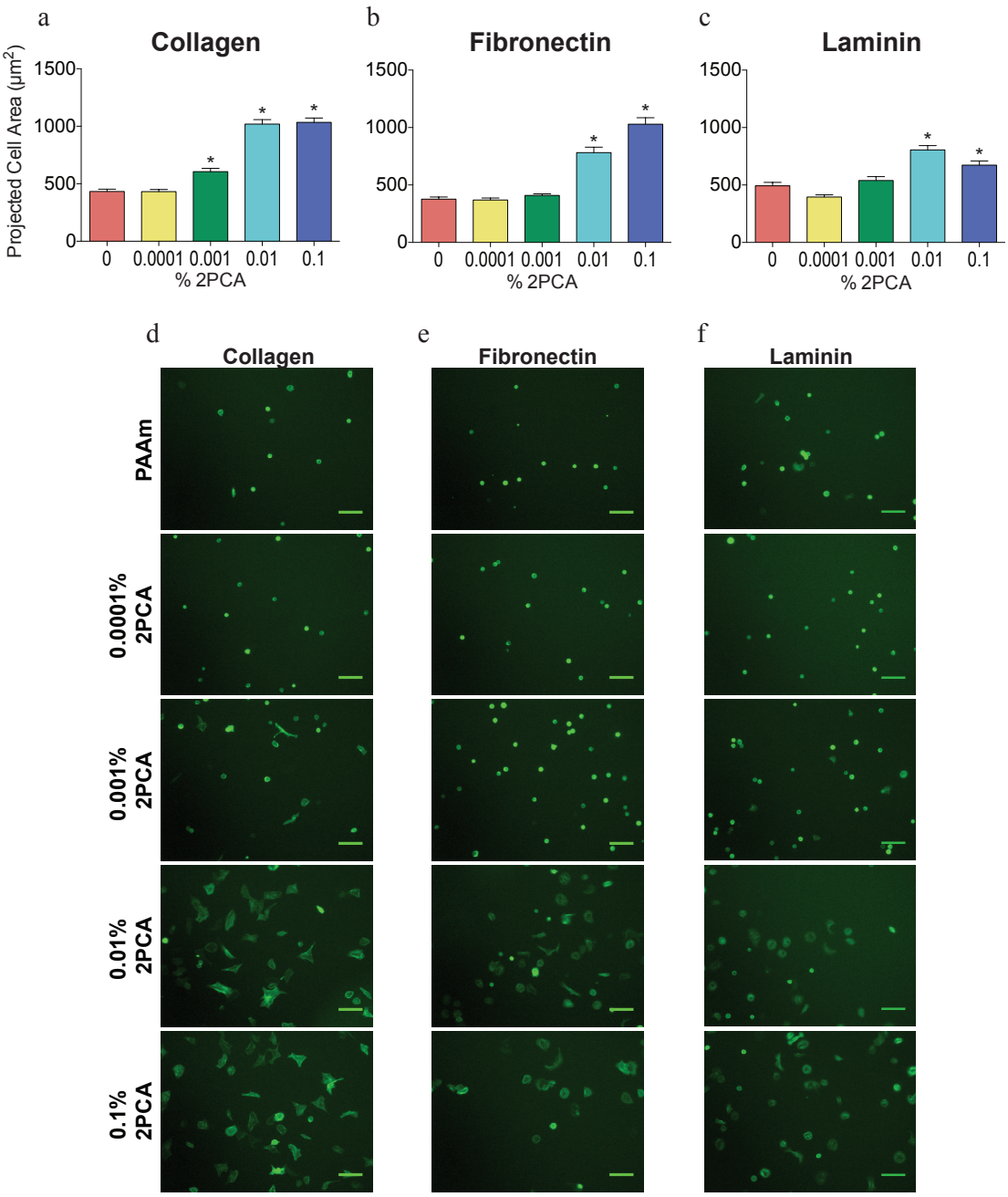


Figure 1.5. Dependence of U2OS cell spreading on PAAm substrates containing variable amounts of 2PCA-acrylamide. Quantification (a-c) and representative fields of view (d-f) are shown for U2OS spreading (green = LifeAct) on 5.8 kPa PAAm substrates. Bars represent mean ± SEM; *, $p < 0.05$ with respect to PAAm; $N \geq 144$, 65 and 106 cells per condition for collagen, fibronectin and laminin, respectively. Scale bars = 100 μm.

substrates conjugated with all three ECM proteins, we chose this formulation for use in subsequent studies.

1.2.4. Cell Response to PAAm Substrate Stiffness

Many cell types, including U2OS cells, increase adhesion and spread area with increases in ECM stiffness.[2,3,7,24-27] As an additional proof of principle, we next asked whether cells respond

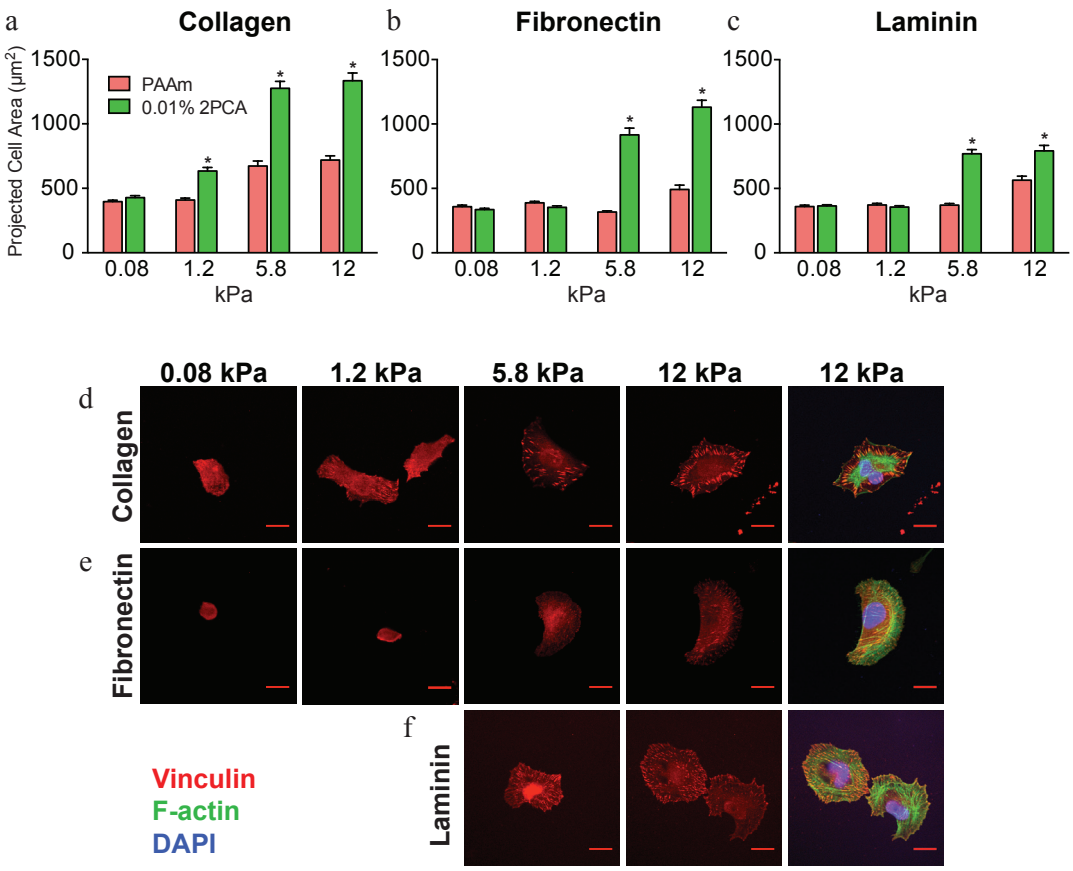


Figure 1.6. Influence of substrate rigidity on the spreading of U2OS cells cultured on PAAm gels with 0.01% 2PCA-acrylamide. (a-c) Cell spreading was quantified after U2OS cells were seeded on PAAm substrates conjugated with collagen (a), fibronectin (b) and laminin (c). Bars represent mean ± SEM; *, $p < 0.05$ with respect to PAAm of corresponding stiffness; $N \geq 122$, 90 and 65 cells per condition for collagen, fibronectin and laminin, respectively. (d-f) High-magnification images of representative U2OS cells cultured on 0.01% 2PCA-PAAm substrates conjugated with collagen (d), fibronectin (e) and laminin (f), and subsequently stained for the focal adhesion protein vinculin (red), F-actin (green), and nuclear DNA (blue). Cell adhesion on 0.08 kPa and 1.2 kPa substrates conjugated with laminin were too weak for staining. Scale bars = 20 μm.

to substrate stiffness when adhesive ligands are end-tethered on the 2PCA-PAAm substrates. We observed greater cell spreading on 2PCA-PAAm substrates relative to non-functionalized PAAm, which further confirms that all three ECM proteins are immobilized on the gel surface through 2PCA. U2OS spreading also dramatically increases with increasing 2PCA-PAAm stiffness, with cells rounded on soft substrates and extensively spread on stiff substrates (Figure 1.6a-c). Cells on stiff substrates show discrete and elongated focal adhesions, which become progressively smaller and more diffuse on softer substrates (Figure 1.6d-f), recapitulating the behavior of U2OS cells cultured on alginate substrates.[24] Thus, 2PCA-PAAm substrates support the expected rigidity-dependent spreading and adhesion seen on more conventional materials.

1.2.5. Collagen Fiber Formation on 2PCA-PAAm Substrates

Of the three matrix proteins considered, collagen is of particular interest because of its high abundance in tissue and ability to self-assemble into a remarkable hierarchy of bioactive higher-order structures, including sheets and bundles.[28-30] Whereas previous studies have shown

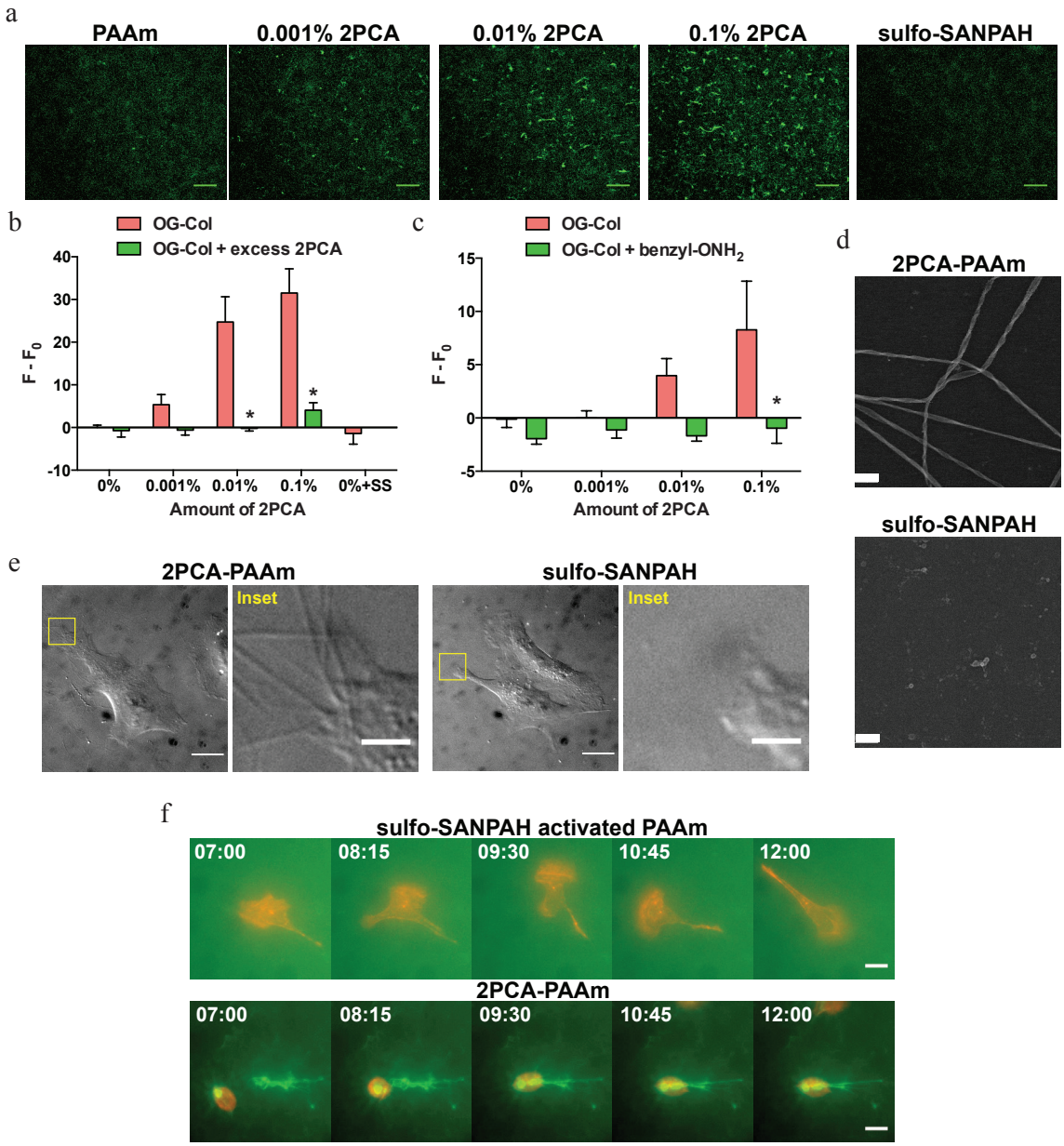


Figure 1.7. Collagen fiber assembly on PAAm substrates. (a) Representative fluorescence images are shown for OG-Col immobilized on 5.8 kPa PAAm gels with various 2PCA-acrylamide content. Scale bars = 100 μm. The reaction was analyzed by fluorescence intensity of OG-Col immobilized on substrates with or without 10 mM piperazine-2PCA (**S5**) (b) and 40 mM benzylalkoxyamine (benzyl-ONH_2) treatment (c). The fluorescence intensity of 0% 2PCA-PAAm without excess 2PCA or benzylalkoxyamine treatment (F_0) was subtracted from the fluorescence intensity (F) of the individual samples. Bars represent mean ± SEM; *, $p < 0.05$ with respect to PAAm of corresponding 2PCA content; $N \geq 6$ gels per condition. (d) Representative SEM (d) and DIC (e) images are shown for 50 μg/mL unlabeled collagen immobilized on 5.8 kPa 0.01% 2PCA-PAAm gels or sulfo-SANPAH activated PAAm gels. SEM scale bars = 0.5 μm. DIC scale bars = 20 μm. DIC inset scale bars = 4 μm. (f) Movie stills showing U2OS behavior on OG-Col immobilized PAAm substrates. 50 μg/mL collagen was conjugated to 5.8 kPa 0.0001% 2PCA-PAAm or sulfo-SANPAH activated PAAm before U2OS cell seeding (t = 0). PAAm with low 2PCA content is shown to highlight interactions between the cells and collagen fibers. Fluorescence time-lapse imaging of LifeAct-tag RFP (red) and OG-Col (green) started 7 h after cell seeding. Fluorescence contrast of sulfo-SANPAH activated PAAm substrate is enhanced to show the lack of fiber formation. Time stamp is in hours:minutes. Scale bars = 20 μm.

that monomeric collagen directly deposited or adsorbed onto rigid surfaces can nucleate the assembly of fibrils and small fibers,[31,32] such structures are in general not observed when collagen is covalently conjugated to PAAm hydrogels using sulfo-SANPAH.[33] A possible explanation is that lysine residues required for triple helix formation (e.g., through interchain salt bridges) are blocked in conjugation.[34,35] Moreover, one might anticipate that laterally immobilizing collagen on the hydrogel surface could sterically preclude interchain assembly. In contrast, 2PCA conjugation chemistry, which targets only the N-terminal amine rather than internal lysines and enables the rest of the chain to protrude freely into solution, would circumvent both of these issues. This raises the intriguing possibility that 2PCA-mediated conjugation could support or facilitate collagen fiber assembly.

To investigate, we fluorescently labeled type I collagen with Oregon Green NHS ester at low stoichiometry (OG-Col), leaving most amines free for subsequent conjugation (N-terminal amine for 2PCA, N-terminal amine plus internal lysines for sulfo-SANPAH). We then incubated 2PCA-PAAm substrates with OG-Col for covalent conjugation. Fluorescence images revealed the formation of fiber-like structures on the hydrogel surface, with the number and intensity of these structures increasing as the 2PCA-acrylamide content of the substrates increased, suggesting controlled 2PCA-dependent immobilization of collagen (Figure 1.7a-c). The presence of more fiber bundles on PAAm substrates with higher 2PCA content is indirect evidence of immobilization of structurally unperturbed collagen on the substrate.

To verify that collagen was conjugated to the PAAm substrate via the 2PCA moiety rather than noncovalently adsorbed, we conducted two control experiments using (1) excess benzylalkoxyamine to quench the aldehyde group of 2PCA prior to incubation in the collagen solution, and (2) excess 2PCA to block the N-terminus of collagen competitively during incubation in the collagen solution (Figure 1.7b,c). Both control surfaces showed reduced fluorescence relative to samples without excess benzylalkoxyamine or 2PCA treatment. The overall fluorescence of control substrates was comparable to that of PAAm substrates without 2PCA. Together, these experiments confirmed that collagen was covalently tethered to 2PCA-PAAm gels through the 2PCA moiety.

Collagen fiber bundles were absent when 2PCA was omitted from the preparation (Figure 1.7a, far left) or when OG-Col was conjugated via sulfo-SANPAH to the PAAm surface (Figure 1.7a, far right). We were not able to quantify the amount of monomeric collagen on the substrates by OG-Col fluorescence. This observation is in agreement with previous reports of fluorescent quantification of collagen density on sulfo-SANPAH activated PAAm substrates.[18] To verify fiber formation independent of fluorescence, we performed scanning electron microscopy (SEM) and differential interference contrast (DIC) microscopy on unlabeled surfaces, which confirmed the abundance of fibrous structures on the 2PCA-PAAm surfaces. In contrast, these features were not found on sulfo-SANPAH activated surfaces (Figure 1.7d,e).

Notably, despite the lack of observable collagen fibers, U2OS cells are able to adhere, spread and migrate on sulfo-SANPAH activated substrates (Figure 1.7f, top). This is expected, since several studies have shown that cells can spread and migrate on substrates coated in denatured collagen.[31-33] Presumably, however, such surfaces would not easily allow cellular deformation and remodeling of the fibers, which is key to many physiologically critical process including tissue compaction, directional 3D migration, and wound healing.[28,36,37] Indeed, time-lapse fluorescence imaging of

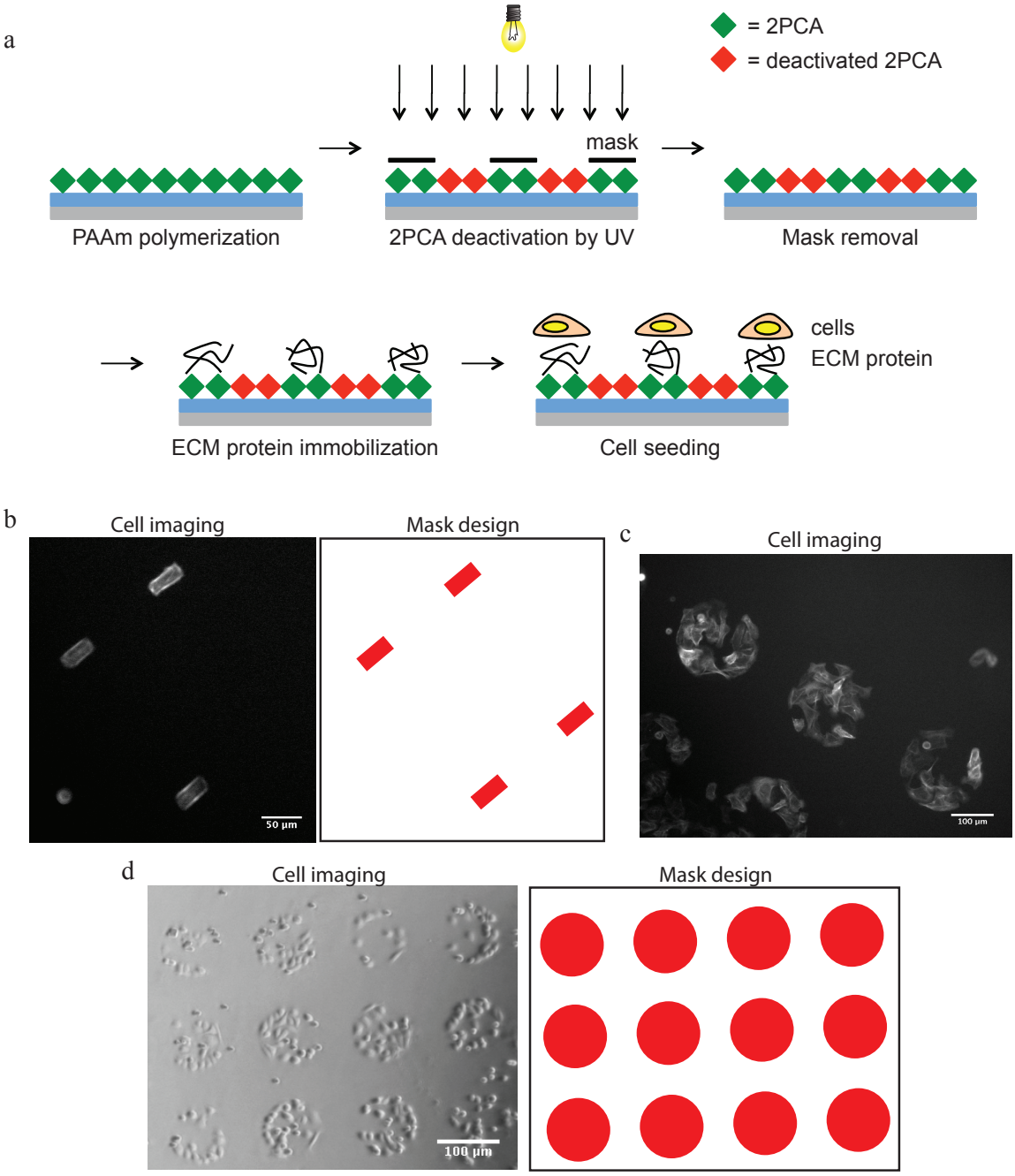


Figure 1.8. ECM protein and U2OS cell patterning on 2PCA-PAAm gels. (a) A scheme is shown for ECM protein patterning through UV deactivation of 2PCA, ECM protein coupling and subsequent cell seeding. Fluorescence microscopy images are shown for single cell (b) and collective cell (c) patterning on 5.8 kPa 2PCA-PAAm substrates (fluorescence = LifeAct). Scale bars = 50 µm and 100 µm for (b) and (c) respectively. (d) Bright field image of collective cell patterning on a 5.8 kPa 2PCA-PAAm substrate is shown. Scale bar = 100 µm. Mask designs are shown to the right of cell images in (b) and (d). Solid red denotes masked areas.

cells cultured on 2PCA-PAAm substrates with fluorescently labeled collagen fibers reveal that cells interact and remodel fibers (Figure 1.7f, bottom). Such remodeling events are not observed on surfaces in which collagen is immobilized by sulfo-SANPAH. Thus, 2PCA-PAAm surfaces have the distinct ability to study stiffness-dependent behavior while at the same time presenting the adhesive moiety as a "living" surface that mimics tissue presentation and allows cellular remodeling.

1.2.6. ECM Protein and Cell Patterning on 2PCA-PAAm Substrates

In addition to substrate stiffness, cell adhesive microenvironment geometry also affects phenotype, including cell shape, migration, proliferation and differentiation. Since both substrate stiffness and spatial organization of ECM proteins affect cell behavior, it would be useful to develop cell culture substrates where these two parameters can be modulated independently. Surface micropatterning has been used extensively to manipulate geometrical constraint of cells.[38-41] Patterning of single cell and island of multiple cells are useful for investigating single cell behavior and how spatially separated cells communicate respectively.[42,43] In particular, micropatterning on PAAm substrates has been mostly accomplished by microcontact-printing,[14,15,40,41,44-47] but drawbacks include incomplete protein transfer from the stamp to PAAm substrates, limited spatial resolution (only down to a few microns), non-site-specific conjugation chemistry, difficulty in large scale production of patterned substrates, and time consuming preparation involving multiple steps. Previous reports of ECM proteins photopatterned on PAAm are more reproducible and less time consuming, but they rely on non-specific adsorption of proteins on the PAAm substrates.[48,49] Hence, there is a strong need for a fast and scalable protein photopatterning method on PAAm substrates.

While investigating cell spread area on 2PCA-PAAm substrates, we noticed that UV irradiation (Sunray 400 SM, full power, 8 min) reduces U2OS cell spreading. This observation suggests that 2PCA can be deactivated by UV irradiation, so we took advantage of this UV deactivation to photopattern proteins. Polymerized 2PCA-PAAm was first sandwiched between a photomask and a silanized coverslip. Then, it was UV irradiated through the mask so that 2PCA was deactivated only in UV-exposed areas. After removal of the mask and immobilization of ECM protein, cells were finally seeded onto the patterned substrate (Figure 1.8a). Cells should only spread on masked areas, where 2PCA is intact to conjugate to ECM proteins.

Because UV light can be refracted through a buffer or air medium, PAAm substrates were placed in direct contact with the photomask to increase patterning resolution. A short irradiation time is needed so that the hydrogel is not dried out during UV irradiation. Hence, we used a deep UV light source from a UV/ozone cleaner to deactivate 2PCA rapidly. Because UV/ozone cleaners generate atomic oxygen, which is a very strong oxidizing agent, 2PCA was likely deactivated through aldehyde oxidation to the carboxylic acid.[50] We noticed that after irradiation, patterns were visible under a light microscope even without protein or cells, albeit extremely subtle and only on the curved edges of the hydrogel. This suggests that the high energy deep UV nonspecifically damages a thin surface layer of PAAm substrates as well. Therefore, we screened irradiation time to achieve efficient cell patterning while minimizing physical changes to the substrate surface. Using fibronectin as the cell adhesive ligand, we were able to pattern single and collective islands of U2OS cells on 2PCA-PAAm for all irradiation times ranging from 2 - 12 min (Figure 1.8b-d). To remove unspecific adhesion of cells on the irradiated area, washes can be further optimized by adjusting timing and rigor of wash steps.

Currently, the patterning resolution is limited by the resolution of the chrome quartz mask.

In theory, ECM proteins other than fibronectin can also be patterned on PAAm substrates, but irradiation times, ECM protein concentrations and wash steps will have to be optimized for efficient patterning. Nonetheless, we have developed a fast and efficient bench-top cell photopatterning strategy that can be performed in most laboratories. More importantly, this is the only reported patterning technique that site-specifically immobilizes proteins on PAAm substrates.

1.3. Conclusions

We have developed a strategy to end-tether full-length ECM proteins to hydrogel surfaces using 2PCA-mediated N-terminal conjugation. This approach recapitulates widely observed relationships between adhesive ligand density, ECM stiffness, and cell spreading, while uniquely allowing the assembly of collagen fibers that cells can remodel as they migrate. Importantly, this conjugation chemistry eliminates multiple covalent anchoring points on the ECM proteins, thereby fully decoupling the protein lateral tether density from bulk ECM stiffness – a major concern in the preparation of such substrates. In that context, it is notable that we observe strong stiffness-dependent behavior on 2PCA-conjugated PAAm substrates, despite the fact that the protein is anchored at a single point. This is consistent with the notion that ECM stiffness regulates cell adhesion and spreading independently of tether density.[18] Irrespective of this interpretation, we expect that this strategy should provide a useful alternative to traditional conjugation chemistries and may be used to immobilize or photopattern a broad variety of proteins and peptides to hydrogel surfaces for both basic and translational investigations.

1.4. Materials and Methods

1.4.1. General Procedures and Materials

Unless stated otherwise, all reagents and solvents used were of analytical grade and were used as received from commercial sources. Type I bovine collagen (PureCol, Advanced BioMatrix), human plasma fibronectin (Millipore) and mouse laminin (Gibco) were also used as received. PBS, pH 7.4, was purchased from Fisher Scientific.

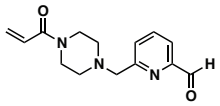
NMR spectra were recorded on a Bruker AVQ-400 spectrometer. ^{1}H NMR chemical shifts are reported as δ in units of parts per million (ppm) relative to $CDCl_3$ (δ 7.26). Multiplicities are reported as: s (singlet), br.s (broad singlet), d (doublet), t (triplet), q (quartet), dd (doublet of doublets), dt (doublet of triplets), br (broad) or m (multiplet). Coupling constants are reported as a *J* value in Hertz (Hz). ^{13}C NMR chemical shifts are reported as δ in units of parts per million (ppm) relative to $CDCl_3$ (δ 77.2).

Epifluorescence images were acquired using an inverted Nikon Eclipse Ti microscope equipped with a 10x air objective. DIC images were acquired using an inverted Nikon TE2000-E2 microscope equipped with a 60x oil objective. Time-lapse images were acquired every 15 min using an inverted Nikon TE2000-E2 microscope equipped with a 20x air objective. Both microscopes are equipped with a programmable stage and an incubator chamber to maintain constant temperature, humidity and CO_2 levels. High magnification images were acquired using a Prairie Technologies upright swept-field confocal microscope equipped with a 60x lens.

1.4.2. ECM Protein Conjugation with 2PCA-Fluorescein in Solution

2PCA-fluorescein was synthesized as previously reported.[23] Stock solutions of 50 mM fluorescein (free acid, Sigma) and 2PCA-fluorescein were prepared using DMSO (Sigma). A 250 mM stock solution of benzylalkoxyamine (Sigma, CDS001502) was prepared by the addition of an appropriate volume of water and adjustment to pH 7.5 using 5 M NaOH. In a 50 μL reaction, 0.5 mg/mL ECM protein was incubated with 2PCA-fluorescein or fluorescein (final concentration 5 mM) in 40 mM phosphate buffer, pH 7.5, at 37 °C without agitation. 2PCA-fluorescein was pre-incubated with benzylalkoxyamine (final concentration, 50 mM) for 4 h at 37 °C before the addition of an ECM protein. After 12 h, all samples were centrifuged at 12,000 rpm to remove any precipitate. Collagen and fibronectin samples were purified using a 10 kDa MWCO Amicon Ultra-4 centrifugal filter spin concentrator (Millipore). Laminin samples were purified by acetone precipitation. Samples were analyzed using SDS-PAGE (7.5% Mini-PROTEAN TGX gel, Bio-Rad, or Bolt 8% Bis-Tris Plus gel, Life Technologies), Typhoon 9410 variable mode imager (Amersham Biosciences) before staining with Coomassie Brilliant Blue R-250 (Bio-Rad) and Gel Doc EZ System (Bio-Rad) after staining.

1.4.3. Synthesis of 2PCA-Acrylamide



6-((4-acryloylpiperazin-1-yl)methyl)picolinaldehyde (5). 6-(piperazin-1-ylmethyl)-2-pyridinecarboxaldehyde (**4**) was synthesized as previously reported.[23] In an oven-dried round bottom flask, 5.5 mmol (1.0 equiv.) of 6-(piperazin-1-ylmethyl)-2-pyridinecarboxaldehyde (**4**) was stirred in 30 mL of dichloromethane. The solution was cooled in an ice-water bath and turned from cloudy to clear upon the addition of triethylamine (Sigma, 1.82 g, 2.5 mL, 18 mmol, 3.3 equiv.). Acryloyl chloride (Sigma, 0.54 g, 487 μL, 6.0 mmol, 1.1 equiv.) was then added dropwise at 0 °C. The reaction mixture was warmed to room temperature and stirred overnight. After evaporation of solvent and triethylamine under reduced pressure, the resulting material was purified by flash chromatography using 1.5% to 3% methanol in dichloromethane to afford product as a colorless viscous oil (0.86 g, 60% yield). ^{1}H NMR (400 MHz, chloroform-*d*): δ, 9.93 (s, 1H), 7.81–7.70 (m, 2H), 7.60 (d, *J* = 7.2 Hz, 1H), 6.46 (dd, *J* = 16.8, 10.6 Hz, 1H), 6.14 (d, *J* = 16.8 Hz, 1H), 5.57 (d, *J* = 10.6 Hz, 1H), 3.68 (s, 2H), 3.61 (s, 2H), 3.50 (s, 2H), 2.48–2.42 (m, 4H). ^{13}C NMR (100 MHz, chloroform-*d*): δ, 193.4, 165.22, 159.0, 152.2, 137.5, 127.8, 127.4, 127.3, 120.4, 63.7, 53.3, 52.7, 45.6, 41.7. HRMS (ESI) calculated for $C_{14}H_{18}N_3O_2$ ($[M+H]^+$) 260.1394, found 260.1388.

1.4.4. Preparation of 2PCA-PAAm Substrates

12 mm #1 circular glass coverslips were plasma cleaned (Harrick Plasma, PDC-32G), then silanized using a solution of 5% v/v acetic acid (Sigma) and 0.3% v/v PlusOne Bind-Silane (GE Healthcare) in ethanol for 3 min at room temperature. The silanized coverslips were then rinsed with 70% ethanol in water before drying with a Kimwipe. To ease the detachment of polymerized PAAm gels, a flat piece of glass was made hydrophobic by spraying with Rain-X (original glass water repellent) and drying with a Kimwipe. A series of 40% acrylamide and 2% bisacrylamide (Bio-Rad) were combined in different percentages and diluted to the appropriate volume with ultrapure water. A stock solution of 1.12 M 2PCA-acrylamide (molar equiv. of 40% acrylamide) was prepared in

acetone and diluted in ultrapure water (Invitrogen) to appropriate concentrations before adding to acrylamide/bisacrylamide solutions. Ammonium persulfate (Bio-Rad, 10% stock solution in water, final concentration 0.1%) and tetramethylethylenediamine (Bio-Rad, 1:1000 v/v) were added to the solutions immediately before sandwiching 30 μL of the polymerization solution between Rain-X treated glass and silanized coverslips. After the solution was allowed to polymerize for 15 to 30 min at room temperature, the gels were removed from the Rain-X treated glass, placed in 24-well plates (Falcon, cat# 353047) and rinsed in PBS for 10 min x 3 times. Substrates containing 2PCA-acrylamide were directly incubated in 400 μL of 50 μg/mL collagen or 25 μg/mL fibronectin or 25 μg/mL laminin in PBS at 37 °C for 12 to 16 h before rinsing in PBS for 10 min x 3 times and subsequent cell seeding. Substrates without 2PCA-acrylamide were submerged in 0.5 mL of 0.5 mg/mL sulfo-SANPAH (Pierce) in PBS. They were then activated with 8 min UV exposure (Sunray 400 SM) and rinsed in PBS 5 min x 3 times before incubating in ECM protein solutions as described above. The pH of ECM protein solutions was always checked and adjusted to pH 7.4, if necessary, prior to incubation with PAAm substrates.

1.4.5. Mechanical Characterization of PAAm Hydrogels

Substrates were prepared as described in section 1.4.4, with the exception that for each substrate, a 480 μL solution was polymerized in between and then detached from two Rain-X treated 25 mm glass coverslips. Anton Paar Physica MCR 301 rheometer with 25-mm parallel plate was used to determine substrate stiffness at 37 °C. The linear regime was determined based on amplitude sweeps over the range c = 0.1 to 10%. Frequency sweeps at 1% strain over 0.1 to 20 Hz were recorded to extract storage, loss and complex moduli. Storage moduli at 0.1 Hz of 0.1% 2PCA-PAAm substrates containing final acrylamide/bisacrylamide (A/B) percentages of 3% A/0.1% B, 4% A/0.075% B, 4% A/0.2% B, 8% A/0.3% B and 15% A/1.2% B were measured to be 0.08, 0.46, 1.2, 5.8 and 12 kPa respectively. At least three independent samples were measured per condition. Note that the stiffness reported here is the storage modulus instead of the Young's modulus.

1.4.6. Cell Line and Reagents

Immediate early promoter of cytomegalovirus (Pcmv IE) and LifeAct fused to TagRFP were amplified by polymerase chain reaction from pCMV-LifeAct-TagRFP plasmid (Ibidi) and PacI/EcoRI restriction sites were incorporated at each end. The PCR product was subcloned into pFUG-IP (kindly provided by D.V. Schaffer, University of California, Berkeley, CA) after removal of the hUbC promoter and EGFP by digesting with PacI/EcoRI. Viral particles were packaged in 293T cells and used to infect U2OS human osteosarcoma cells (ATTC) at a multiplicity of infection of 1.5 IU/cell. Cells expressing the vectors were sorted on a DAKO-Cytomation MoFlo High Speed Sorter based on RFP fluorescence. Cells were cultured in DMEM (Gibco, cat# 11965) with 10% fetal bovine serum (JR Scientific), 1% penicillin/streptomycin (Thermo Fisher Scientific) and 1% MEM non-essential amino acids solution (Life Technologies Corporation) in a 37 °C incubator in the presence of 5% CO_2. U2OS-RFP-LifeAct was used in all experiments described in this chapter.

1.4.7. Cell Spreading Experiments

Trypsinized U2OS were seeded on substrates at a density of 9,000 cells/cm^2 in serum-containing culture medium describe above. Cells were allowed to adhere and spread for 2 h before removal of non-adhered cells by aspiration. Fresh culture media was added and live cells were imaged 3-5 h after seeding. At least 5 fields of view were acquired per substrate. Using ImageJ,

projected cell area was determined based on Lifeact-RFP signal.

1.4.8. Immunofluorescence Staining

For immunostaining, all reagents were diluted in PBS. After fixing with 4% paraformaldehyde, cells were washed twice with PBS and permeabilized with 0.5% Triton-X 100 (EMD Biosciences) in 5% goat serum (Gibco) for 15 min at room temperature. Cells were rinsed twice and blocked in 5% goat serum for 45 min at room temperature. Cells were then stained for focal adhesions using mouse monoclonal anti-vinculin IgG (Sigma, V9131, 1:200 dilution) in 1% goat serum overnight at 4 °C. After rinsing in 1% goat serum twice, cells were incubated with Alexa Fluor 633 labeled goat anti-mouse IgG (ThermoFisher, A21052, 1:400 dilution) in 1% goat serum for 1 h at room temperature. Finally, F-actin and the nucleus were stained with Alexa Fluor 546 conjugated phalloidin (Thermo Fisher, A22283, 1:200 dilution) and DAPI (ThermoFisher, 2.5 μg/mL) for 20 min at room temperature.

1.4.9. Characterization of Surface Available Ligands

To label collagen with a fluorophore, 1 mL of 3 mg/mL type I collagen was mixed with 9 mL of 0.1 M carbonate buffer, pH 9.2, and 13 μL of 26 mM Oregon Green 488 carboxylic acid, succinimidyl ester (Molecular Probes, stock solution prepared in DMF). The collagen solution was kept at 4 °C at all times to prevent gelation. The reaction was placed on a rotator for 14 h at 4 °C before using a 10 kDa MWCO Amicon Ultra-4 centrifugal filter spin concentrator (Millipore) to remove unconjugated fluorophore and exchange into 0.01 N HCl. Oregon Green-labeled collagen (OG-Col) was further purified using a size exclusion column (NAP-10, GE Healthcare). Based on a BCA assay and absorbance at 488 nm, the extent of labeling was estimated to be 4 fluorophores per collagen monomer on average. 2PCA-acrylamide containing substrates were incubated with 50 μg/mL total collagen (1:4, Oregon Green labeled: unlabeled) as described in section 1.4.4. As control experiments, 2PCA-PAAm substrates were treated with 10 mM piperazine-2PCA (**4**) during OG-Col incubation or preincubated with 40 mM benzylalkoxyamine in 50 mM acetate buffer, pH 5.0, for 20-24 h before OG-Col immobilization. After rinsing, U2OS cells were seeded onto the substrates and allowed to adhere for 3 h before acquiring fluorescence images. To ensure unbiased determination of focal planes, focusing was done on adhered cells using bright field. Fluorescence images were processed using rolling ball background subtraction and overall fluorescence was quantified with ImageJ software (NIH).

1.4.10. Scanning Electron Microscopy

PAAm substrates were prepared as described in section 1.4.4, with the exception that for each substrate, a 10 μL solution was polymerized in between a 5 mm Rain-X treated coverslip and silicon wafers pre-treated with hydrophobic solution (OMS OptoChemicals). After overnight incubation at 37 °C in 50 μg/mL collagen, the hydrogel substrates were fixed in 2% glutaraldehyde for 1 h, rinsed in buffer (10 min x 3 times), then incubated in 1% osmium tetroxide, for 1 h at room temperature in 0.1 M sodium cacodylate buffer at pH 7.2. After rinsing again in buffer (5 min x 3 times), the samples were dehydrated in ethanol, dried using the critical-point technique (AutoSamdri 815, Tousimis), and sputter-coated with approximately 20 nm of gold and palladium (Tousimis) before acquiring images using a Hitachi S-5000 scanning electron microscope.

1.4.11. Protein and Cell Patterning

0.01% 2PCA-PAAm was prepared as described in section 1.4.4 using a 12 mm circular coverslip. A 7 μL droplet of H_2O was then sandwiched between a polymerized hydrogel surface and a chrome quartz photomask (aBeam Technologies) to ensure complete contact between the two surfaces. The sandwich was then irradiated for 2 to 12 min with deep UV (Jelight 42 UVO cleaner). After carefully removing the hydrogel from the mask using a razor blade and water, it was swelled in PBS for 20 min at room temperature. At this stage, patterns on the 2PCA-PAAm substrate can be confirmed with a light microscope. The hydrogel was then conjugated with 25 μg/mL fibronectin as described in section 1.4.4 and seeded with U2OS as described in section 1.4.7.

1.4.12. Statistical Analysis

All images and data are representative of the results of at least three or more independent biological experiments. Data are reported as mean ± SEM unless stated otherwise. Statistical significance was determined using one-way ANOVA followed by the Tukey multiple comparisons test. A Student's unpaired *t* test was used if statistical comparisons were made between two sets of data. The significance level was set at $p < 0.05$.

1.5. References

1. Discher, D. E., Janmey, P., and Wang, Y. L. (2005) Tissue cells feel and respond to the stiffness of their substrate, *Science 310*, 1139-1143.

2. Engler, A., Bacakova, L., Newman, C., Hategan, A., Griffin, M., and Discher, D. (2004) Substrate compliance versus ligand density in cell on gel responses, *Biophys. J. 86*, 617-628.

3. Ulrich, T. A., de Juan Pardo, E. M., and Kumar, S. (2009) The mechanical rigidity of the extracellular matrix regulates the structure, motility, and proliferation of glioma cells, *Cancer Res. 69*, 4167-4174.

4. Keung, A. J., de Juan-Pardo, E. M., Schaffer, D. V., and Kumar, S. (2011) Rho GTPases mediate the mechanosensitive lineage commitment of neural stem cells, *Stem Cells 29*, 1886-1897.

5. Keung, A. J., Dong, M., Schaffer, D. V., and Kumar, S. (2013) Pan-neuronal maturation but not neuronal subtype differentiation of adult neural stem cells is mechanosensitive, *Sci. Rep. 3*, 1817.

6. Ahearne, M. (2014) Introduction to cell-hydrogel mechanosensing, *Interface Focus 4*, 2.

7. Pelham, R. J., Jr., and Wang, Y. (1997) Cell locomotion and focal adhesions are regulated by substrate flexibility, *Proc. Natl. Acad. Sci. USA 94*, 13661-13665.

8. Tse, J. R., and Engler, A. J. (2010) Preparation of hydrogel substrates with tunable mechanical properties, *Curr. Protoc. Cell Biol. 47*, 10.16.1-10.16.16.

9. Levental, I., Georges, P. C., and Janmey, P. A. (2007) Soft biological materials and their impact on cell function, *Soft Matter 3*, 299-306.

10. Georges, P. C., and Janmey, P. A. (2005) Cell type-specific response to growth on soft materials, *J. Appl. Physiol. (1985) 98*, 1547-1553.

11. Gribova, V., Crouzier, T., and Picart, C. (2011) A material's point of view on recent developments of polymeric biomaterials: control of mechanical and biochemical properties, *J. Mater. Chem. 21*, 14354-14366.

12. Kandow, C. E., Georges, P. C., Janmey, P. A., and Beningo, K. A. (2007) Polyacrylamide hydrogels for cell mechanics: steps toward optimization and alternative uses, *Methods Cell Biol. 83*, 29-46.

13. Pless, D. D., Lee, Y. C., Roseman, S., and Schnaar, R. L. (1983) Specific cell-adhesion to immobilized glycoproteins demonstrated using new reagents for protein and glycoprotein immobilization, *J. Biol. Chem. 258*, 2340-2349.

14. Damljanovic, V., Lagerholm, B. C., and Jacobson, K. (2005) Bulk and micropatterned conjugation of extracellular matrix proteins to characterized polyacrylamide substrates for cell mechanotransduction assays, *Biotechniques 39*, 847-851.

15. Grevesse, T., Versaevel, M., Circelli, G., Desprez, S., and Gabriele, S. (2013) A simple route to functionalize polyacrylamide hydrogels for the independent tuning of mechanotransduction cues, *Lab Chip 13*, 777-780.

16. Hynd, M. R., Frampton, J. P., Dowell-Mesfin, N., Turner, J. N., and Shain, W. (2007) Directed cell growth on protein-functionalized hydrogel surfaces, *J. Neurosci. Methods 162*, 255-263.

17. Trappmann, B., Gautrot, J. E., Connelly, J. T., Strange, D. G., Li, Y., Oyen, M. L., Cohen Stuart, M. A., Boehm, H., Li, B., Vogel, V., Spatz, J. P., Watt, F. M., and Huck, W. T. (2012) Extracellular-matrix tethering regulates stem-cell fate, *Nat. Mater. 11*, 642-649.

18. Wen, J. H., Vincent, L. G., Fuhrmann, A., Choi, Y. S., Hribar, K. C., Taylor-Weiner, H., Chen, S., and Engler, A. J. (2014) Interplay of matrix stiffness and protein tethering in stem cell differentiation, *Nat. Mater. 13*, 979-987.

19. Kumar, S. (2014) Cellular mechanotransduction: stiffness does matter, *Nat. Mater. 13*, 918-920.

20. Chen, Y. X., Triola, G., and Waldmann, H. (2011) Bioorthogonal chemistry for site-specific labeling and surface immobilization of proteins, *Acc. Chem. Res. 44*, 762-773.

21. Wong, L. S., Khan, F., and Micklefield, J. (2009) Selective covalent protein immobilization: strategies and applications, *Chem. Rev. 109*, 4025-4053.

22. Bhagawati, M., and Kumar, S. (2014) Biofunctionalization of hydrogels for engineering the cellular microenvironment. In *Micro- and Nanoengineering of the Cell Surface* (Karp, J., and Zhao, W., Ed.), pp 315-348, William Andrew, Oxford.

23. MacDonald, J. I., Munch, H. K., Moore, T., and Francis, M. B. (2015) One-step site-specific modification of native proteins with 2-pyridinecarboxyaldehydes, *Nat. Chem. Biol. 11*, 326-331.

24. Chaudhuri, O., Gu, L., Darnell, M., Klumpers, D., Bencherif, S. A., Weaver, J. C., Huebsch, N.,

and Mooney, D. J. (2015) Substrate stress relaxation regulates cell spreading, *Nat. Commun. 6*, 6364.

25. Pathak, A., and Kumar, S. (2012) Independent regulation of tumor cell migration by matrix stiffness and confinement, *Proc. Natl. Acad. Sci. USA 109*, 10334-10339.

26. Kim, Y., and Kumar, S. (2014) CD44-mediated adhesion to hyaluronic acid contributes to mechanosensing and invasive motility, *Mol. Cancer Res. 12*, 1416-1429.

27. Rape, A. D., Zibinsky, M., Murthy, N., and Kumar, S. (2015) A synthetic hydrogel for the high-throughput study of cell-ECM interactions, *Nat. Commun. 6*, 8129.

28. Ulrich, T. A., Jain, A., Tanner, K., MacKay, J. L., and Kumar, S. (2010) Probing cellular mechanobiology in three-dimensional culture with collagen-agarose matrices, *Biomaterials 31*, 1875-1884.

29. Pathak, A., and Kumar, S. (2013) Transforming potential and matrix stiffness co-regulate confinement sensitivity of tumor cell migration, *Integr. Biol. 5*, 1067-1075.

30. Badylak, S. F. (2005) Regenerative medicine and developmental biology: the role of the extracellular matrix, *Anat. Rec. 287B*, 36-41.

31. Elliott, J. T., Tona, A., Woodward, J. T., Jones, P. L., and Plant, A. L. (2003) Thin films of collagen affect smooth muscle cell morphology, *Langmuir 19*, 1506-1514.

32. McDaniel, D. P., Shaw, G. A., Elliott, J. T., Bhadriraju, K., Meuse, C., Chung, K. H., and Plant, A. L. (2007) The stiffness of collagen fibrils influences vascular smooth muscle cell phenotype, *Biophys. J. 92*, 1759-1769.

33. Dingal, P. C. D. P., Bradshaw, A. M., Cho, S., Raab, M., Buxboim, A., Swift, J., and Discher, D. E. (2015) Fractal heterogeneity in minimal matrix models of scars modulates stiff-niche stem-cell responses via nuclear exit of a mechanorepressor, *Nat. Mater. 14*, 951-960.

34. Persikov, A. V., Ramshaw, J. A. M., Kirkpatrick, A., and Brodsky, B. (2005) Electrostatic interactions involving lysine make major contributions to collagen triple-helix stability, *Biochemistry 44*, 1414-1422.

35. Gurry, T., Nerenberg, P. S., and Stultz, C. M. (2010) The contribution of interchain salt bridges to triple-helical stability in collagen, *Biophys. J. 98*, 2634-2643.

36. Ulrich, T. A., Lee, T. G., Shon, H. K., Moon, D. W., and Kumar, S. (2011) Microscale mechanisms of agarose-induced disruption of collagen remodeling, *Biomaterials 32*, 5633-5642.

37. Lu, P., Takai, K., Weaver, V. M., and Werb, Z. (2011) Extracellular matrix degradation and remodeling in development and disease, *Cold Spring Harb. Perspect. Biol. 3*. 12.

38. Vignaud, T., Blanchoin, L., and Thery, M. (2012) Directed cytoskeleton self-organization, *Trends Cell Biol. 22*, 671-682.

39. Thery, M. (2010) Micropatterning as a tool to decipher cell morphogenesis and functions, *J. Cell Sci. 123*, 4201-4213.

40. Parker, K. K., Brock, A. L., Brangwynne, C., Mannix, R. J., Wang, N., Ostuni, E., Geisse, N. A., Adams, J. C., Whitesides, G. M., and Ingber, D. E. (2002) Directional control of lamellipodia extension by constraining cell shape and orienting cell tractional forces, *FASEB J. 16*, 1195-1204.

41. Versaevel, M., Grevesse, T., and Gabriele, S. (2012) Spatial coordination between cell and nuclear shape within micropatterned endothelial cells, *Nat. Commun. 3*. 671.

42. Kilian, K. A., Bugarija, B., Lahn, B. T., and Mrksich, M. (2010) Geometric cues for directing the differentiation of mesenchymal stem cells, *Proc. Natl. Acad. Sci. USA 107*, 4872-4877.

43. Frimat, J. P., Sisnaiske, J., Subbiah, S., Menne, H., Godoy, P., Lampen, P., Leist, M., Franzke, J., Hengstler, J. G., van Thriel, C., and West, J. (2010) The network formation assay: a spatially standardized neurite outgrowth analytical display for neurotoxicity screening, *Lab Chip 10*, 701-709.

44. Polio, S. R., Rothenberg, K. E., Stamenovic, D., and Smith, M. L. (2012) A micropatterning and image processing approach to simplify measurement of cellular traction forces, *Acta Biomaterialia 8*, 82-88.

45. Rape, A. D., Guo, W. H., and Wang, Y. L. (2011) The regulation of traction force in relation to cell shape and focal adhesions, *Biomaterials 32*, 2043-2051.

46. Zhang, J., Guo, W. H., Rape, A., and Wang, Y. L. (2013) Micropatterning Cell Adhesion on Polyacrylamide Hydrogels. In *Cell-Cell Interactions: Methods and Protocols, 2nd Edition, Vol 1066* (Baudino, T. A., Ed), pp 147-156, Humana Press, New York.

47. Tang, X., Ali, M. Y., and Saif, M. T. A. (2012) A novel technique for micro-patterning proteins and cells on polyacrylamide gels, *Soft Matter 8*, 7197-7206.

48. Vignaud, T., Ennomani, H., and Thery, M. (2014) Polyacrylamide hydrogel micropatterning, *Methods Cell Biol. 120*, 93-116.

49. Tseng, Q. Z., Wang, I., Duchemin-Pelletier, E., Azioune, A., Carpi, N., Gao, J., Filhol, O., Piel, M., Thery, M., and Balland, M. (2011) A new micropatterning method of soft substrates reveals that different tumorigenic signals can promote or reduce cell contraction levels, *Lab Chip 11*, 2231-2240.

50. Vig, J. R. (1985) UV/ozone cleaning of surfaces, *J. Vac. Sci. Technol. A 3*, 1027-1034.

Chapter 2

Site-Specific Photopatterning of an Intrinsically Disordered Protein on Glass

Abstract

Intrinsically disordered proteins (IDPs) have emerged as an important class of protein biomaterials. Due to their lack of well-defined structure and high degree of flexibility, it has been challenging to elucidate how IDP conformation depends on environmental conditions. Site-specific patterning of IDPs on a two-dimensional substrate would facilitate this process by rendering this problem amenable to surface characterization techniques such as atomic force microscopy. To this end, we have developed a protein patterning technique based on the photoinitiated azidophenol conjugation with aniline functionalized substrates. We successfully labeled the N-terminus of an IDP with azidophenol and used this moiety to photopattern the IDP to a glass surface. Finally, we characterized the efficacy of this patterning technique using fluorescence imaging and laser ablation inductively coupled plasma mass spectrometry. This bench-top photolithography method is fast and generalizable to most proteins without the need for protein engineering or specialized equipment.

2.1. Introduction

2.1.1. Intrinsically Disordered Proteins (IDPs)

Throughout the 20th century, well-defined protein structure was classically considered to be necessary for protein function. In the past two decades, however, a class of proteins termed intrinsically disordered proteins (IDPs) has emerged. They do not possess well-defined three-dimensional structures, are structurally flexible and, depending on their environment, adopt an ensemble of conformations, which is important for function.[1-6] Their amino acid compositions generally differ significantly from globular proteins. Most IDPs are highly charged and hydrophilic, lacking a compacted hydrophobic core. Although they were recognized much more recently than globular proteins, IDPs are highly abundant, with the fraction of disordered residues increasing with increasing size of the proteome.[7] Several studies have predicted that 30-50% eukaryotic proteins possess long (> 41 residues) disordered segments while 12% of all eukaryotic proteins are fully disordered.[4,7,8] Despite their abundance, relatively little is known about their structural conformations compared to globular proteins. The major reason is that IDPs cannot be studied using conventional structural techniques, such as protein crystallography, because of their flexibility. Therefore, there remains a need for a system to study the dynamic conformations of IDPs in order to shed light on their structure-function relationships.

Neurofilments (NFs) are the most abundant cytoskeletal protein of large, myelinated axons. They assemble into an intracellular network to protect axons from mechanical stress. They are composed of three subunits: neurofilament heavy (NFH), medium (NFM) and light (NFL), which are predicted to be 112.5, 102.5 and 61.5 kDa respectively in humans, and can heteropolymerize with other intermediate filaments such as α-internexin or peripherin.[9,10] NF proteins are co-assembled to form filaments with diameter of ca. 10 nm. The C-terminal sidearm domains of NFH and NFM (NFH-SA and NFM-SA) are fully disordered and laterally project from the filament backbone. Interestingly, the serine residues within their repetitive lysine-serine-proline (KSP) motif can be heavily phosphorylated. These IDPs have been suggested to be important for axonal transport, axonal growth, stabilizing filament network, and determining interfilament spacing, but their roles are still unclear.[11-16]

Not only is the understanding of the dynamic conformation of NF's disordered sidearm domains and how it relates to their physiological role important, they possess unique characteristics that render them useful as biomaterials. Unlike globular proteins, IDPs adopt an extended conformation due to their lack of defined secondary structure. Unlike synthetic polymers, recombinantly expressed IDPs are homogeneous and their sequences can be engineered to provide desired physical properties. Introduction of modification handles or protein fusions can also be accomplished using molecular biological tools. Moreover, many IDP structures can be induced in certain environments to promote specific substrate binding.[3-6] In principle, they could also serve as substrates for enzymes, offering a functionality that has been challenging to incorporate into synthetic polymers. Thus, IDPs are very attractive platforms from which to build smart materials. IDPs based on NFH sidearm domains, in particular, have been grafted onto silicon oxide 2D substrates and gold nanoparticles to form polymer brushes with the goal of studying their environment-dependent conformational changes and intermolecular interactions, as well as generating stimuli-responsive smart materials.[17,18]

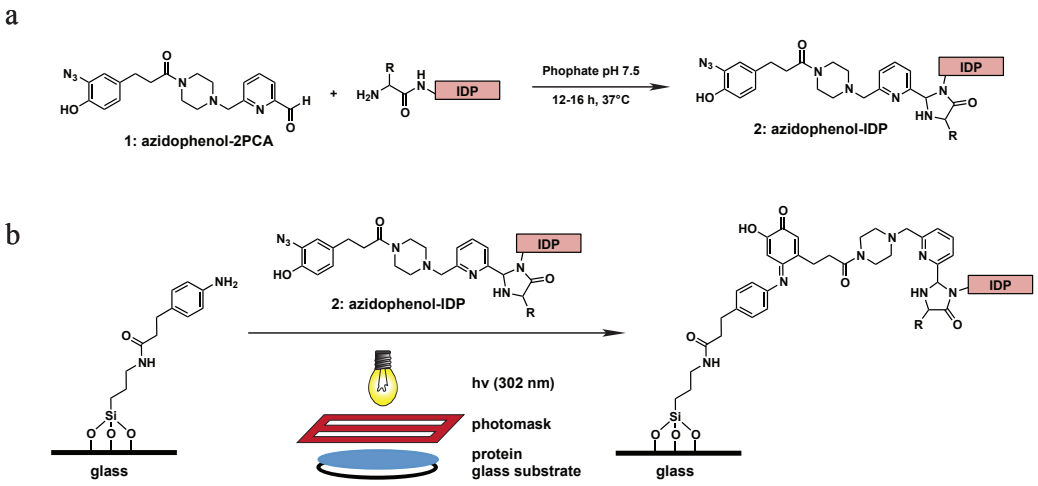


Figure 2.1. N-terminal specific photopatterning of intrinsically disordered proteins (IDP) on glass. (a) Schematic for the N-terminal specific conjugation of azidophenol-2PCA to IDP is shown. (b) Schematic for the photoactivated conjugation between azidophenol-IDP and aniline functionalized glass substrates is shown.

2.1.2. Site-Specific Protein Photopatterning Techniques

Currently, there are numerous protein patterning methodologies,[19-21] which typically employ microcontact printing,[22,23] photolithography,[24,25] dip-pen nanolithography,[26,27] electron beam lithography,[28,29] or inkjet printing,[30,31] but they are not site-specific. In order to pattern IDPs for the generation of uniform polymer brushes with maximal conformational flexibility, it is important to develop patterning chemistry that is specific to the N- or C- terminus. Chemoselective immobilization techniques have been coupled with aforementioned patterning techniques for site-selective protein patterning.[32-36] These techniques, however, require protein engineering, such as optimization of N-terminal residues, installation of peptide fusion tags or unique chemical handles. Furthermore, these patterning techniques are not only time-consuming, but they also demand expensive equipment.

With the goal of developing a site-specific protein patterning strategy without the need for protein engineering or special equipment, we utilized a photoinitiated biocompatible coupling reaction between azidophenols and anilines.[37] Our approach involves the covalent conjugation between IDP, which is labeled with azidophenol on the N-terminus, and an aniline, which is functionalized on a glass substrate (Figure 2.1). In this chapter, we describe the four steps in the development of this novel protein patterning method: 1) cloning, expression and purification of IDP, 2) selective modification of the N-terminus with azidophenol, 3) UV-irradiation of desired regions for covalent protein conjugation on aniline-glass, and 4) analysis of patterned substrates to evaluate photopatterning efficacy.

2.2. Results and Discussion

2.2.1. Cloning and Expression of IDP

We engineered an intrinsically disordered protein (IDP) construct containing the sidearm domain (residues 431-1058) of the rat neurofilament heavy protein (NFH, Protein Data Bank identification code F1LRZ7, see Appendix A for the preparation of a truncated NFH-SA as a drug

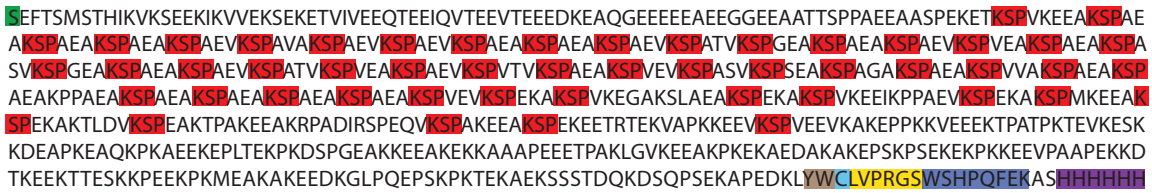


Figure 2.2. Engineered intrinsically disordered protein (IDP) sequence derived from disordered C-terminal sidearm domain of the heavy subunit of the neurofilament complex. This IDP contains 50 KSP repeats (red) which can be phosphorylated. N-terminal serine (green) or proline and cysteine (light blue) near the C-terminus are introduced as conjugation handles. Tyrosine and tryptophan (brown) were introduced for protein quantification by absorbance at 280 nm. Strep-tag (blue) and 6xHistidine-tag (purple) were introduced to facilitate protein purification. Thrombin cleavage site (yellow) was introduced for cleavage of affinity tags.

delivery agent). The C-terminus of the designed sequence includes a tyrosine and a tryptophan for the determination of protein concentration by absorbance at 280 nm, a unique cysteine for site-specific labeling, a Strep tag and a 6xHis tag for protein purification, as well as a thrombin cleavage site for the removal of affinity tags (see Figure 2.2 and Section 2.4.2 for details of cloning). This amphiphilic protein is highly charged with 141 negatively charged residues and 133 positively charged residues at neutral pH, resulting in a theoretical pI of ca. 6 including affinity tags. A pEcoli-Cterm 6xHN expression vector was modified to add tags before cloning in the neurofilament heavy side-arm (NFH-SA) gene. To explore various N-terminal conjugation chemistries, two constructs, one with N-terminal serine and the other with N-terminal proline, were made via site-directed mutagenesis. Notably, protein expression levels were low or not observed until the sequence between the ribosome binding site and the start codon was shorten to the original sequence in the pEcoli-Cterm 6xHN vector. The construct of a previously reported recombinant expression of NFH-SA has four cysteine residues in the N-terminus for covalent surface immobilization and a poly-His-Asn tag at the C-terminus for purification.[17] In contrast, the current constructs have a unique cysteine near the C-terminus, both Strep tag and a 6xHis tag for protein purification, as well as aromatic residues.

2.2.2. Purification and Characterization of IDP

A previous report of the purification of poly-His-Asn tagged NFH-SA utilized a Ni-NTA resin.[17] The Ni-NTA purified samples needed further purification by ion exchange chromatography due to a lower molecular weight impurity's affinity for Ni-NTA. Strep-Tactin affinity purification of Strep-tagged proteins has been shown to result in purer protein compared with Ni-NTA purification of 6xHis-tagged proteins,[38,39] so we used a Strep-Tactin column for the purification of the Strep-tagged IDP. Notably, Coomassie stained SDS-PAGE gel shows that the 70 kDa NFH-SA band ran similarly to the 120 kDa marker in the protein ladder. As previous reports of similar observations have suggested, this is likely due to the lower affinity of highly charged IDPs for sodium dodecyl sulfate (SDS) and their extended conformations.[3] Comparable to previously reported Ni-NTA affinity purification of 6xHis tagged NFH-SA, NFH-SA was not only in the elution fractions, but also in the flow through and wash fractions (Figure 2.3a). Increasing resin volumes of Ni-NTA or Strep-Tactin from 0.4 to 1.0 mL did not increase protein isolation yields, suggesting that protein loading did not exceed resin capacity (Figure 2.3b,c). One hypothesis is that the IDP is so dynamic and flexible that it has large excluded volume and therefore prevents interactions with the resin.[40,41] Despite the low yield, IDP of high purity was eluted from the Strep-Tactin column without the need for further purification. The identities of IDP with N-terminal serine or proline were confirmed by

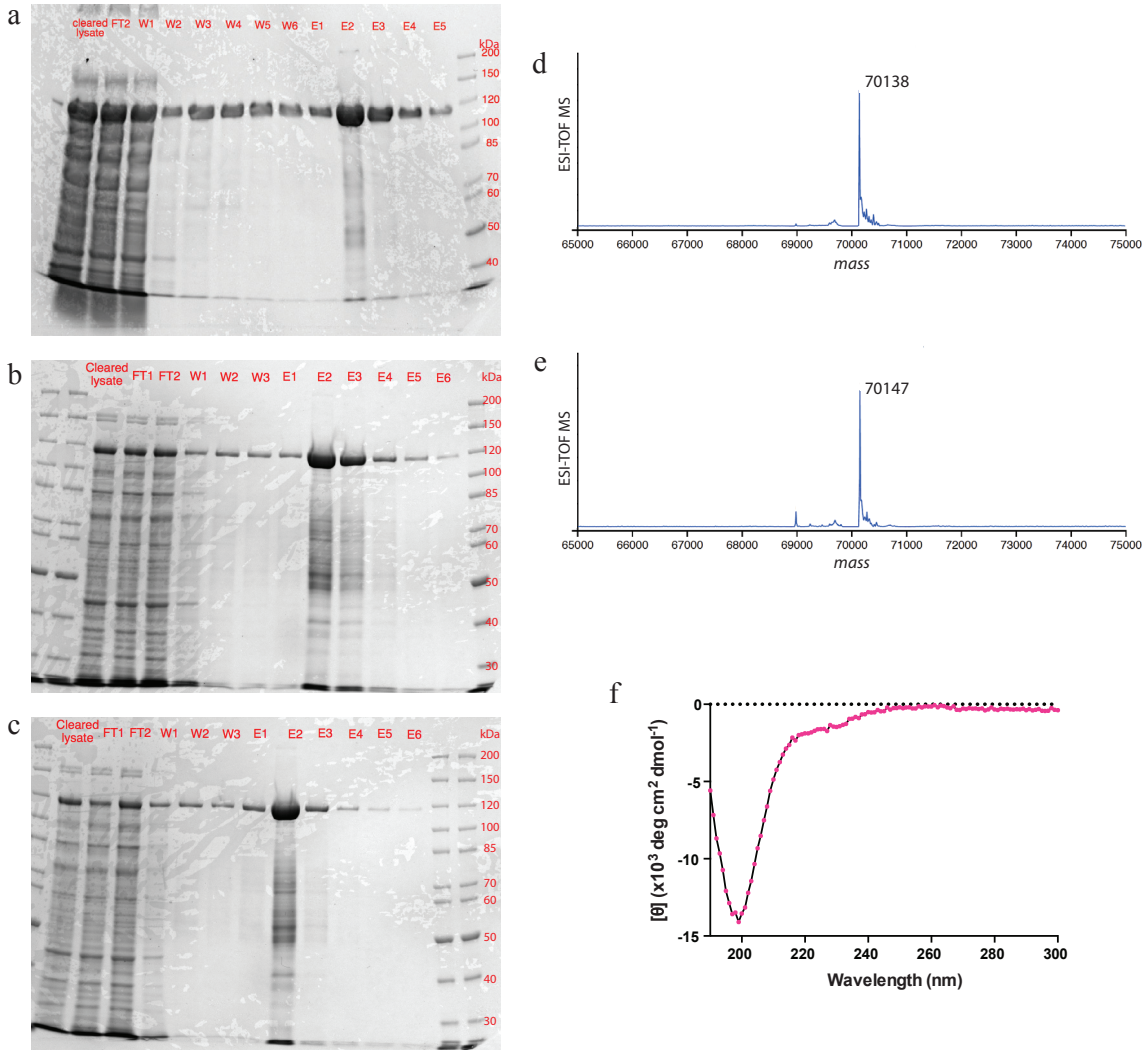


Figure 2.3. Purification and characterization of IDP. (a) Coomassie stained SDS-PAGE image is shown for large-scale purification of IDP (N-terminal serine) using a Strep-Tactin cartridge. Cleared cell lysate was loaded onto the cartridge, which was subsequently washed with buffer without *d*-desthiobiotin (FT: flow through; W1-W6: wash fractions 1-6). IDP was eluted (E1-E6) with buffer containing 5 mM *d*-desthiobiotin. Coomassie stained SDS-PAGE images are shown for small-scale gravitational Strep-Tactin column purification of IDP (N-terminal proline) using 0.4 mL (b) or 1.0 mL (c) Strep-Tactin resin. The same amount of lysate is loaded onto each of the two columns. Deconvoluted ESI-TOF MS spectra of IDP with N-terminal serine (d) and N-terminal proline (e) are shown. Expected masses of IDP: 70140 Da (with N-terminal serine) and 70150 (with N-terminal proline). (f) Far-ultraviolet circular dichroism spectrum of 1.7 μM IDP (N-terminal serine) in 10 mM pH 7.8 phosphate buffer is shown.

electrospray ionization time-of-flight mass spectrometry (ESI-TOF MS, Figure 2.3d,e). The far-UV circular dichroism (CD) spectrum of NFH-SA shows a minimum at ca. 200 nm, which is typical of an unfolded polypeptide chain with no secondary structures (Figure 2.3f).[42] We were able to isolate IDP with high purity using Strep-Tactin resin in one-step. Mass spectrometry and CD confirmed the protein identities and the lack of secondary structures.

2.2.3. Modification of IDP

After testing two N-terminal conjugation reactions with N-terminal serine and proline

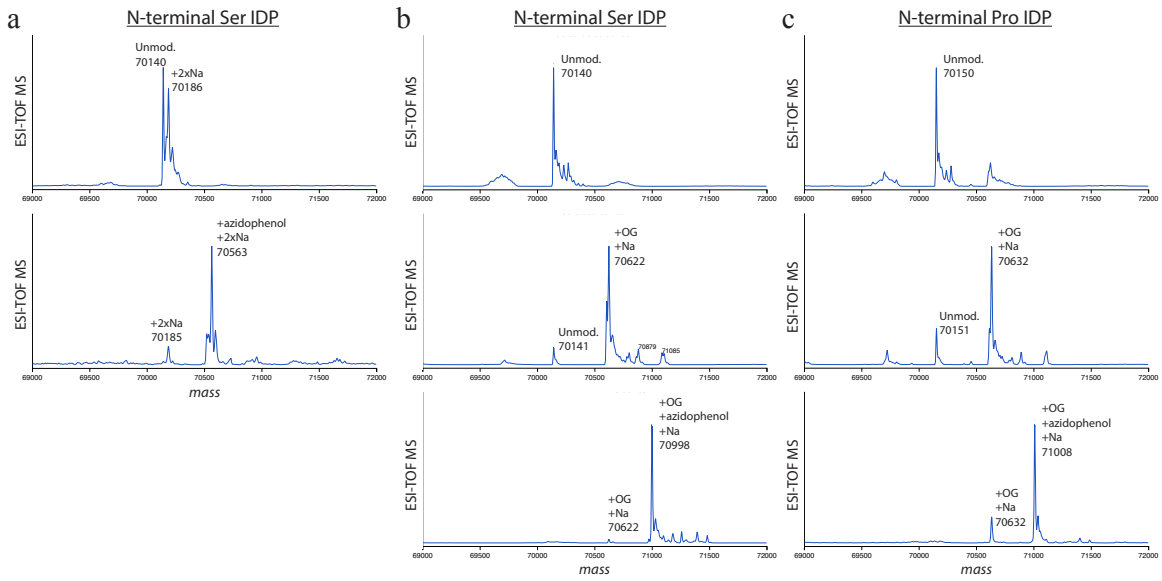


Figure 2.4. Deconvoluted ESI-TOF MS spectra showing small molecule conjugation reactions with IDP. (a) Unlabeled IDP with N-terminal serine was conjugated with azidophenol-2PCA **1**. IDP with N-terminal serine (b) and N-terminal proline (c) were first modified with Oregon Green 488 maleimide. After removal of excess dye, they were then reacted with azidophenol-2PCA **1** to afford fluorescently labeled azidophenol-IDP **2**.

IDP, 2-pyridinecarboxaldehyde (2PCA) was able to modify both N-termini in better yield than transamination using *N*-methylpyridinium-4-carboxaldehyde benzenesulfonate salt (Rapoport's salt).[43-45] We found that the 2PCA concentration for optimal modification is 5 mM for N-terminal serine and 10 mM for N-terminal proline. In previous studies, 2PCA modifies N-terminal proline to the lowest extent compared with all other amino acids, likely due to steric constraint of proline's rigid pyrrolidine side chain.[43] We synthesized an azidophenol-2PCA (**1**, Figure 2.1a) by coupling azidophenol NHS ester and piperazine-2PCA. Using azidophenol-2PCA, we were able to label IDP N-terminus with the azidophenol moiety (azidophenol-IDP) in high yields (Figure 2.4a). The unique cysteine near the C-terminus can be selectively labeled with fluorophores for fluorescence visualization or labeled with metal-bound chelators for quantitative analysis by laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS). Dual labeling of IDP was achieved by first incubating the protein with a fluorophore-maleimide or chelator-maleimide. After removal of excess maleimide and confirmation of cysteine labeling with mass spectrometry, singly labeled IDP was incubated in a solution of azidophenol-2PCA to yield doubly labeled IDP (Figures 2.4b,c and 2.13a). In the case of the Europium (Eu)-DOTA maleimide conjugation, a small peak corresponding to the addition of two maleimide molecules is present (Figure 2.13a). This is likely due to the maleimide reacting with one of the 129 lysine residues. Nevertheless, we were able to install single or double labels on IDP in good yields.

2.2.4. Photopatterning and Visualization of IDP on Aniline-Glass

Due to its ability to aggregate at high concentrations (as well as salt-free conditions, Figure 2.5e), IDP was photopatterned in solution between the photomask and aniline-glass at a concentration of no more than 50 μM. Aniline-glass was fabricated as described previously without modifications of the protocol.[37] To allow for visualization of protein patterns on glass substrates, fluorescent maleimides were conjugated to IDP either before (Figure 2.5a) or after (Figure 2.5b)

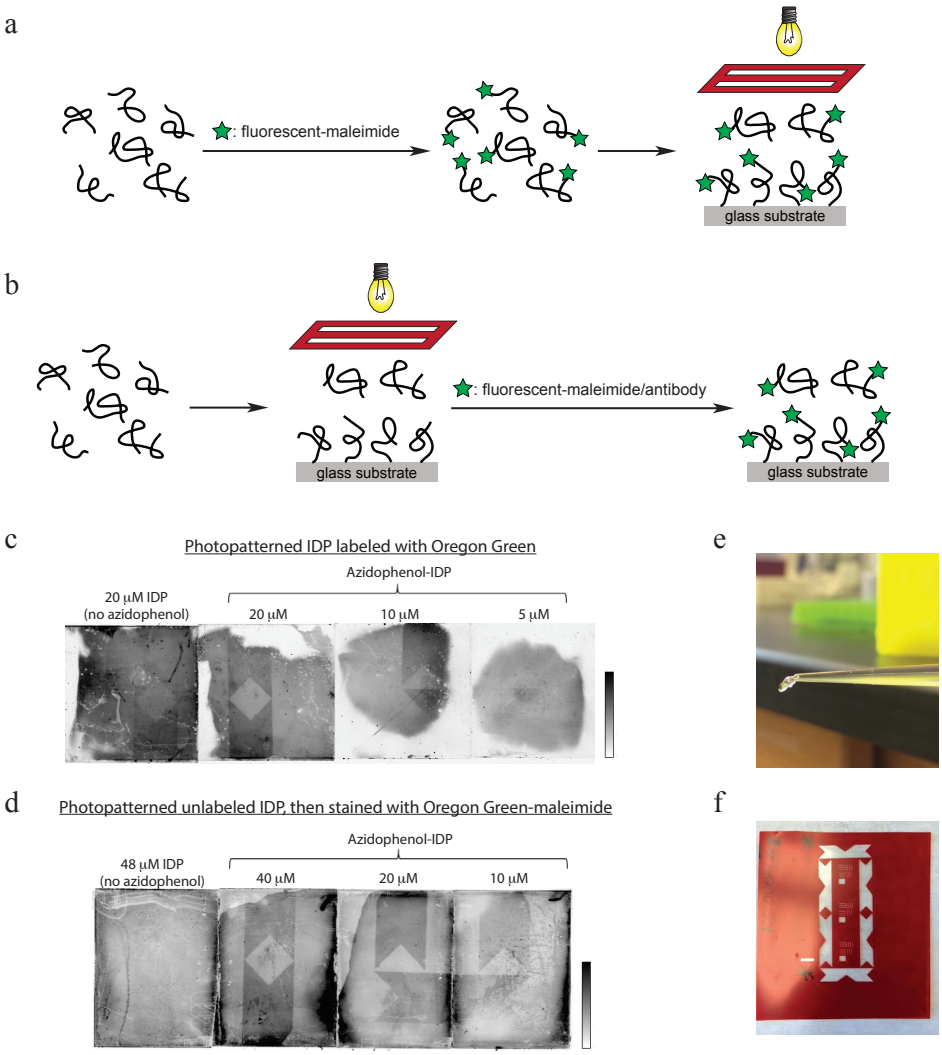


Figure 2.5. Photopatterning of azidophenol-IDP on aniline modified glass substrates. (a,b) To visualize patterned substrates, IDP can be fluorescently labeled before the photopatterning reaction (a) or patterned substrates can be stained with fluorescent maleimide or antibody after the photopatterning reaction (b). (c) Fluorescence image of photopatterning reactions between aniline glass and various concentrations of Oregon Green labeled azidophenol IDP visualized by method (a) are shown. (d) Fluorescence image of photopatterning reactions between aniline glass and various concentrations of unlabeled azidophenol IDP visualized by method (b) are shown. IDP was in 10 mM pH 7.8 phosphate buffer during irradiation. (e) IDP gelation at high concentrations and/or in water no salt. (f) The photomask (ca. 10 x 10 cm) used in this study is shown.

patterning to label its unique cysteine. When IDP was fluorescently labeled before UV irradiation, even though azidophenol-IDP was successfully patterned in a concentration dependent manner, Oregon Green labeled IDP without the azidophenol moiety was also patterned to a similar extent (Figure 2.5c). A possible explanation is that UV irradiation of Oregon Green, which has significant absorbance at the UV range, results in nonspecific radical reaction between the dye and aniline. When unlabeled IDP was first UV irradiated and patterned on glass substrates, which was subsequently stained with Oregon Green maleimide, the IDP control without azidophenol shows minimal pattern (Figure 2.5d). This supports the hypothesis that when pre-labeled, the Oregon Green moiety rather than the azidophenol moiety was responsible for patterning (Figure 2.5c).

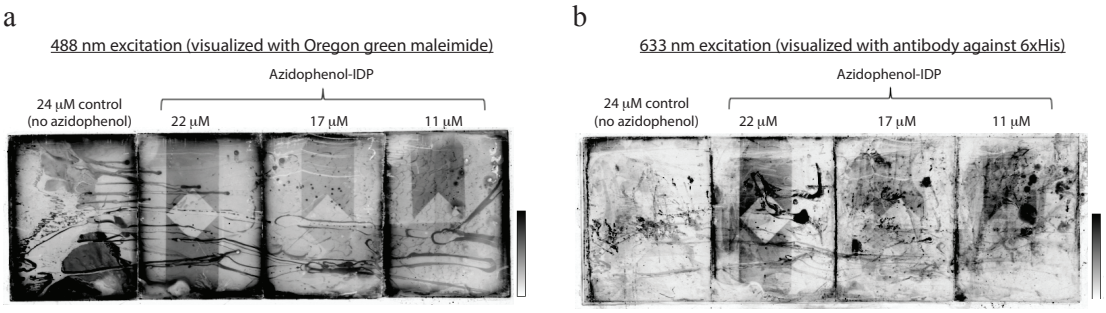


Figure 2.6. Visualization methods of photopatterning reactions. (a) Fluorescence image of substrates using 488 nm excitation to visualize fluorescence signal from Oregon Green maleimide is shown. (b) The same substrates were immunostained and imaged using 633 nm excitation to visualize fluorescence signal from antibody against 6xHis tag.

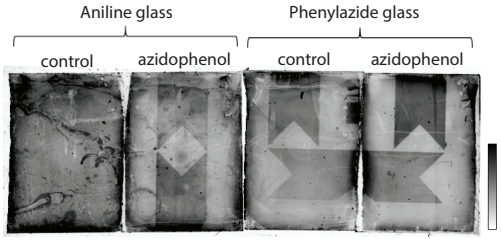


Figure 2.7. Comparison of photopatterning on aniline and phenylazide modified glass substrates. Fluorescence image of 22 μM unlabeled IDP with and without azidophenol in 50 mM pH 7.5 phosphate buffer irradiated on aniline or azide modified glass is shown. Glass substrates were stained with 4 mM Oregon Green maleimide for fluorescence visualization.

Because UV irradiation of unlabeled IDP leads to azidophenol-IDP patterning in a concentration dependent manner with minimal patterning of the control IDP, we used this method in subsequent studies.

We compared two methods of fluorescent visualization of surface-immobilized IDP: 1) direct conjugation of IDP to a small molecule fluorophore (Figure 2.6a) and 2) immunofluorescence using a primary antibody against 6xHis tag and a fluorescent secondary antibody (Figure 2.6b). Both methods confirm the expected patterns, but the immunofluorescence image shows unpredictable fluorescent patches, which may arise from nonspecific adhesion and aggregation of antibodies. We therefore used small molecule fluorescent maleimide-based labeling for substrate visualization in subsequent studies.

While testing various control experiments for the photopatterning reaction, we compared aniline-glass and phenylazide-glass without tris(2-carboxyethyl)phosphine (TCEP) reduction of phenylazide (Figure 2.7). Notably, IDPs with and without azidophenol moiety were both patterned onto phenylazide-glass. Upon UV irradiation, phenylazide is known to generate a reactive nitrene, which can participate in a variety of nonspecific chemistries via insertion into C—H, O—H, or N—H bonds on proteins.[46] This explains why proteins can be nonspecifically photopatterned onto phenylazide glass.

2.2.5. Optimization of Photopatterning Reaction Conditions

After confirming photopatterning of IDP, we optimized the reaction by screening salt concentrations, buffers, pH and UV irradiation time. The patterning reaction offers best contrast

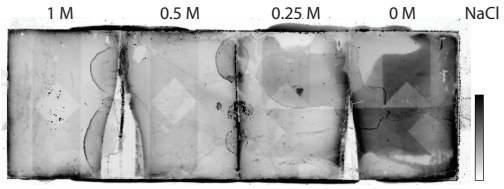


Figure 2.8. Effects of salt concentration on photopatterning reaction between azidophenol-IDP and aniline glass substrates. Azidophenol-IDP (20 μM) in 10 mM pH 7.8 phosphate buffer was irradiated through a mask. Glass substrates were then stained with 3 mM Oregon Green maleimide and imaged.

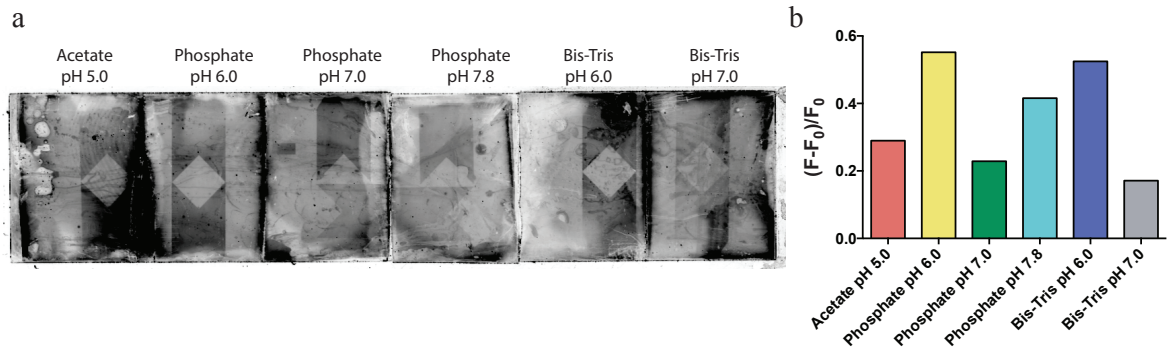


Figure 2.9. Effects of buffer and pH on photopatterning reaction between azidophenol-IDP and aniline glass substrates. (a) Azidophenol-IDP (10 μM) in 50 mM of various buffers were irradiated through mask. Glass substrates were then stained with 4 mM Oregon Green maleimide and imaged. (b) Fluorescence signal above background was quantified by fluorescence intensity of irradiated area (F) subtracted by fluorescence intensity of masked area (F_0) and divided by F_0.

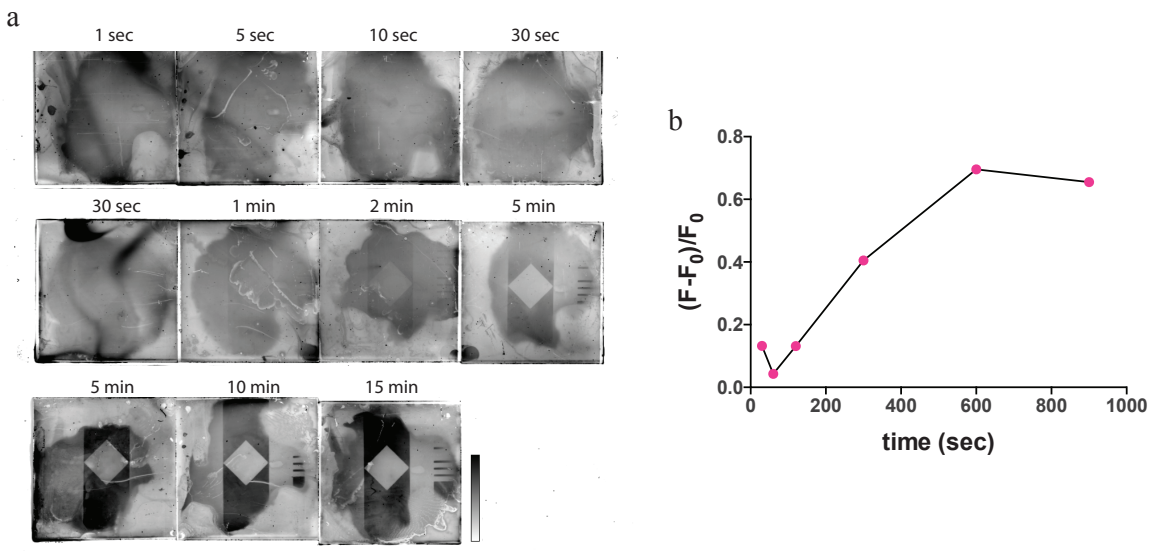


Figure 2.10. Time course of photopatterning reaction between azidophenol-IDP and aniline glass substrates. (a) Fluorescence images of 10 μM unlabeled IDP in 50 mM pH 6.0 phosphate buffer irradiated for various durations on aniline glass substrates is shown. Glass substrates were stained with 4 mM Oregon Green maleimide for fluorescence visualization. (b) Fluorescence signal above background was quantified by fluorescence intensity of the irradiated area (F) subtracted by the fluorescence intensity of masked area (F_0) and divided by F_0. Data from duplicate samples (30 sec and 5 min) were averaged. Irradiation times shorter than 30 sec were not quantified due to the lack of signal above background.

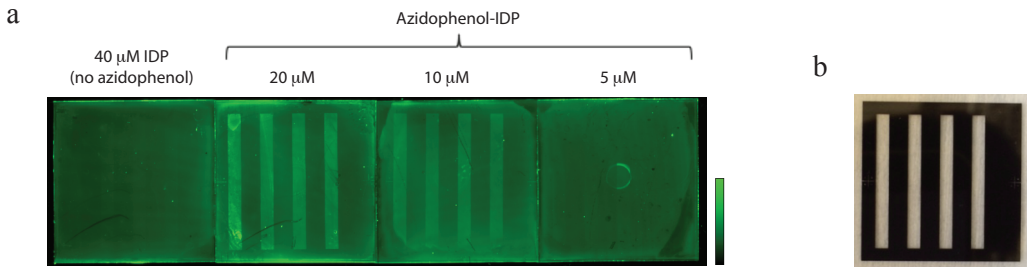


Figure 2.11. Optimized photopatterning reaction of azidophenol-IDP on aniline modified glass substrates. (a) Fluorescence image of photopatterning reactions between aniline glass and various concentrations of unlabeled azidophenol IDP in 50 mM pH 6.0 phosphate buffer are shown. The reaction was irradiated for 10 min and the glass substrates were stained with 4 mM Oregon Green maleimide before imaging. (b) The photomask used in this study is shown. Each transparent (white) bar on the photomask is 21 x 2 mm.

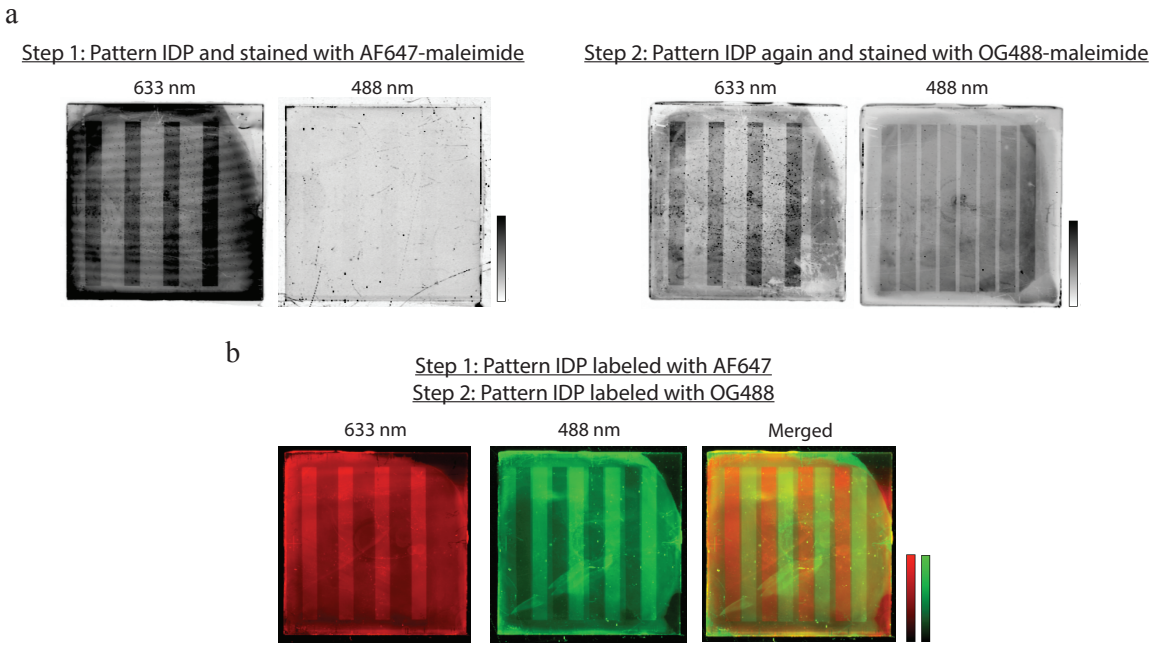


Figure 2.12. Dual protein photopatterning. (a) Unlabeled IDP (20 μM) in 50 mM pH 6.0 phosphate buffer was first patterned (10 min irradiation) on aniline glass substrate, which was then stained with Alexa Fluor 647 (AF647) maleimide and imaged. Unlabeled IDP was photopatterned again with a complementary photomask (offset to the right of the first photomask) using the same conditions as above. The substrate was then stained with Oregon Green 488 (OG488) maleimide and imaged with 633 nm and 488 nm excitation. (b) Using the same reaction conditions above, AF647 labeled azidophenol-IDP was first photopatterned on an aniline glass substrate. OG488 labeled azidophenol-IDP was then patterned and imaged.

at low salt concentration (no addition of NaCl) compared with 0.25, 0.5 and 1 M NaCl, so no salt was added in subsequent studies (Figure 2.8). Fluorescence quantification of photopatterning performed in different buffers at different pH values indicates that pH 6.0 (phosphate or Bis-Tris buffer) is the optimal pH for this reaction (Figure 2.9), consistent with previous reports using this reaction.[37] Therefore, the patterning reactions were done in pH 6.0 phosphate buffer in subsequent experiments. Finally, the dependence of photopatterning efficacy on UV irradiation time was evaluated and quantified. Using a 6 W handheld UV lamp, patterns were not visible when irradiation was shorter than 30 sec. Beyond 30 sec, however, the patterning signal compared with background increased with increasing irradiation times until 10 min, when the signal plateaued (Figure 2.10). Hence, an irradiation time of 10 min was applied for subsequent studies. IDP photopatterning with

the optimized conditions and various IDP concentrations is shown in Figure 2.11.

2.2.6. Dual-Protein Patterning

With the validation of single protein patterning in hand, we next tested if we could pattern multiple proteins using this photolithography technique. As a proof of concept, we patterned the same IDP twice, but labeled them with different fluorophores. Using the optimized method above, we successfully patterned azidophenol-IDP and stained the patterned substrate with Alexa Fluor 647-maleimide. The pattern was observable with the 633 nm but not the 488 nm filter, showing that the fluorescence of Alexa Fluor 647 does not bleed into the 488 nm channel (Figure 2.12a,

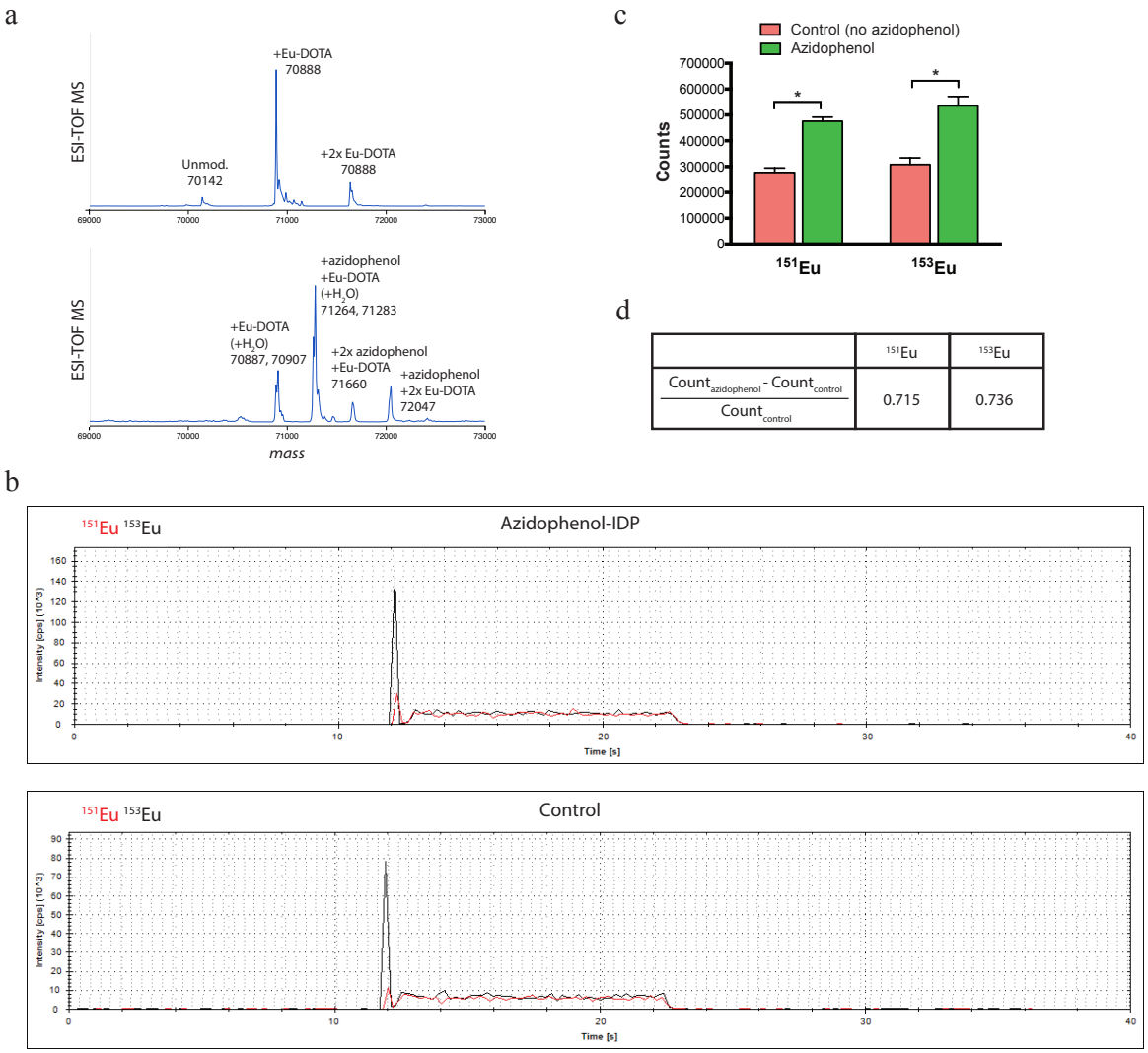


Figure 2.13. Quantification of photopatterned proteins on glass substrates using laser-ablation ICP-MS. (a) Deconvoluted ESI-TOF MS spectra are shown for IDP conjugated with Eu-DOTA-maleimide (top) and subsequently reacted with azidophenol-2PCA **1** to afford azidophenol-IDP labeled with Eu for ICP-MS analysis. (b-d) Glass substrates photopatterned with 20 μM IDP with and without azidophenol using optimized reaction conditions were analyzed by laser-ablation ICP-MS. (b) Representative ^{151}Eu (red) and ^{153}Eu (black) signal intensity traces of laser-ablation ICP-MS across glass substrates patterned using IDP with (top) and without (bottom) azidophenol. The scales of the y-axes are different. (c) Quantification of laser ablation ICP-MS data is shown. Bars represent mean ± 95% confidence intervals. $N = 5$ areas of ablation for each of the two samples. * denotes $p < 0.0001$ (Student's unpaired t test). (d) Quantification of ^{151}Eu and ^{153}Eu signal (IDP with azidophenol) above background (IDP without azidophenol).

left). A second round of azidophenol-IDP was then patterned using a different photomask with an offset to the right of the first photomask. After staining the patterned substrate with Oregon Green 488-maleimide, the 488 nm channel showed comparable fluorescence intensities in areas of the first and second patterned IDP (Figure 2.12a, right). These results can be explained by the nonspecific adhesion of excess OG-488 maleimide on the initially patterned IDP/fluorophore. To rule out the possibility that these results arise from nonspecific adhesion of excess dye, we patterned two rounds of azidophenol-IDP that is pre-labeled with different dyes. The resulting substrate shows two different colors of fluorescence corresponding to the two complementary photomasks used (Figure 2.12b). Even though the dual-patterning experiment with prelabled IDPs did result in a substrate with expected fluorescence patterns, more control experiments needed to be done to verify that the photoreaction was specific to the azidophenol moiety instead of photoreactivity of the dyes on IDP, as mentioned in the single protein patterning experiment.

2.2.7. Quantitative Analysis of Surface-Patterned IDP

We have relied on fluorescence to evaluate the efficacy of the photolithography technique thus far, but fluorescence intensity is not ideal for surface quantification because it is subject to quenching with the substrate surface and dyes in close proximity, which depends on the conformation of immobilized IDP (extended brush or collapsed mushroom) as well as the surface density of IDP. To evaluate the density of immobilized IDP independent of fluorescence, we utilized laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS), which is a highly sensitive surface chemical analysis technique with detection limit typically in the parts per billion (ppb) range for rare earth metals.[47,48] We labeled IDP with Eu-chelated DOTA maleimide (Figure 2.13a and Section 2.2.3) and photopatterned it on aniline-glass. Using a laser beam focused on the patterned surface, surface immobilized Eu was ablated into fine particles, which were transported to an ICP-MS instrument for the analysis of signal intensity from two Eu isotopes ^{151}Eu and ^{153}Eu (Figure 2.13b). The signals from substrates patterned with IDP with and without azidophenol were quantified by

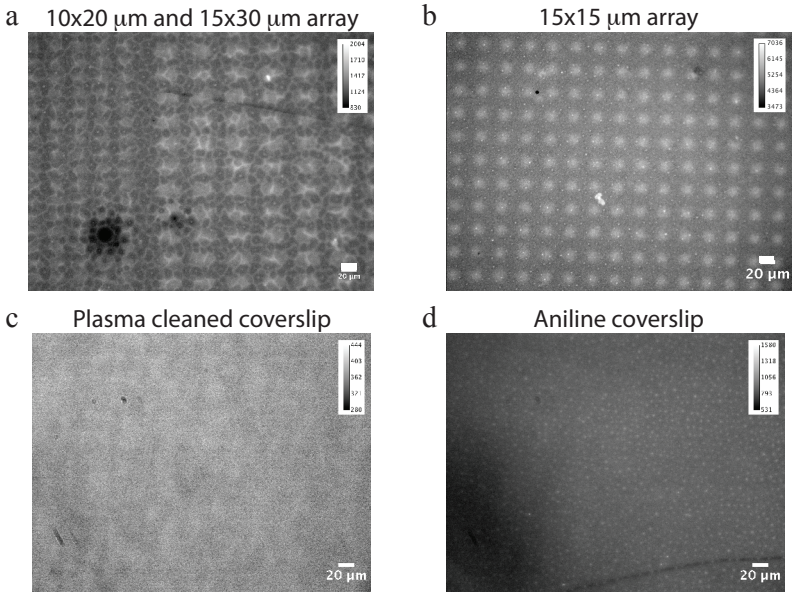


Figure 2.14. High-magnification fluorescence characterization of glass substrates. (a, b) Epifluorescence images are shown for glass substrates photopatterned with 20 μM azidophenol-IDP, then stained with 4 mM Oregon Green 488 maleimide. (c, d) Plasma cleaned and aniline modified glass substrates were also stained with 4 mM Oregon Green 488 maleimide and imaged. Scale bars = 20 μm.

separately integrating the signals from ^{151}Eu and ^{153}Eu intensity traces (Figure 2.13c,d). Due to the lack of appropriate matrix-matched standards for calibration, absolute quantification of Eu surface density was not possible, so relative quantification compared with control from IDP without azidophenol was determined. Eu signal above background for both isotopes were ca. 0.7, which agrees well with fluorescence quantification (Figure 2.10b). The low signal above background may also explain why we were unable to detect protein patterns using atomic force microcopy.

2.2.8. Fluorescence Microscopy Analysis of Glass Substrates

A signal to background ratio of 0.7 needs to be improved for useful applications of this protein patterning technique. To investigate the reasons for such low signal to background, we used a mask with features in the micron range to pattern IDP and an epifluorescence microscope for imaging. Patterned substrates did indeed show expected fluorescent array from the photomask used. The fluorescence was not even within the UV irradiated or masked areas, however, implying the presence of surface defects (Figure 2.14a,b). Aniline-functionalized surface stained with Oregon Green also showed such uneven fluorescence while plasma cleaned surface does not, suggesting that surface defects arise from aniline functionalization (Figure 2.14c,d). Because low patterning signal to background may be due to surface defects, improving the aniline functionalization step may improve the efficacy of this novel patterning technique. Low patterning signal to background may also be due to nonspecific adsorption of proteins onto the glass substrate, which is a common problem in protein patterning methods. A more quantitative technique, such as protein radiolabeling with ^{125}I, is required for determining the absolute surface density, which will indicate whether the patterned protein density is low or the nonspecific absorption of the IDP is high.

2.3. Conclusions

In this chapter, we developed a site-specific protein photolithographic patterning technique using the reaction between light-activated azidophenols and aniline functionalized glass. This novel bench-top method does not require expensive equipment and takes only 30 min per pattern. IDP from the neurofilament heavy protein was engineered and its overexpression was optimized. Protein purification to high purity was achieved using a single affinity chromatographic step. The N-terminal serine or proline of IDP was successfully labeled with the azidophenol moiety to good yields via 2PCA conjugation. We were able to label the engineered IDP in two specific locations for downstream analysis of the protein patterned substrates. Other than fluorescence analysis, we developed a LA-ICP-MS technique to quantify the relative density of immobilized proteins. Although IDP was successfully patterned, the signal to background needs to be improved for this technique to be useful. Future efforts will focus on the fabrication of high-density aniline functionalized glass with even surface, as well as preventing nonspecific protein absorption on the substrate.

2.4. Materials and Methods

2.4.1. General Procedures and Materials

Unless otherwise noted, all reagents and solvents used were of analytical grade and were used

as received from commercial sources. UV-visible spectrometry was performed using a NanoDrop 1000 (Thermo Scientific). Photomasks were either custom ordered (Mylar, Fineline Imaging) or made with a sticker cutter. Epifluorescence images were acquired using an inverted Nikon Eclipse Ti microscope equipped with a 20x air objective.

NMR spectra were recorded on a Bruker AVQ-400 spectrometer. ^{1}H NMR chemical shifts are reported as δ in units of parts per million (ppm) relative to $CDCl_3$ (δ 7.26). Multiplicities are reported as: s (singlet), br.s (broad singlet), d (doublet), t (triplet), q (quartet), dd (doublet of doublets), dt (doublet of triplets), br (broad) or m (multiplet). Coupling constants are reported as a *J* value in Hertz (Hz). ^{13}C NMR chemical shifts are reported as δ in units of parts per million (ppm) relative to $CDCl_3$ (δ 77.2).

High-resolution electrospray ionization mass spectrometry (HR-ESI-MS) data were obtained at the UC Berkeley QB3/Chemistry Mass Spectrometry Facility. Protein mass spectrometry data were collected on an Agilent 6224 TOF LC/MS equipped with a Turbospray ion source and connected to an Agilent 1260 Infinity Series HPLC system and a Poroshell 300SB-C18 column (Agilent Technologies). Protein mass reconstruction was performed on the charge ladder with Mass Hunter software (Agilent Technologies).

Sodium dodecyl sulfate-poly(acrylamide) gel electrophoresis (SDS-PAGE) analysis was performed using 7.5% Mini-PROTEAN TGX gels (Bio-Rad). SDS loading buffer containing β-mercaptoethanol was added to all electrophoresis samples, which were then boiled at 95 °C for 5 min to ensure complete denaturation and disulfide-bond reduction. Coomassie Brilliant Blue R-250 (Bio-Rad) and Gel Doc EZ System (Bio-Rad) were used to stain and image SDS-PAGE gels.

2.4.2. Cloning of IDP (NFH-SA)

The expression vector was constructed by modifying the multiple cloning site of pEcoli-Cterm 6xHN vector (Clontech, Cat. no. 631417) to introduce EcoRI and HindIII restriction enzyme recognition sequences, an N-terminal serine (after methionine cleavage), a cysteine near the C-terminus, aromatic residues, thrombin cleavage site and affinity tags. These features were introduced between the NcoI and NotI sites on the original pEcoli-Cterm 6xHN vector using the following forward primers:

VF1: 5'- CATGGGTTAAGAAGGAATG -3'

VF2: 5'- AGCGAATTCAGGCTTGTCGACAAGCT -3'

VF3: 5'- TTATTGGTGCCTGGTTCCGCGTGGC -3'

VF4: 5'- TCTTGGAGCCACCCGCAGTTCGAAAAAGC -3'

VF5: 5'- TAGCCATCACCATCACCATCACTAAGC -3'

and reverse primers:

VR1: 5'- CCAATTCTTCCTTACTCGCTTAAGTCC -3'

VR2: 5'- GAACAGCTGTTCGAAATAACCACGGACC -3'

VR3: 5'- AAGGCGCACCGAGAACCTCGGTGGGCGT -3'

VR4: 5'- CAAGCTTTTTCGATCGGTAGTGGTA -3'

VR5: 5'- GTGGTAGTGATTCGCCGG -3'.

The 5' ends of all primers, except for VF1 and VR5, were phosphorylated using T4 polynucleonide kinase (T4 PNK, NEB M0201S). After inactivation of T4 PNK, 50 pmol each of the phosporylated primers, VF1 and VR5 were mixed in a PCR tube, heated at 95 C° for 5 min and cooled to room temperature overnight to allow for annealing. Using T4 DNA ligase (NEB M0202T), the annealed DNA was inserted into purified pEcoli-Cterm 6xHN vector digested with NcoI-HF (NEB R3193S) and NotI (NEB R3189S). The resulting modified expression vector was then transformed into 5-α competent *Escherichia coli* cells (NEB, C2987H) and plated on LB plates with 100 μg/mL ampicillin. After plasmid isolation and sequence confirmation, the modified expression vector was digested with EcoRI-HF (NEB, R3101S) and HindIII-HF (NEB, R3104S) and subsequently purified. NFH-SA (IDP) gene, cloned in pClneo with a HindIII restriction enzyme site,[17,49] was excised by ApoI (NEB, R0566S) and HindIII-HF digestion. The purified gene insert was then ligated to the doubly digested modified expression vector using T4 DNA ligase. The resulting modified expression vector with the NFH-SA gene insert was then transformed into 5-α competent *Escherichia coli* cells. Sequences of the entire NFH-SA gene in plasmids isolated from the transformants were confirmed using the T7 promoter sequencing forward primer and following primers:

S1: 5'- CAGCTGAGGTCAAATCTCCAGTGGA -3'

S2: 5'- CCGTGAAGGAAGGTGCAAAATCCCT -3'

S3: 5'- GCCAAGAAGGAAGAGGCTAAAGAGA -3'.

A frameshift is needed for desired protein expression, so site-directed mutagenesis (QuikChange method) was performed to insert one base pair using PfuUltra II Fusion HS DNA Polymerase (Agilent, 600670), DpnI (NEB, R0176L), dNTP mix (NEB, N0447L), XL1-Blue competent *Escherichia coli* cells (QB3 MacroLab) and the following primers:

forward: 5'- GGAATGAGCGAATTC<u>A</u>CCTCCATGTCCACTC -3'

reverse: 5'- GAGTGGACATGGAGG<u>T</u>GAATTCGCTCATTCC -3'.

The sequence between the ribosome binding site and the start codon needed to be optimized to increase protein expression levels. This was achieved using sequential Quikchange mutagenesis with the following primers:

forward: 5'- GAAGGAGATATACCGTTAAGAATGAGCGAATTCACC -3'

reverse: 5'- GGTGAATTCGCTCATTCTTAACGGTATATCTCCTTC -3'

and

forward: 5'- CTTTAAGAAGGAGATATACCATGAGCGAATTCACCTCC -3'

reverse: 5'- GGAGGTGAATTCGCTCATGGTATATCTCCTTCTTAAAG -3'.

IDP with N-terminal proline was also constructed using Quikchange mutagenesis with the following primers:

forward: 5'- AAGAAGGAGATATACCATGCCCGAATTCACCTCCATGTCC -3'

reverse: 5'- GGACATGGAGGTGAATTCGGGCATGGTATATCTCCTTCTT -3'.

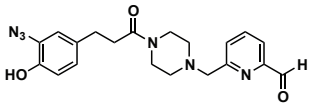
2.4.3. Expression and Purification of IDP

For protein expression, plasmids were transformed into Rosetta (DE3) competent *Escherichia coli* strain (Novagen, 70954) and plated on LB plates with 100 μg/mL ampicillin and 34 μg/mL chloroamphenicol. A single colony or frozen cell stock was grown overnight in LB with antibiotics and used to inoculate LB at a 1:100 dilution. The culture was shaken at 37 °C and was induced by addition of 0.75 mM (final concentration) isopropyl β-D-1-thiogalactopyranoside (IPTG) when its OD_{600} reached 0.5-0.6. The cells were harvested 3-5 h after induction by centrifugation at 5000 rpm for 15 min and flash frozen for storage. Cell pellet from a 2 L culture was resuspended in 15 mL of lysis buffer (100 mM pH 8.0 Tris, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, 1 mM PMSF) and lysed by sonication. Cell debris was removed by centrifugation at 14000 rpm for 30 min at 4 °C. For large-scale purification, the clarified lysate was then applied to a Strep-Tactin cartridge (IBA, 2-1234-001) using an Akta Pure M FPLC system (GE Healthcare) at 4 °C. After protein injection, the column was washed with wash buffer (100 mM pH 8.0 Tris, 150 mM NaCl, 1 mM EDTA) and Strep-tagged IDP was eluted with wash buffer containing 5 mM *d*-desthiobiotin (Sigma). Flow through and wash fractions containing IDP were re-injected on the Strep-Tactin column to increase the yield of purified IDP. The elution fractions were combined and then buffer exchanged into 10 mM pH 7.8 phosphate buffer or 10 mM pH. 6.5 Bis Tris buffer using spin concentrators with a 10 kDa molecular weight cut-off (Amicon Ultra, EMD Millipore) and/or a NAP-5 Sephadex size exclusion column (GE Healthcare). Protein concentrations were determined using the absorbance at 280 nm and the theoretical extinction coefficient 12490 $M^{-1}cm^{-1}$. This protocol yielded ca. 5 mg pure IDP per liter culture.

2.4.4. Circular Dichroism

CD spectrum was recorded from 190 to 250 nm on a Jasco-720 spectropolarimeter at 25 °C using a cuvette with 1 mm path length. The signal of 1.7 μM IDP with N-terminal serine in 10 mM pH 7.8 potassium phosphate buffer was buffer subtracted.

2.4.5. Synthesis of Azidophenol-2PCA



6-((4-(3-(3-azido-4-hydroxyphenyl)propanoyl)piperazin-1-yl)methyl)picolinaldehyde (1). *N*-hydroxysuccinimidyl-3-(4-hydroxy-3-azidophenyl)propanoate (azidophenol-NHS ester)[37] and 6-(piperazin-1-ylmethyl)-2-pyridinecarboxaldehyde (piperazine-2PCA)[43] were synthesized as previously reported. Azidophenol-NHS ester (100 mg, 0.33 mmol, 1.0 equiv.) and piperazine-2PCA (109 mg, 0.39 mmol, 1.2 equiv.) were added to a 20 mL scintillation vial. DMF (2 mL) was then added to the vial to dissolve solid reagents. Upon addition of triethylamine (0.136 mL, 1.17 mmol, 3.5 equiv.), the solution color instantly changed from pale yellow to dark brown with precipitate. After stirring the reaction mixture at room temperature in the dark for 2 h, DMF was removed by rotary evaporation. The resulting residue was dissolved in dichloromethane and washed with a saturated solution of sodium bicarbonate (2x), brine and dried with Na_2SO_4. The solvent was

removed under reduced pressure. Purification by flash column chromatography (silica gel, 3% methanol in dichloromethane) yielded product **1** as a sticky off white solid (92 mg, 0.23 mmol, 71%). ^{1}H NMR (400 MHz, chloroform-*d*) δ 10.05 (s, 1H), 7.87 (d, *J* = 4.5 Hz, 2H), 7.78-7.59 (m, 1H), 6.97-6.72 (m, 3H), 6.32 (br.s, 1H), 3.79 (s, 2H), 3.69 (s, 2H), 3.47 (t, *J* = 5.1 Hz, 2H), 2.90 (t, *J* = 7.7 Hz, 2H), 2.60 (t, *J* = 7.7 Hz, 2H), 2.50 (dt, *J* = 22.9, 5.2 Hz, 4H). ^{13}C NMR (101 MHz, chloroform-*d*) δ 193.4, 170.6, 152.4, 146.2, 137.6, 133.9, 127.6, 125.9, 120.6, 118.6, 116.2, 63.7, 53.2, 52.9, 45.3, 41.4, 34.9, 30.6. HRMS (ESI) calculated for $C_{20}H_{23}N_6O_3$ ([M+H]$^+$) 395.1826, found 395.1822. Using a NanoDrop, the extinction coefficient at 280 nm was determined to be 4290 $M^{-1}cm^{-1}$, which was used to quantify azidophenol-IDP conjugate sample concentrations.

2.4.6. Protein Modification

Procedure for cysteine alkylation with fluorescent maleimides. A solution of IDP (20 μM) in 50 mM pH 7.5 phosphate buffer was incubated with 10 equiv. of Oregon Green 488 maleimide (Thermo Fisher O6034) or Alexa Fluor 647 maleimide (Thermo Fisher A20347) for 1 h at room temperature. Excess dye was removed using a spin concentrator with a 10 kDa molecular weight cut-off.

Procedure for cysteine alkylation with Eu-DOTA maleimide. $EuCl_3$ in H_2O (12 mM final concentration, 1.0 equiv.) was added to a solution of DOTA-GA maleimide (CheMatech, C117, 1.05 equiv.) in 10 mM pH 4.5 acetate buffer. The reaction was incubated at room temperature for 15 min. A solution of IDP (70 μM final concentration, 1.0 equiv.) in 50 mM pH 6.0 Bis-Tris buffer was incubated with the crude Eu-chelated DOTA maleimide solution (30 equiv.). Observed gelated IDP was removed with a 0.22 μm centrifugal filter (EMD Millipore) before removing excess maleimide with a spin concentrator with a 10 kDa molecular weight cut-off.

Procedure for azidophenol-2PCA conjugation. A solution of IDP (20 μM) in 50 mM pH 7.5 phosphate buffer was incubated at 37 °C for 12-14 h with 5 mM or 10 mM azidophenol-2PCA for N-terminal Serine and N-terminal proline IDP respectively. Undissolved particulates were then removed using a 0.22 μm centrifugal filter and the unreacted azidophenol-2PCA was removed with a spin concentrator with a 10 kDa molecular weight cut-off.

2.4.7. IDP Photopatterning on Glass Surfaces

Aniline-glass substrates were prepared using glass slides or coverslips (Fisherbrand, plain, 25 x 75 x 1 mm or 25 x 25 mm No. 2), 3-(4-azidophenyl)-N-(3-trimethoxysilylpropyl) propanamide (azidophenyl silane) and TCEP for the reduction azidophenyl moiety to aniline as described previously.[37] Aniline-glass slides were cut into quarters for photopatterning experiments. IDP solution (7 μL) was sandwiched between aniline-glass and a photomask and irradiated using a 6 W handheld 302 nm (UVB) UV lamps (UVP, USA) placed ca. 1 cm above the photomask for 5 min on a bench top. After UV exposure, the photomask was removed and the patterned glass substrate was placed on a coverslip rack inside a glass beaker. The beaker was then filled with 0.4% SDS solution and was shaken on the orbital shaker for 10 min. The glass substrate was then rinsed in PBS for 5 min and H_2O briefly before it was dried with a stream of N_2.

2.4.8. Analysis of Patterned Substrates

Visualization of patterned substrates with fluorphore maleimides. To visualize the

patterned substrates by fluorescence, 7 μL of 4 mM fluorescent maleimide in 50 mM pH 7.5 phosphate buffer was sandwiched between two patterned glass substrates. After 1 h of incubation at room temperature in the dark, the substrates were rinsed and dried as described above. Alternatively, patterned IDP can be visualized by immunofluorescence staining using antibody against 6xHis tag. Patterned substrates were first blocked with 1% bovine serum albumin (BSA) in PBS for 1 h. They were then incubated in 6xHis epitope tag antibody (1:500 dilution in 1% BSA, Pierce MA1-21315) for 30 min. After rinsing in PBS 5 times for 5 min each, the substrates were blocked again with 1% BSA for 30 min and incubated in goat anti-mouse secondary antibody labeled with Alexa Fluor 633 (1:1000 dilution in 1% BSA, Thermo Fisher A-21052) for 1 h. The substrates were finally rinsed in PBS 5 times for 5 min each and briefly in water. All immunostaining incubation steps were done at room temperature with shaking. Fluorescent images of pattern substrates were collected using a Typhoon 9410 variable mode imager (Amersham Biosciences) with a thin layer of H_2O between the patterned surface and the scanner. Fluorescence intensities were quantified with ImageJ software (NIH) by averaging three or more UV exposed or unexposed areas on each substrate.

Quantification of immobilized IDP using LA-ICP-MS. Eu-DOTA labeled IDP with or without azidophenol was patterned on aniline glass substrates as described above. The LA-ICP-MS analysis was conducted by the NWR213 laser ablation system (ESI) coupled to a iCAP Q ICP-MS system. Laser ablation of the glass surface was performed using a focused Nd:YAG laser beam (wavelength 213 nm, repetition frequency 20 Hz, laser spot diameter 100 μm, scanning speed 50 μm/s, laser frequency 2.00 J/cm^2) with dwell time of 0.1 s. The relative amounts of ^{151}Eu and ^{153}Eu were quantified by integrating signal intensities for 10 s, excluding initial spike.

2.5. References

1. Wright, P. E., and Dyson, H. J. (1999) Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm, *J. Mol. Biol. 293*, 321-331.

2. Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., Oldfield, C. J., Campen, A. M., Ratliff, C. R., Hipps, K. W., Ausio, J., Nissen, M. S., Reeves, R., Kang, C. H., Kissinger, C. R., Bailey, R. W., Griswold, M. D., Chiu, M., Garner, E. C., and Obradovic, Z. (2001) Intrinsically disordered protein, *J. Mol. Graph. Model. 19*, 26-59.

3. Habchi, J., Tompa, P., Longhi, S., and Uversky, V. N. (2014) Introducing protein intrinsic disorder, *Chem. Rev. 114*, 6561-6588.

4. Tompa, P. (2012) Intrinsically disordered proteins: a 10-year recap, *Trends Biochem. Sci. 37*, 509-516.

5. Oldfield, C. J., and Dunker, A. K. (2014) Intrinsically disordered proteins and intrinsically disordered protein regions, *Annu. Rev. Biochem. 83*, 553-584.

6. Uversky, V. N., Gillespie, J. R., and Fink, A. L. (2000) Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins: Struct., Funct., Genet. 41*, 415-427.

7. Lobanov, M. Y., and Galzitskaya, O. V. (2015) How common is disorder? Occurrence of

disordered residues in four domains of life, *Int. J. Mol. Sci. 16*, 19490-19507.

8. Bogatyreva, N. S., Finkelstein, A. V., and Galzitskaya, O. V. (2006) Trend of amino acid composition of proteins of different taxa, *J. Bioinform. Comput. Biol. 4*, 597-608.

9. Yuan, A. D., Rao, M. V., Veeranna, and Nixon, R. A. (2012) Neurofilaments at a glance, *J. Cell Sci. 125*, 3257-3263.

10. Perrot, R., Berges, R., Bocquet, A., and Eyer, J. (2008) Review of the multiple aspects of neurofilament functions, and their possible contribution to neurodegeneration, *Mol. Neurobiol. 38*, 27-65.

11. Kumar, S., Yin, X. H., Trapp, B. D., Hoh, J. H., and Paulaitis, M. E. (2002) Relating interactions between neurofilaments to the structure of axonal neurofilament distributions through polymer brush models, *Biophys. J. 82*, 2360-2372.

12. Kumar, S., and Hoh, J. H. (2004) Modulation of repulsive forces between neurofilaments by sidearm phosphorylation, *Biochem. Biophys. Res. Commun. 324*, 489-496.

13. Rao, M. V., Garcia, M. L., Miyazaki, Y., Gotow, T., Yuan, A. D., Mattina, S., Ward, C. M., Calcutt, N. A., Uchiyama, Y., Nixon, R. A., and Cleveland, D. W. (2002) Gene replacement in mice reveals that the heavily phosphorylated tail of neurofilament heavy subunit does not affect axonal caliber or the transit of cargoes in slow axonal transport, *J. Cell Biol. 158*, 681-693.

14. Rao, M. V., Campbell, J., Yuan, A. D., Kumar, A., Gotow, T., Uchiyama, Y., and Nixon, R. A. (2003) The neurofilament middle molecular mass subunit carboxyl-terminal tail domains is essential for the radial growth and cytoskeletal architecture of axons but not for regulating neurofilament transport rate, *J. Cell Biol. 163*, 1021-1031.

15. Garcia, M. L., Lobsiger, C. S., Shah, S. B., Deerinck, T. J., Crum, J., Young, D., Ward, C. M., Crawford, T. O., Gotow, T., Uchiyama, Y., Ellisman, M. H., Calcutt, N. A., and Cleveland, D. W. (2003) NF-M is an essential target for the myelin-directed "outside-in" signaling cascade that mediates radial axonal growth, *J. Cell Biol. 163*, 1011-1020.

16. Garcia, M. L., Rao, M. V., Fujimoto, J., Garcia, V. B., Shah, S. B., Crum, J., Gotow, T., Uchiyama, Y., Ellisman, M., Calcutt, N. A., and Cleveland, D. W. (2009) Phosphorylation of highly conserved neurofilament medium KSP repeats is not required for myelin-dependent radial axonal growth, *J. Neurosci. 29*, 1277-1284.

17. Srinivasan, N., Bhagawati, M., Ananthanarayanan, B., and Kumar, S. (2014) Stimuli-sensitive intrinsically disordered protein brushes, *Nat. Commun. 5*, 5145.

18. Pregent, S., Lichtenstein, A., Avinery, R., Laser-Azogui, A., Patolsky, F., and Beck, R. (2015) Probing the interactions of intrinsically disordered proteins using nanoparticle tags, *Nano Lett. 15*, 3080-3087.

19. Hung, A. M., Noh, H., and Cha, J. N. (2010) Recent advances in DNA-based directed assembly on surfaces, *Nanoscale 2*, 2530-2537.

20. Mendes, P. M., Yeung, C. L., and Preece, J. A. (2007) Bio-nanopatterning of surfaces, *Nanoscale Res. Lett. 2*, 373-384.

21. Blawas, A. S., and Reichert, W. M. (1998) Protein patterning, *Biomaterials 19*, 595-609.

22. Ricoult, S. G., Nezhad, A. S., Knapp-Mohammady, M., Kennedy, T. E., and Juncker, D. (2014) Humidified microcontact printing of proteins: universal patterning of proteins on both low and high energy surfaces, *Langmuir 30*, 12002-12010.

23. Bernard, A., Renault, J. P., Michel, B., Bosshard, H. R., and Delamarche, E. (2000) Microcontact printing of proteins, *Adv. Mater. 12*, 1067-1070.

24. Mooney, J. F., Hunt, A. J., McIntosh, J. R., Liberko, C. A., Walba, D. M., and Rogers, C. T. (1996) Patterning of functional antibodies and other proteins by photolithography of silane monolayers, *Proc. Natl. Acad. Sci. USA 93*, 12287-12291.

25. Sorribas, H., Padeste, C., and Tiefenauer, L. (2002) Photolithographic generation of protein micropatterns for neuron culture applications, *Biomaterials 23*, 893-900.

26. Lee, K. B., Park, S. J., Mirkin, C. A., Smith, J. C., and Mrksich, M. (2002) Protein nanoarrays generated by dip-pen nanolithography, *Science 295*, 1702-1705.

27. Lim, J. H., Ginger, D. S., Lee, K. B., Heo, J., Nam, J. M., and Mirkin, C. A. (2003) Direct-write dip-pen nanolithography of proteins on modified silicon oxide surfaces, *Angew. Chem. Int. Ed. 42*, 2309-2312.

28. Zhang, G. J., Tanii, T., Zako, T., Hosaka, T., Miyake, T., Kanari, Y., Funatsu, T., and Ohdomari, I. (2005) Nanoscale patterning of protein using electron beam lithography of organosilane self-assembled monolayers, *Small 1*, 833-837.

29. Ballav, N., Thomas, H., Winkler, T., Terfort, A., and Zharnikov, M. (2009) Making protein patterns by writing in a protein-repelling matrix, *Angew. Chem. Int. Ed. 48*, 5833-5836.

30. Wilson, W. C. Jr., and Boland, T. (2003) Cell and organ printing 1: protein and cell printers, *Anat. Rec. A Discov. Mol. Cell Evol. Biol. 272*, 491-496.

31. McWilliam, I., Chong Kwan, M., and Hall, D. (2011) Inkjet printing for the production of protein microarrays, *Methods Mol. Biol. 785*, 345-361.

32. Tolstyka, Z. P., Richardson, W., Bat, E., Stevens, C. J., Parra, D. P., Dozier, J. K., Distefano, M. D., Dunn, B., and Maynard, H. D. (2013) Chemoselective immobilization of proteins by microcontact printing and bio-orthogonal click reactions, *Chembiochem 14*, 2464-2471.

33. Reynolds, N. P., Tucker, J. D., Davison, P. A., Timney, J. A., Hunter, C. N., and Leggett, G. J. (2009) Site-specific immobilization and micrometer and nanometer scale photopatterning of yellow fluorescent protein on glass surfaces, *J. Am. Chem. Soc. 131*, 896-897.

34. Xia, S., Cartron, M., Morby, J., Bryant, D. A., Hunter, C. N., and Leggett, G. J. (2016) Fabrication of nanometer- and micrometer-scale protein structures by site-specific immobilization of

histidine-tagged proteins to aminosiloxane films with photoremovable protein-resistant protecting groups, *Langmuir 32*, 1818-1827.

35. Smith, J. J., Conrad, D. W., Cuneo, M. J., and Hellinga, H. W. (2005) Orthogonal site-specific protein modification by engineering reversible thiol protection mechanisms, *Protein Sci. 14*, 64-73.

36. Christman, K. L., Broyer, R. M., Tolstyka, Z. P., and Maynard, H. D. (2007) Site-specific protein immobilization through N-terminal oxime linkages, *J. Mater. Chem. 17*, 2021-2027.

37. El Muslemany, K. M., Twite, A. A., ElSohly, A. M., Obermeyer, A. C., Mathies, R. A., and Francis, M. B. (2014) Photoactivated bioconjugation between *ortho*-azidophenols and anilines: a facile approach to biomolecular photopatterning, *J. Am. Chem. Soc. 136*, 12600-12606.

38. Lichty, J. J., Malecki, J. L., Agnew, H. D., Michelson-Horowitz, D. J., and Tan, S. (2005) Comparison of affinity tags for protein purification, *Protein Express. Purif. 41*, 98-105.

39. Schmidt, T. G. M., and Skerra, A. (2007) The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins, *Nat. Protoc. 2*, 1528-1535.

40. Soranno, A., Koenig, I., Borgia, M. B., Hofmann, H., Zosel, F., Nettels, D., and Schuler, B. (2014) Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments, *Proc. Natl. Acad. Sci. USA 111*, 4874-4879.

41. Santner, A. A., Croy, C. H., Vasanwala, F. H., Uversky, V. N., Van, Y. Y. J., and Dunker, A. K. (2012) Sweeping away protein aggregation with entropic bristles: intrinsically disordered protein fusions enhance soluble expression, *Biochemistry 51*, 7250-7262.

42. Uversky, V. N. (2009) Intrinsically disordered proteins and their environment: effects of strong denaturants, temperature, pH, counter ions, membranes, binding partners, osmolytes, and macromolecular crowding, *Protein J. 28*, 305-325.

43. MacDonald, J. I., Munch, H. K., Moore, T., and Francis, M. B. (2015) One-step site-specific modification of native proteins with 2-pyridinecarboxyaldehydes, *Nat. Chem. Biol. 11*, 326-331.

44. Witus, L. S., Netirojjanakul, C., Palla, K. S., Muehl, E. M., Weng, C. H., Iavarone, A. T., and Francis, M. B. (2012) Site-specific protein transamination using *N*-methylpyridinium-4-carboxaldehyde, *J. Am. Chem. Soc. 135*, 17223-17229.

45. Palla, K. S., Witus, L. S., Mackenzie, K. J., Netirojjanakul, C., and Francis, M. B. (2015) Optimization and expansion of a site-selective *N*-methylpyridinium-4-carboxaldehyde-mediated transamination for bacterially expressed proteins, *J. Am. Chem. Soc. 137*, 1123-1129.

46. Schrock, A. K., and Schuster, G. B. (1984) Photochemistry of phenyl azide - chemical-properties of the transient intermediates, *J. Am. Chem. Soc. 106*, 5228-5234.

47. Russo, R. E., Mao, X. L., Liu, H. C., Gonzalez, J., and Mao, S. S. (2002) Laser ablation in analytical chemistry - a review, *Talanta 57*, 425-451.

48. Durrant, S. F., and Ward, N. I. (2005) Recent biological and environmental applications of laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS), *J. Anal. At. Spectrom. 20*, 821-829.

49. Ackerley, S., Thornhill, P., Grierson, A. J., Brownlees, J., Anderton, B. H., Leigh, P. N., Shaw, C. E., and Miller, C. C. J. (2003) Neurofilament heavy chain side arm phosphorylation regulates axonal transport of neurofilaments, *J. Cell Biol. 161*, 489-495.

Chapter 3

Green and Red Fluorescent Sensors for Imaging Sodium Levels in Living Cells

Abstract

In this chapter, we present the synthesis and spectroscopic characterization of green and red fluorescent indicators, TG-Sodium and TM-Sodium, which are selective for Na^+ over excess amounts of K^+ and other biologically relevant cations. Both sensors have K_d values appropriate for sensing intracellular Na^+ levels. TM-Sodium displays a 40-fold turn-on in response to Na^+ in the presence of K^+, the greatest dynamic range of all reported fluorescent Na^+ sensors. Confocal microscopy experiments establish that TM-Sodium is the first small-molecule red-emitting Na^+ sensor that is able to report sodium levels in living cells.

The X-ray crystal structure described in this chapter was solved in collaboration with Teera Chantarojsiri.

3.1. Introduction

Organisms devote a significant portion of their metabolic energy to maintain a large Na^+ gradient between the inside and outside of cells. Cells in the resting state have $[Na^+]_{intracellular}$ in the 10-40 mM range while $[Na^+]_{extracellular}$ is in the 120-450 mM range. This Na^+ gradient provides energy for most transport systems across the cell membrane and is therefore critical in powering uptake and efflux of intracellular ions, such as H^+, K^+ and Ca^{2+}, and molecules such as neurotransmitters and glucose. Na^+ is not only essential in maintaining solute homeostasis, it is also the major charge carrier during action potentials in electrically excitable cells, such as neurons, astrocytes and muscle cells.[1-8]

Molecular imaging with fluorescent sensors selective for Na^+ is an attractive technique to study intracellular sodium. Molecular imaging is a non-destructive technique that measures Na^+ levels with spatial and temporal resolution. By contrast, common techniques used to study Na^+ in biological systems, such as flame photometry, atomic absorption spectroscopy, ^{22}Na counting, electron microprobe analysis, ^{23}Na NMR, and Na^+ selective microelectrodes, do not provide such information.[9,10]

There are several reported fluorescent Na^+ sensors suitable for use in aqueous systems. They are, however, excited in the ultraviolet-visible or green region and/or have apparent dissociation constants (K_d) in the 200 mM range, which is too high to sense Na^+ fluxes in most biological systems.[9,11-19] Moreover, the most widely used Na^+ sensors, SBFI and Sodium Green, consist of one receptor and two fluorophores and the large size of these sensors limit their membrane permeability.[9,19-21] Recently, ANG-2, a green Na^+ sensor developed by TEFLabs, became commercially available.[15] There has been one reported red Na^+ sensor, CoroNa Red, but it has a K_d of 200 mM and it localizes to the mitochondria due to the net positive charge of the sensor.[18,22-24] CoroNa Red is no longer commercially available, likely due to its high K_d, which limits its use in biological systems.[18,19] In this report, we present the synthesis and characterization of a green fluorescent Na^+ sensor, TG-Sodium, which is based on the TokyoGreen platform, and a red fluorescent Na^+ sensor, TM-Sodium, which is based on the TokyoMagenta platform.[25,26] Both Na^+ sensors have K_d values appropriate for sensing intracellular Na^+ levels. Compared to ultraviolet-visible and green counterparts, red fluorescent sensors allow deeper tissue penetration, minimize photodamage to biological samples and minimize interference from background autofluorescence by biomolecules.[27] They also allow dual-color imaging with green fluorophores such as GFP. Furthermore, TM-Sodium is significantly smaller than other published sensors as it only consists of one receptor and one fluorophore. The lower molecular weight TG-Sodium and TM-Sodium may have improved membrane and tissue permeability.

3.2. Results and Discussion

3.2.1. Design and Synthesis of Sodium-Responsive Fluorescent Sensors

Our probe design is based on the receptor that is developed by Minta *et al*. The ring size of 1,7-diaza-15-crown-5 provides selectivity for Na^+ over K^+. The two ortho-methoxy groups provide ligands that cap the top and bottom of the crown ether, resulting in increased affinity to Na^+.[9] Due to the ease of synthesizing molecules with symmetry, the reported syntheses of this receptor provides two identical handles for conjugation to fluorophores.[9,28] In order to reduce the size of the

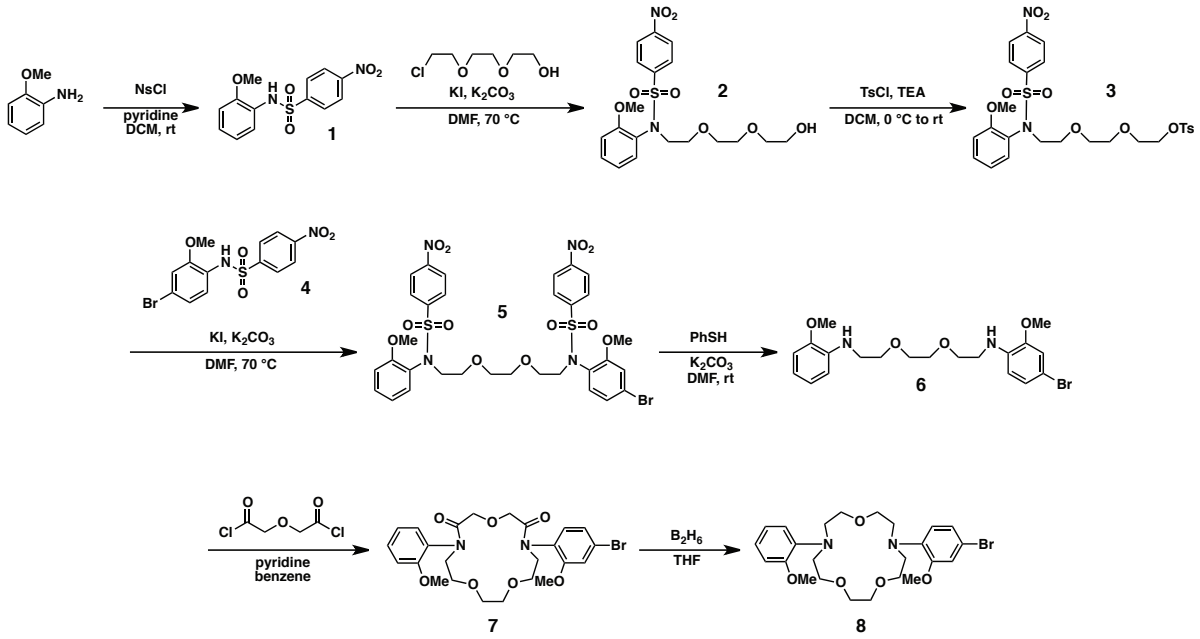


Figure 3.1. Synthetic route to the sodium receptor **8**.

sensors, asymmetric Na^+ receptor **8** with a single bromine handle was synthesized. The key step is the formation of **7** by macrocyclization of the asymmetric linear precursor **6** with diglycolyl chloride (Figure 3.1). Owing to the rigidity of macrocycle **7**, the ^{1}H NMR analysis of **7** is inconclusive, so the X-ray structural analysis was employed to confirm its identity (Figure 3.2 and Section 3.6). With only a single aryl bromide handle, Na^+ receptor **8** can be coupled to TBS-protected xanthone **9** or silicon-xanthone **10** to yield the green TG-Sodium and red TM-Sodium, respectively (Figure 3.3). High temperature ^{1}H NMR was necessary to characterize both sensors (Figure 3.4 and Figure 3.5).

3.2.2. Spectroscopic Properties and Responses to Sodium

We first synthesized and characterized TG-Sodium *in vitro*. To mimic physiological

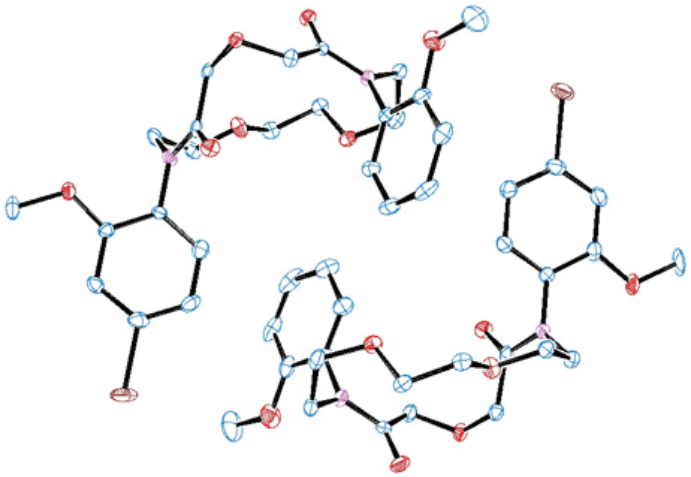


Figure 3.2. Thermal ellipsoid plot of **7**. Red, pink, blue and brown ellipsoids represent oxygen, nitrogen, carbon and bromine atoms, respectively. Hydrogen atoms have been omitted for clarity.

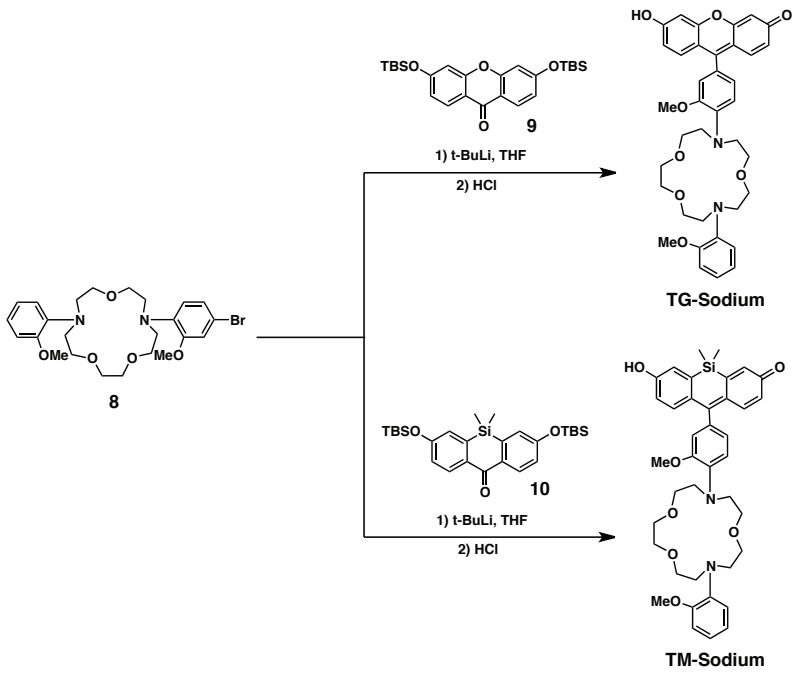


Figure 3.3. Synthetic routes to TG-Sodium and TM-Sodium.

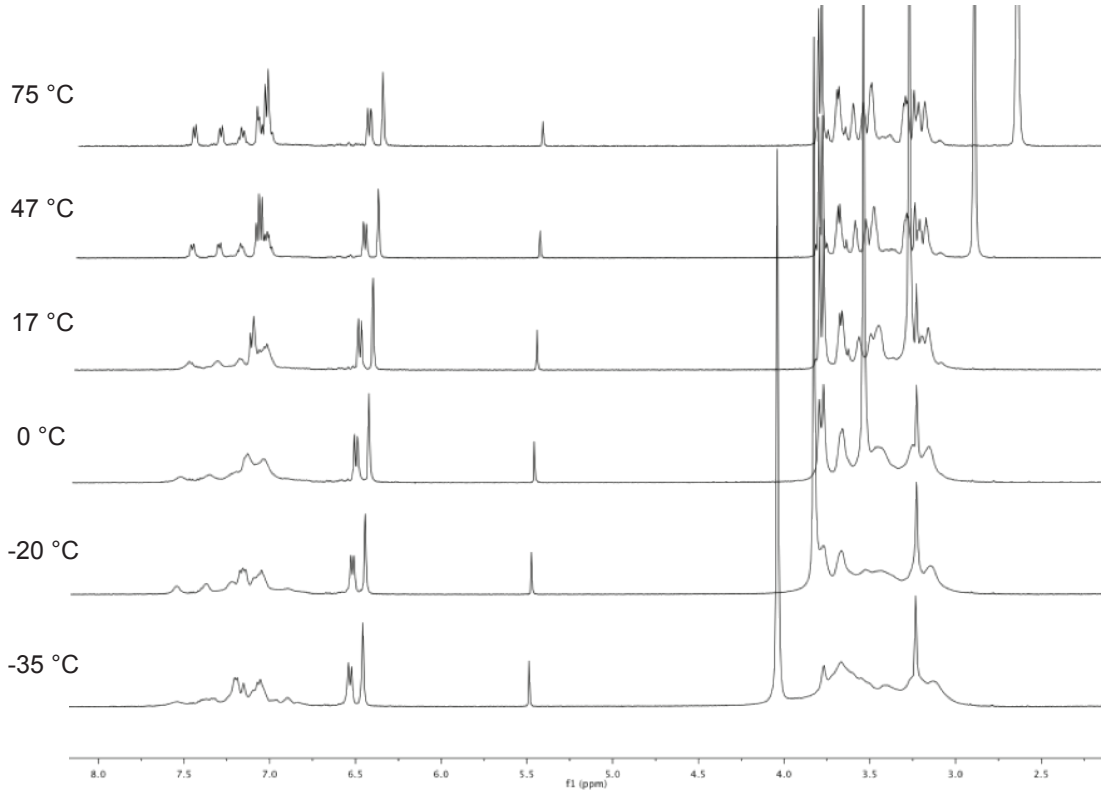


Figure 3.4. Variable temperature ^{1}H NMR of TG-Sodium. Spectra were acquired in acetonitrile-d_3 with minimal methanol-d_4 at 500 MHz.

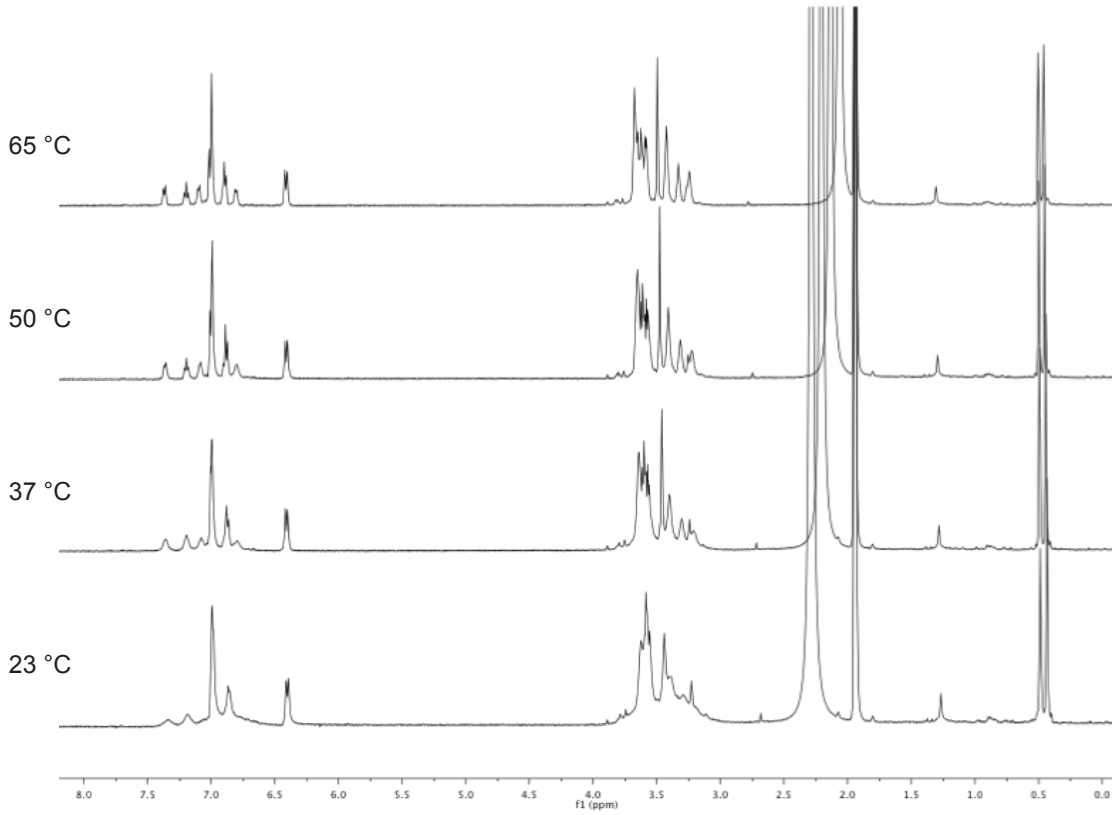


Figure 3.5. Variable temperature ^{1}H NMR of TM-Sodium. Spectra were acquired in acetonitrile-d_3 with minimal methanol-d_4 at 500 MHz.

conditions, Na^+ titration is conducted in aqueous buffers with various Na^+ concentrations while maintaining the ionic strength at 135 mM with K^+.[9] In 10 mM pH 7.5 MOPS buffer with 0 mM Na^+ and 135 mM K^+, TG-Sodium exhibits an absorption peak centered at 495 nm with a maximum fluorescence emission at 512 nm (Figure 3.6a, Figure 3.7a and Table 3.1). The quantum yield (**Φ**) of apo-TG-Sodium is low (**Φ** = 0.0037) due to efficient photoinduced electron transfer (PeT) quenching.[14,29-32] Upon increasing Na^+ concentrations to 135 mM, its absorption and fluorescence emission wavelengths remain at 495 nm and 512 nm, respectively. Because of less efficient PeT quenching, however, **Φ** of TG-Sodium in the presence of 40000 equivalent of Na^+ increases to 0.12, resulting in a 20-fold increase in fluorescence (Figure 3.7a and Table 3.1).

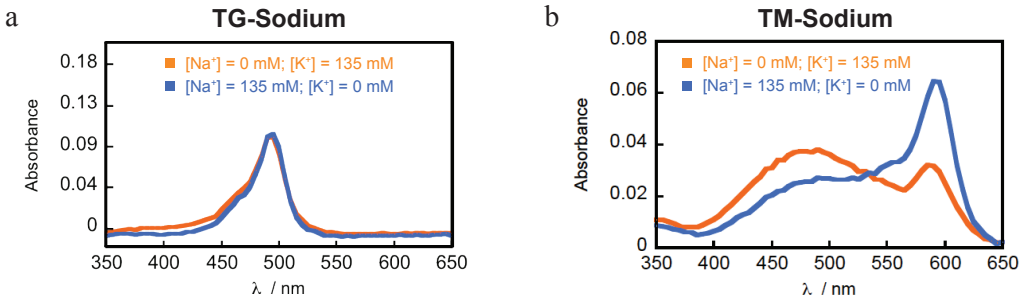


Figure 3.6. Absorption spectra of 5 μM TG-Sodium (a) and TM-Sodium (b). Spectra for [Na^+] of 0 mM and 135 mM in the presence of K^+ are shown. Spectra were acquired in 10 mM pH 7.5 MOPS.

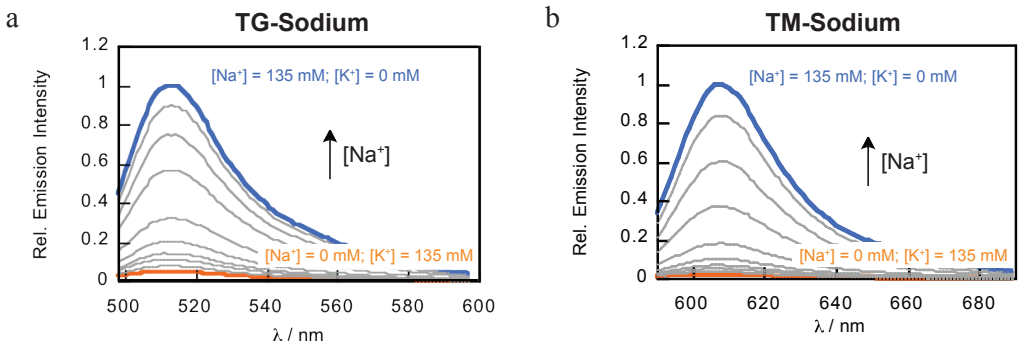


Figure 3.7. Normalized fluorescence responses of 5 μM TG-Sodium (a) and TM-Sodium (b) to Na^+ in the presence of K^+. Spectra for [Na^+] of 0, 1.35, 2.73, 4.13, 7.11, 15.0, 33.8, 67.5, 101.3 and 135 mM, while [Na^+] + [K^+] = 135 mM, are shown. Spectra were acquired in 10 mM pH 7.5 MOPS with excitation at 488 nm (a) or 580 nm (b).

		ε / $cm^{-1}M^{-1}$	Φ
TG-Sodium	0 equiv. Na^+	2.93×10^4	0.0037
	40000 equiv. Na^+	2.70×10^4	0.12
TM-Sodium	0 equiv. Na^+	3.44×10^3	0.0028
	40000 equiv. Na^+	1.85×10^4	0.11

Table 3.1. Extinction coefficients and quantum yields of TG-Sodium and TM-Sodium in 10 mM pH 7.5 MOPS in the absence and presence of Na^+. Sensor concentrations ranging from 2 μM to 8 μM were used. In order to adjust pH to 7.5 with KOH, MOPS buffer at pH 7.5 contains approximately 5 mM K^+.

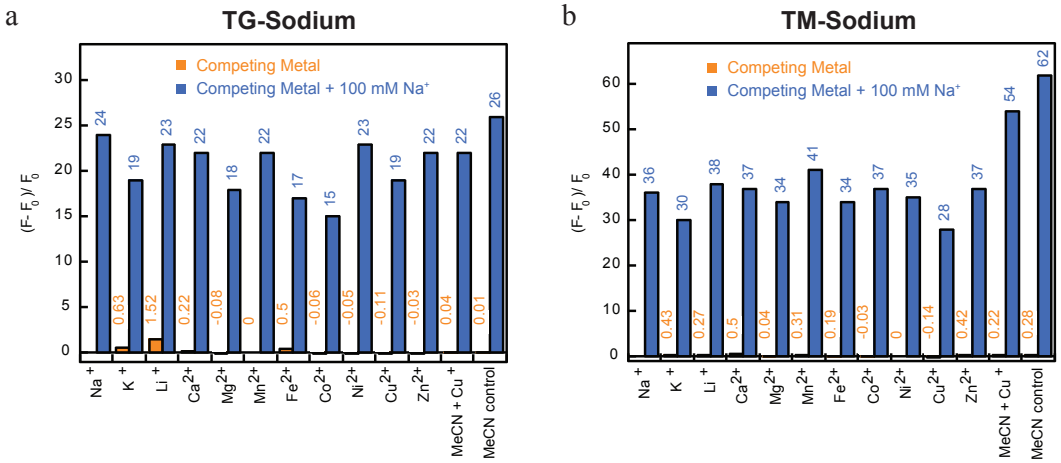


Figure 3.8. Normalized fluorescence responses of TG-Sodium (a) and TM-Sodium (b) to various metal ions. Bars represent the difference between final integrated fluorescence emission (F) and initial integrated emission (F_0) divided by the initial integrated emission (F_0). Orange bars represent the addition of an excess of biologically relevant metal ion (200 mM for K^+; 5 mM for Li^+, Ca^{2+} and Mg^{2+}; 100 μM for all other cations) to a 5 μM solution of TG-Sodium or TM-Sodium. Blue bars represent the subsequent addition of 100 mM Na^+ to the solution. Because Cu^+ was added in the form of $[Cu(MeCN)_4][PF_6]$ dissolved in an acetonitrile stock solution (2 mM), data for the acetonitrile control are also shown. In the case of TM-Sodium, greater turn-on for the addition of Na^+ to the solution containing Cu^+ and acetonitrile is due to the presence of acetonitrile and not Cu^+. Excitation was provided at 488 nm for TG-Sodium and 580 nm for TM-Sodium. The collected emission was integrated over 498 to 650 nm for TG-Sodium and 590 to 700 nm for TM-Sodium. Spectra were acquired in 10 mM pH 7.5 MOPS.

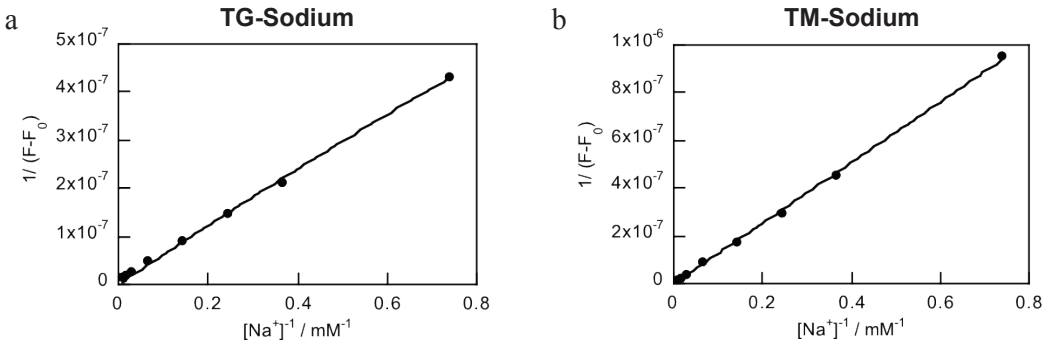


Figure 3.9. Benesi-Hildebrand plots for the complexation of TG-Sodium (a) and TM-Sodium (b) with Na^+. F is the integrated fluorescence emission in the presence of Na^+ while F_0 is the integrated fluorescence emission in the absence of Na^+. Linearity indicates 1:1 complexation between the sensor and Na^+.[17,35] Excitation was provided at 488 nm for TG-Sodium and 580 nm for TM-Sodium. The emission intensities were integrated from 498 nm to 650 nm for TG-Sodium and from 590 nm to 700 nm for TM-Sodium. Spectra were acquired in 10 mM pH 7.5 MOPS. Benesi-Hildebrand plots were generated with the Na^+ titration data set shown in Figure 3.6.

Due to lower LUMO energy, the silicon substituted TG-Sodium analogue, TM-Sodium, is expected to be bathochromically-shifted in both absorption and emission compared to TG-Sodium.[26,33,34] In 10 mM pH 7.5 MOPS buffer with 0 mM Na^+ and 135 mM K^+, TM-Sodium displays red-shifted absorption profile with a major absorption peak at 490 nm and a minor absorption peak at 590 nm (Figure 3.6b and Table 3.1). Upon excitation at 580 nm, TM-Sodium emits weakly at 608 nm. **Φ** of apo-TM-Sodium is 0.0028. In MOPS buffer with 135 mM Na^+, its absorption peak at 590 nm became the major band while its absorption peak at 490 nm became the minor band. Upon addition of Na^+, the maximum emission of TM-Sodium remains at 608 nm. In the presence of 40000 equivalent of Na^+, **Φ** of TM-Sodium drastically increases to 0.11. TM-Sodium has a 40-fold turn-on with 135 mM Na^+ in the presence of K^+, the highest of all reported fluorescent Na^+ sensors (Figure 3.7b and Table 3.1).

The fluorescence response of TG-Sodium and TM-Sodium are selective towards Na^+ with minimal interference from biologically relevant amounts of other metal ions (Figure 3.8). Even physiological concentrations of K^+ (up to 200 mM) only minimally interfere with Na^+ detection. The linearity of the Benesi-Hildebrand plots of both sensors indicates a 1:1 complexation between sensor and Na^+ (Figure 3.9).[17,35] In the presence of K^+, TG-Sodium and TM-Sodium have K_d values

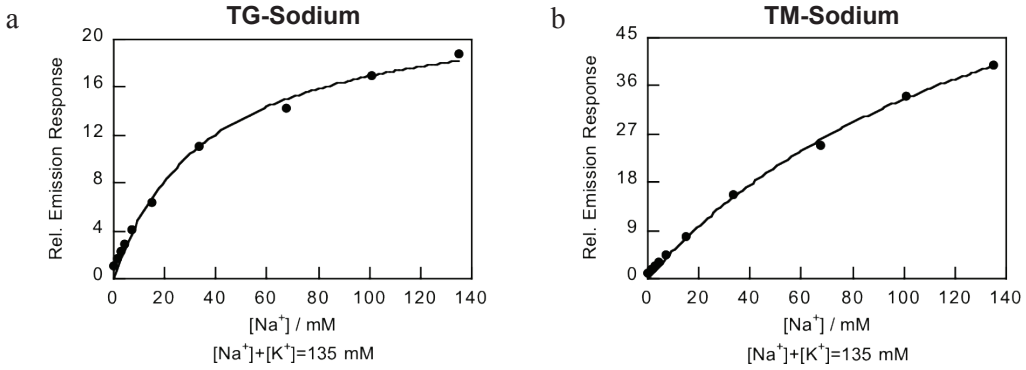


Figure 3.10. Normalized fluorescence responses of 5 μM TG-Sodium (a) and TM-Sodium (b) to Na^+ for K_d value determination. The observed K_d values in the presence of K^+ are 25 mM for TG-Sodium and 47 mM for TM-Sodium. These graphs were generated with the Na^+ titration data set shown in Figure 3.6.

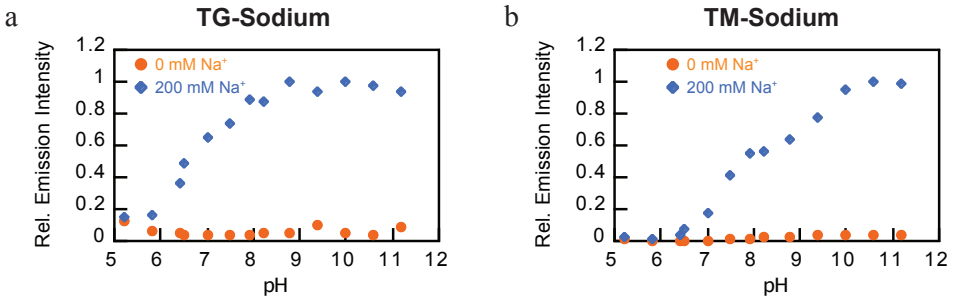


Figure 3.11. Normalized fluorescence responses of TG-Sodium (a) amd TM-Sodium (b) to Na^+ at different pHs. Spectra shown are for 5 μM sensor in the absence of Na^+ and presence of 200 mM Na^+ at pH of 5.2, 5.8, 6.4, 6.5, 7.0, 7.5, 7.9, 8.2, 8.8, 9.4, 10.0, 10.6 and 11.2. Excitation was provided at 488 nm for TG-Sodium and 580 nm for TM-Sodium. The collected emission was integrated over 498 to 650 nm for TG-Sodium and 590 to 700 nm for TM-Sodium. Buffers spanning the pH range were made with 50 mM phthalate, phosphate and MOPS.

for Na^+ of 25 mM and 47 mM, respectively, which are both appropriate for sensing Na^+ fluxes in living cells (Figure 3.10). The pH profile of TG-Sodium and TM-Sodium agree with the reported pK_a values of TokyoGreen and TokyoMagenta, which are 6.2 and 6.8, respectively (Figures 3.11).[25,26] Because both the apo and Na^+-bound sensors are more fluorescent in basic conditions, the fold turn-on of TM-Sodium to Na^+ remains constant (38- to 43-fold) between pH 7 and pH 8. Hence, TM-Sodium is suitable for use in Na^+ detection in cells.

3.2.3. Fluorescence Detection of Sodium in Living Cells

The only reported red sodium sensor, CoroNa Red, has not been widely used because of its predominant localization to the mitochondria and its high K_d.[18] Therefore, we next sought to

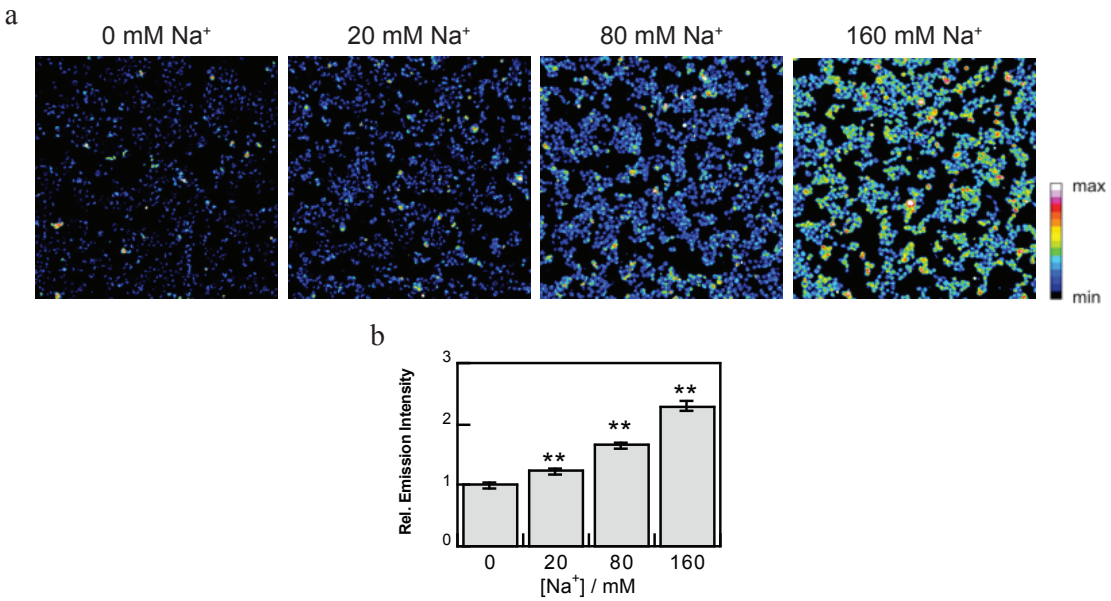


Figure 3.12. Confocal microscopy images (a) and quantifications (b) of Na^+ titration in human embryonic kidney (HEK293T) cells stained with TM-Sodium. (a) Cells were stained with 20 μM TM-Sodium and then treated with 10 μM gramicidin in PBS with various [Na^+] for 20 min before imaging. (b) The bar graph shows the quantification of mean fluorescence intensity of each condition normalized to the control condition ($N = 6$ fields of cells per condition). Bars represent mean ± standard deviation; ** denotes $p < 0.01$.

evaluate whether TM-Sodium is able to report Na^+ levels in living cells. Gramicidin is a pore-forming antibiotic that allows equilibration of Na^+ inside and outside cells. In-cell calibration with gramicidin in HEK293T cells demonstrates that TM-Sodium can reliably report Na^+ levels in a concentration-dependent manner in living cells (Figure 3.12). Notably, TM-Sodium displays uniform staining of HEK293T cells.

3.3. Conclusions

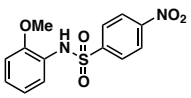
We have presented the synthesis and characterization of a green Na^+ sensor, TG-Sodium, and a red Na^+ sensor, TM-Sodium. TM-Sodium features a robust and selective 40-fold turn-on in response to Na^+, even in the presence K^+. Experiments with HEK293T cells demonstrate that is it capable of detecting changes in Na^+ levels in a concentration dependent manner. TM-Sodium is the first reported small molecule red sodium sensor that has the appropriate K_d to be used for Na^+ detection in cells. We are currently developing second generation of Na^+ sensors with improved pH profile for biological studies.

3.4. Materials and Methods

3.4.1. General Procedures and Materials

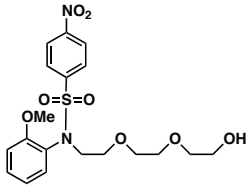
All reactions utilizing air- or moisture-sensitive reagents were performed in dried glassware under an atmosphere of dry N_2. Silica gel P60 (SiliCycle) and Brockmann basic aluminum oxide (Sigma-Aldrich) were used for column chromatography. SiliCycle 60 F_{254} silica gel (SiliCycle) and basic aluminum oxide 60 F_{254} (EMD) were used for analytical thin layer chromatography and visualized by fluorescence quenching under UV light or by staining with iodine. 4-Bromo-2-methoxyaniline was purchased from Oakwood Products (West Columbia, SC) and all other chemicals were purchased from Sigma-Aldrich (St. Louis, MO) and used as received. Xanthone **9**[36] and xanthone **10**[26] were synthesized according to literature procedures. ^1H and ^{13}C NMR spectra for characterization of new compounds were collected in $CDCl_3$, CD_3CN or CD_3OD (Cambridge Isotope Laboratories) at 25 °C (or indicated otherwise) at the reported frequency at the College of Chemistry NMR Facility at the University of California, Berkeley. All chemical shifts are reported in parts per million and referenced to the residual solvent peak from $CDCl_3$ at 7.26 ppm, or CD_3CN at 1.94 ppm. Splitting patterns are indicated as follows: br, broad, d, doublet, dd, doublet of doublets, dt, doublet of triplets, m, multiplet, q, quartet, s, singlet, t, triplet. High-resolution electrospray ionization mass spectral (HRMS-ESI) analyses were carried out at the QB3 Mass Spectrometry Facility at the University of California, Berkeley.

3.4.2. Small Molecule Synthesis

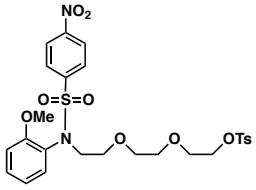


***N*-(2-Methoxyphenyl)-4-nitrobenzenesulfonamide (1).** *O*-anisidine (5 g, 40.6 mmol, 1.0 equiv.) and pyridine (64 mL) were dissolved in 200 mL of dichloromethane. Nosyl chloride (13.5 g, 60.9 mmol, 1.5 equiv.) dissolved in 150 mL of dichloromethane was then added into the reaction

mixture dropwise over 20 min at room temperature. The reaction mixture was stirred for 15 h at room temperature. It was then washed with water (3x), brine and dried over Na_2SO_4, and the solvent was removed by rotary evaporation. The crude residue was subsequently recrystallized in absolute ethanol and yellow crystals were obtained (10.8 g, 35 mmol, 86%). ^{1}H NMR (300 MHz, chloroform-*d*) δ 8.24 (d, *J* = 9.0, 2H), 7.91 (d, *J* = 9.0, 2H), 7.56 (dd, *J* = 7.9, 1.6 Hz, 1H), 7.11 (td, *J* = 7.9, 1.7 Hz, 1H), 7.03 (s, 1H), 6.94 (td, *J* = 7.8, 1.3 Hz, 1H), 6.74 (dd, *J* = 8.2, 1.3 Hz, 1H), 3.62 (s, 3H). ^{13}C NMR (101 MHz, chloroform-*d*) δ 150.26, 150.08, 145.00, 128.69, 126.78, 124.79, 124.07, 122.42, 121.45, 110.93, 55.76. HRMS-ESI calculated for $C_{13}H_{11}O_5N_2S$ (M-H$^+$), 307.0394, found, 307.0392.

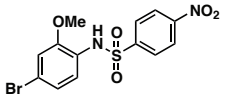


***N*-(2-(2-(2-Hydroxyethoxy)ethoxy)ethyl)-*N*-(2-methoxyphenyl)-4-nitrobenzenesulfonamide (2). 1** (26 g, 84 mmol, 1.0 equiv.), K_2CO_3 (46.4 g, 336 mmol, 4.0 equiv.) and KI (30 g, 181 mmol, 2.2 equiv.) were dissolved in 300 mL of DMF. The mixture was heated to 70 °C and 2-[2-(2-chloroethoxy) ethoxy]ethanol (17 g, 101 mmol, 1.2 equiv.) was then added to the heated reaction mixture in one portion. The reaction mixture was stirred at 70 °C for 15 h. Most of the DMF was removed by rotary evaporation. The residue was dissolved in ethyl acetate and washed with water. The aqueous layer was extracted with diethyl ether (2x). The combined organic layer was washed with water (2x), brine and dried over Na_2SO_4, and the solvent was removed by rotary evaporation to give brown oil. Purification by a plug of silica gel (100% ethyl acetate) gave **2** as a light brown oil (37 g, 0.84 mmol, quantitative yield). ^{1}H NMR (400 MHz, chloroform-*d*) δ 8.27 (d, *J* = 8.8, 2H), 7.85 (d, *J* = 8.8, 2H), 7.40 – 7.28 (m, 2H), 6.97 (td, *J* = 7.6, 1.3 Hz, 1H), 6.77 (dd, *J* = 8.3, 1.3 Hz, 1H), 3.71 (dq, *J* = 4.9, 1.6 Hz, 2H), 3.57 (ddd, *J* = 8.6, 4.9, 2.8 Hz, 10H), 3.34 (s, 3H). HRMS-ESI calculated for $C_{19}H_{24}O_8N_2NaS$ (M+Na$^+$), 463.1146, found, 463.1147.

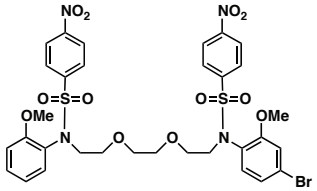


2-(2-(2-(*N*-(2-Methoxyphenyl)-4-nitrophenylsulfonamido)ethoxy)ethoxy)ethyl 4-methylbenzenesulfonate (3). *Product* **2** (5.5 g, 12.5 mmol, 1.0 equiv.) was dissolved in 100 mL of dichloromethane and cooled to 0 °C. Tosyl chloride (4.8 g, 25 mmol, 2.0 equiv.) was then added to the solution as a solid. Lastly, triethylamine (2.5 g, 3.5 mL , 25 mmol, 2.0 equiv.) was added to the reaction mixture, which was allowed to stir at 0 °C and warm up to room temperature overnight. The reaction mixture was cooled to 0 °C again and quenched with 80 mL of water for 10 min. It was then extracted with dichloromethane (2x). The combined organic layer was washed with water (2x), brine and dried over Na_2SO_4, and the solvent was removed by rotary evaporation. Purification by flash column chromatography (silica gel, 20%-40% ethyl acetate/hexanes) gave **3** as a viscous light yellow oil (6.7 g, 11.3 mmol, 90% yield). ^{1}H NMR (300 MHz, chloroform-*d*) δ 8.26 (d, *J* = 8.8 Hz, 2H), 7.84 (d, *J* = 8.8 Hz, 2H), 7.78 (d, *J* = 8.3 Hz, 2H), 7.37 – 7.24 (m, 4H), 6.95 (td, *J* = 7.6, 1.3 Hz, 1H), 6.77 (dd, *J* = 8.8, 1.3 Hz, 1H), 4.17 – 4.08 (m, 2H), 3.67 – 3.59 (m, 2H), 3.55 – 3.41 (m,

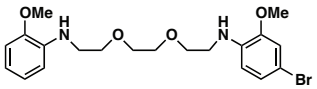
8H), 3.34 (s, 3H), 2.43 (s, 3H). ^{13}C NMR (75 MHz, chloroform-*d*) δ 155.97, 149.86, 146.31, 145.01, 133.63, 133.10, 130.62, 130.02, 128.92, 128.14, 125.82, 123.70, 121.10, 111.82, 70.83, 70.43, 69.41, 69.25, 68.91, 54.96, 49.61, 21.83. HRMS-ESI calculated for $C_{26}H_{30}O_{10}N_2NaS_2$ ($M+Na^+$), 617.1234, found, 617.1238.



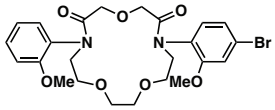
***N*-(4-Bromo-2-methoxyphenyl)-4-nitrobenzenesulfonamide (4).** Pyridine (39 mL) and 4-bromo-2-methoxyaniline (5 g, 25 mmol, 1.0 equv.) were dissolved in 120 mL of dichloromethane. Nosyl chloride (6.6 g, 30 mmol, 1.2 equiv.) dissolved in 100 mL of dichloromethane was then added into the reaction mixture dropwise over 30 min at room temperature. The reaction mixture was stirred for 7 h at room temperature. It was then washed with water (2x), brine and dried over Na_2SO_4, and the solvent was removed by rotary evaporation. Purification by flash column chromatography (silica gel, 75% dichloromethane/hexanes) gave **3** as a yellow solid (9.3 g, 24 mmol, 96% yield). ^{1}H NMR (300 MHz, chloroform-*d*) δ 8.32 – 8.21 (m, 2H), 7.99 – 7.86 (m, 2H), 7.44 (d, *J* = 8.5 Hz, 1H), 7.08 (dd, *J* = 8.6, 2.1 Hz, 1H), 6.96 (br s, 1H), 6.88 (d, *J* = 2.1 Hz, 1H), 3.64 (s, 3H). ^{13}C NMR (101 MHz, chloroform-*d*) δ 150.56, 150.42, 144.81, 128.68, 124.53, 124.27, 124.08, 123.37, 119.49, 114.66, 56.19. HRMS-ESI calculated for $C_{13}H_{10}O_5N_2BrS$ ($M-H^+$), 384.9499, found, 384.9497.



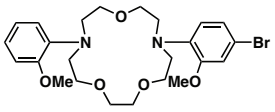
***N*-(4-Bromo-2-methoxyphenyl)-*N*-(2-(2-(2-(*N*-(2-methoxyphenyl)-4-nitrophenylsulfonamido)ethoxy)ethoxy)ethyl)-4-nitrobenzenesulfonamide (5).** K_2CO_3 (29 g, 210 mmol, 4.0 equiv.), KI (8.7 g, 52 mmol, 1.0 equiv.), **3** (34 g, 58 mmol, 1.1 equiv.) and **4** (20 g, 52 mmol, 1.0 equiv) were dissolved in 200 mL of DMF. The reaction mixture was stirred at 70 °C for 15 h. Most of the DMF was removed by rotary evaporation. The resulting residue was dissolved in dichloromethane and washed with water (3′), brine and dried over Na_2SO_4. The solvent was removed by rotary evaporation to give off-white solid, which was added into boiling absolute ethanol. After cooling mixture to 0 °C, insoluble pure product was collected as a white fluffy powder (36 g, 45 mmol, 85% yield). ^{1}H NMR (400 MHz, chloroform-*d*) δ 8.34 – 8.21 (m, 4H), 7.91 – 7.79 (m, 4H), 7.36 – 7.28 (m, 2H), 7.17 (d, *J* = 8.3 Hz, 1H), 7.07 (dd, *J* = 8.3, 2.1 Hz, 1H), 7.00 – 6.88 (m, 2H), 6.78 (d, 1H), 4.13 – 3.46 (m, 8H), 3.44 (s, 4H), 3.36 (s, 3H), 3.33 (s, 3H). ^{13}C NMR (101 MHz, chloroform-*d*) δ 156.62, 155.98, 149.99, 149.89, 146.31, 146.14, 134.75, 133.70, 130.67, 128.94, 128.92, 125.81, 125.23, 124.35, 124.02, 123.84, 123.72, 121.13, 115.66, 111.85, 70.39, 70.36, 69.17, 55.46, 54.98, 49.60, 49.56. HRMS-ESI calculated for $C_{32}H_{33}O_{12}N_4BrNaS_2$ ($M+Na^+$), 831.0612, found, 831.0626.



4-Bromo-2-methoxy-*N*-(2-(2-(2-((2-methoxyphenyl)amino)ethoxy)ethoxy)ethyl)aniline (6). 5 (5.5 g, 6.8 mmol, 1.0 equiv.), K_2CO_3 (9.4 g, 68 mmol, 10 equiv.) and 70 mL of DMF were added into a 20 mL scintillation vial. Thiophenol (5.3 g, 4.9 mL, 48 mmol, 7.0 equiv.) was then added to the reaction mixture. After stirring for 4 h at room temperature, the reaction mixture was poured into water and extracted with ethyl acetate. The aqueous layer was extracted with ethyl acetate. The combined organic layer was washed with water (2x), brine and dried over Na_2SO_4, and the solvent was removed by rotary evaporation. Purification by flash column chromatography (silica gel, 20% ethyl acetate/hexanes) gave **6** as a viscous light brown oil (2.6 g, 5.9 mmol, 87% yield). ^{1}H NMR (300 MHz, chloroform-*d*) δ 6.95 (dd, *J* = 8.4, 2.1 Hz, 1H), 6.87 (td, *J* = 7.5, 1.6 Hz, 1H), 6.83 (d, *J* = 2.1 Hz, 1H), 6.76 (dd, *J* = 7.9, 1.5 Hz, 1H), 6.70 (dd, *J* = 7.4, 1.6 Hz, 1H), 6.67 – 6.59 (m, 1H), 6.45 (d, *J* = 8.4 Hz, 1H), 4.58 (s, 2H), 3.81 (s, 3H), 3.79 (s, 3H), 3.74 (q, *J* = 5.2 Hz, 4H), 3.67 (s, 4H), 3.33 (t, *J* = 5.4 Hz, 2H), 3.28 (t, *J* = 5.4 Hz, 2H).^{13}C NMR (75 MHz, chloroform-*d*) δ 147.70, 147.12, 138.10, 137.37, 123.73, 121.24, 116.78, 112.79, 110.78, 110.03, 109.49, 107.97, 70.39, 70.35, 69.83, 69.60, 55.66, 55.41, 43.32, 43.23.

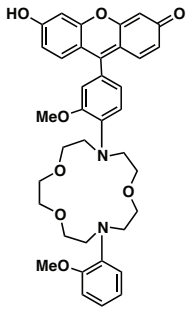


7-(4-Bromo-2-methoxyphenyl)-13-(2-methoxyphenyl)-1,4,10-trioxa-7,13-diazacyclopentadecane-8,12-dione (7). Two pressure equalizing addition funnels with threaded valves were charged with **6** (1.0 g, 2.3 mmol, 1.0 equiv.) and diglycolyl chloride (0.39 g, 2.3 mmol, 1.0 equiv) dissolved in anhydrous benzene separately (80 mL each). They were simultaneously added dropwise into a stirred solution of anhydrous pyridine (1.9 mL) and anhydrous benzene (200 mL) over 24 h. After addition has completed, the reaction mixture was heated to reflux for 5 h. After the solvent was removed under vacuum, the residue was dissolved in dichloromethane and washed (2x) with 1 M HCl. It was then washed with water and brine, and dried over Na_2SO_4. The solvent was removed by rotary evaporation to give an off white foamy solid. Purification by column chromatography (basic alumina, 0-5% methanol in acetonitrile) gave **7** as a foamy white solid (0.39 g, 0.73 mmol, 32% yield). Recrystallization of purified **7** in absolute ethanol and diethyl ether as co-solvent at -20 °C gave shiny colorless crystals appropriate for X-ray crystallography. ^{1}H NMR (400 MHz, chloroform-*d*) δ 7.67 – 6.62 (m, 7H), 5.06 – 2.65 (m, 22H). HRMS-ESI calculated for $C_{24}H_{30}O_7N_2Br$ ($M+H^+$), 537.1231, found, 537.1231.

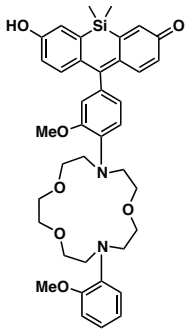


7-(4-Bromo-2-methoxyphenyl)-13-(2-methoxyphenyl)-1,4,10-trioxa-7,13-diazacyclopentadecane (8). Product **7** (0.81 g, 1.5 mmol, 1.0 equiv.) was dissolved in 30 mL of anhydrous THF and sodium borohydride (0.34 g, 9.1 mmol, 6.0 equiv.) was added to the solution as a solid. Boron trifluoride diethyl etherate (47% solution, 1.7 g, 1.5 mL, 12 mmol, 8.0 equiv)

dissolved in 20 mL of anhydrous THF was then added dropwise into the reaction mixture over 15 min at room temperature. The resulting mixture was then stirred at reflux for 2 h. After the reaction mixture was cooled to room temperature, it was poured into water and the pH was brought up to 7 using 3 M KOH. The mixture was then extracted with dichloromethane (2x). The combined organic layer was then washed with water and brine, and dried over Na_2SO_4. The solvent was removed by rotary evaporation to give a viscous white oil (0.77 g, 1.5 mmol, quantitative yield), which was used in the next step without further purification. ^{1}H NMR (500 MHz, chloroform-*d*) δ 7.03 (d, *J* = 7.8 Hz, 1H), 7.00 – 6.81 (m, 5H), 3.82 (s, 3H), 3.81 (s, 3H), 3.72 – 3.59 (m, 10H), 3.50 – 3.37 (m, 10H). ^{13}C NMR (101 MHz, chloroform-*d*) δ 153.68, 153.12, 140.61, 139.80, 123.71, 122.49, 122.21, 121.18, 120.92, 115.28, 114.36, 111.98, 71.41, 71.18, 70.98, 70.81, 69.97, 69.80, 55.86, 55.58, 53.38, 53.26, 52.61. HRMS-ESI calculated for $C_{24}H_{33}O_5N_2BrNa$ (M+Na$^+$), 531.1465, found, 531.1469.



6-Hydroxy-9-(3-methoxy-4-(13-(2-methoxyphenyl)-1,4,10-trioxa-7,13-diazacyclopentadecan-7-yl)phenyl)-3*H*-xanthen-3-one (TG-Sodium). In the glovebox, **8** (0.102 g, 0.20 mmol, 3.0 equiv.) was dissolved in 1 mL of THF and xanthone **9** (30 mg, 0.066 mmol, 1.0 equiv.) was dissolved in 5 mL of THF. Both dissolved compounds were then brought out of the glove box. Product **8** in THF was cooled to -78 °C and 1.7 M *tert*-butyllithium in pentane (129 mL, 0.22 mmol, 3.3 equiv.) was then added to the cooled solution dropwise. The resulting mixture was stirred for 8 min at -78 °C. Xanthone **9** dissolved THF was then added to the reaction mixture dropwise at -78 °C. After addition has completed, the reaction mixture was warmed to room temperature and stirred for 30 min. 2 M HCl (2 mL) was then added to the reaction mixture, which was allowed to stir for 1 h at room temperature. The mixture was poured into a saturated solution of $NaHCO_3$ and extracted with ethyl acetate (2x). The combined organic layer was then washed with water (2x) and dried over Na_2SO_4. The solvent was removed by rotary evaporation to give a red film. Purification by column chromatography (basic alumina, 10-30% methanol in DCM) gave **TG-Sodium** as a red film (13 mg, 0.020 mmol, 31%). ^{1}H NMR (500 MHz, acetonitrile-d_3 and 2 drops of methanol-d_4, 75°C) δ 7.47 (d, *J* = 8.1 Hz, 1H), 7.32 (d, *J* = 8.0 Hz, 1H), 7.20 (t, *J* = 8.0 Hz, 1H), 7.14 – 6.96 (m, 6H), 6.45 (dd, *J* = 9.7, 2.1 Hz, 2H), 6.37 (d, *J* = 2.2 Hz, 2H), 3.83 (s, 3H), 3.81 (s, 3H), 3.72 (q, *J* = 5.9 Hz, 4H), 3.63 (t, *J* = 5.0 Hz, 2H), 3.57 (t, *J* = 5.2 Hz, 2H), 3.52 (q, *J* = 6.1, 5.7 Hz, 4H), 3.32 (dt, *J* = 9.4, 4.6 Hz, 4H), 3.27 (br s, 1H), 3.25 (t, *J* = 4.9 Hz, 2H), 3.21 (t, *J* = 5.0 Hz, 2H). HRMS-ESI calculated for $C_{37}H_{41}O_8N_2$ (M+H$^+$), 641.2857, found, 641.2869.



7-Hydroxy-10-(3-methoxy-4-(13-(2-methoxyphenyl)-1,4,10-trioxa-7,13-diazacyclopentadecan-7-yl)phenyl)-5,5-dimethyldibenzo[*b*,*e*]silin-3(5*H*)-one (TM-Sodium). In the glovebox, **8** (0.153 g, 0.30 mmol, 3.0 equiv.) was dissolved in 1 mL of THF and silicon xanthone **10** (50 mg, 0.10 mmol, 1.0 equiv.) was dissolved in 5 mL of THF. Both dissolved compounds were then brought out of the glove box. Product **8** in THF was cooled to -78 °C and 1.7 M *tert*-butyllithium in pentane (194 mL, 0.33 mmol, 3.3 equiv.) was then added to the cooled solution dropwise. The resulting mixture was stirred for 8 min at -78 °C. Silicon xanthone **10** dissolved THF was then added to the reaction mixture dropwise at -78 °C. After addition has completed, the reaction mixture was warmed to room temperature and stirred for 30 min. HCl (2 mL of 2 M solution) was then added to the reaction mixture, which was allowed to stir for 1 h at room temperature. The mixture was poured into a saturated solution of $NaHCO_3$ and extracted with ethyl acetate (2x). The combined organic layers were then washed with water (2x) and dried over Na_2SO_4. The solvent was removed by rotary evaporation. The resulting brown oil was purified by column chromatography (basic alumina, 4-10% methanol in DCM) to give crude **TM-Sodium** as a red/blue oil (33 mg, 48% crude yield), which was further purified by HPLC with Agilent Prep-C18 column to give **TM-Sodium** as a dark red film. ^{1}H NMR (500 MHz, acetonitrile-d_3 and 2 drops of methanol-d_4, 65°C) δ 7.37 (d, J = 8.1 Hz, 1H), 7.20 (t, J = 7.9 Hz, 1H), 7.10 (d, J = 7.9 Hz, 1H), 7.01 (m, 5H), 6.89 (m, 2H), 6.81 (d, J = 7.5 Hz, 1H), 6.41 (d, J = 8.5 Hz, 2H), 3.71 – 3.56 (m, 15H), 3.50 (s, 3H), 3.43 (d, J = 5.4 Hz, 4H), 3.34 (d, J = 5.8 Hz, 2H), 3.26 (d, J = 11.7 Hz, 2H), 0.51 (s, 3H), 0.46 (s, 3H). HRMS-ESI calculated for $C_{39}H_{47}O_7N_2Si$ ($M+H^+$), 683.3147, found, 683.3140.

3.4.3. Spectroscopic Materials and Methods

Millipore water was used to prepare all aqueous solutions. For all studies, sensors were diluted in aqueous buffers from a 2 mM stock in DMSO. All *in vitro* spectroscopic measurements were performed in 10 mM MOPS buffer, pH 7.5, unless indicated otherwise. Fluorescence spectra were obtained using a Quanta Master 4 L-format scanning spectrofluorometer equipped with an LPS-220B 75 W xenon lamp and power supply, A-1010B lamp housing with an integrated igniter, switchable 814 photon-counting/analog photomultiplier detection unit, and MD5020 motor driver (Photon Technology International, Inc.) and UV spectra were acquired using a Cary Bio50 UV spectrophotometer (Varian). Samples for emission and absorption measurements were contained in 1 cm × 0.1 cm quartz cuvette (Starna). Absorption and emission spectra for TG-Sodium, TM-Sodium, fluorescein and cresyl violet were obtained over a range of concentrations where a linear correlation between concentration and absorption was observed. Extinction coefficients were obtained by plotting absorbance versus sensor concentration. Quantum yields were determined by reference to fluorescein ($\Phi_{standard}$ = 0.79 in 0.1 M NaOH) for TG-Sodium and cresyl violet ($\Phi_{standard}$

= 0.54 in methanol) for TM-Sodium. Quantum yields were calculated according to the equation $\Phi_{sample} = \Phi_{standard}$ ($Grad_{sample}$/ $Grad_{standard}$)(ηsample/ηstandard), where Φ is quantum yield, Grad is the slope of the plot of absorbance versus integrated emission intensity, and η is the refractive index of the solvent. The apparent dissociation constants (K_d) for Na^+ in the presence of K^+ were estimated using the following equation: $(F - F_{min}) / (F_{max} - F_{min}) = [Na^+]/(K_d + [Na^+])$, where F is the observed fluorescence, F_{max} is the fluorescence of the sensor in 135 mM Na^+, and F_{min} is the fluorescence of the sensor in 135 mM K^+.

3.4.4. Preparation and Staining of Cell Culture

HEK 293T cells were maintained in exponential growth as a monolayer in Dulbecco's Modified Eagle Mediaum (DMEM, Invitrogen) supplemented with 10% fetal bovine serum (FBS, Hyclone), and incubated at 37 °C in 5% CO_2. One day before imaging, the cells were passaged and plated in phenol red-free medium on a 4-well Lab Tek borosilicate chambered coverglass slides (Nunc) and allowed to grow to 70% confluence. Before the imaging experiments, cells were incubated with 20 μM TM-Sodium (from 2 mM stock in DMSO) in DMEM at 37 °C for 60 min. TM-Sodium in DMEM was then removed from wells and cells were treated with 10 μM gramicidin in DPBS containing 0, 20, 80 or 160 mM Na^+ for 30 min before being imaged.

3.4.5. Confocal Imaging Experiments

Confocal fluorescence imaging studies were performed with a Zeiss laser scanning microscope 710 with a 10× air objective lens, with Zen 2009 software (Carl Zeiss). TM-Sodium was excited at 594 nm with a He-Ne laser, and emission collected using a META detector between 597 and 734 nm. The cells were imaged at 37 °C. The focal plane with highest fluorescence intensity was selected for each field of cells. Image analysis was performed in ImageJ (National Institute of Health) and Zen 2009 software (Carl Zeiss).

3.5. References

1. Bers, D. M., Barry, W. H., and Despa, S. (2003) Intracellular Na^+ regulation in cardiac myocytes, *Cardiovasc. Res. 57*, 897-912.

2. Bers, D. M., and Despa, S. (2006) Cardiac myocytes $Ca2^+$ and Na^+ regulation in normal and failing hearts, *J. Pharmacol. Sci. 100*, 315-322.

3. Kirischuk, S., Parpura, V., and Verkhratsky, A. (2012) Sodium dynamics: another key to astroglial excitability? *Trends Neurosci. 35*, 497-506.

4. Maragakis, N. J., and Rothstein, J. D. (2004) Glutamate transporters: animal models to neurologic disease, *Neurobiol. Dis. 15*, 461-473.

5. Murphy, E., and Eisner, D. A. (2009) Regulation of intracellular and mitochondrial sodium in health and disease, *Circ. Res. 104*, 292-303.

6. Parpura, V., and Verkhratsky, A. (2012) Homeostatic function of astrocytes: Ca^{2+} and Na^+ signalling, *Transl. Neurosci. 3*, 334-344.

7. Rose, C. R. (1997) Intracellular Na^+ regulation in neurons and glia: functional implications, *Neuroscientist 3*, 85-88.

8. Rose, C. R., and Karus, C. (2013) Two sides of the same coin: Sodium homeostasis and signaling in astrocytes under physiological and pathophysiological conditions, *Glia 61*, 1191-1205.

9. Minta, A., and Tsien, R. Y. (1989) Fluorescent indicators for cytosolic sodium, *J. Biol. Chem. 264*, 19449-19457.

10. Harootunian, A. T., Kao, J. P. Y., Eckert, B. K., and Tsien, R. Y. (1989) Fluorescence ratio imaging of cytosolic free Na+ in individual fibroblasts and lymphocytes, *J. Biol. Chem. 264*, 19458-19467.

11. de Silva, A. P., Gunaratne, H. Q. N., Gunnlaugsson, T., and Nieuwenhuizen, M. (1996) Fluorescent switches with high selectivity towards sodium ions: correlation of ion-induced conformation switching with fluorescence function, *Chem. Commun. 16*, 1967-1968.

12. Gunnlaugsson, T., Nieuwenhuyzen, M., Richard, L., and Thoss, V. (2002) Novel sodium-selective fluorescent PET and optically based chemosensors: towards Na^+ determination in serum, *J. Chem. Soc., Perkin Trans 2*, 141-150.

13. He, H. R., Mortellaro, M. A., Leiner, M. J. P., Young, S. T., Fraatz, R. J., and Tusa, J. K. (2003) A fluorescent chemosensor for sodium based on photoinduced electron transfer, *Anal. Chem. 75*, 549-555.

14. Kenmoku, S., Urano, Y., Kanda, K., Kojima, H., Kikuchi, K., and Nagano, T. (2004) Rational design of novel photoinduced electron transfer type fluorescent probes for sodium cation, *Tetrahedron 60*, 11067-11073.

15. Lamy, C. M., and Chatton, J. Y. (2011) Optical probing of sodium dynamics in neurons and astrocytes, *Neuroimage 58*, 572-578.

16. Martin, V. V., Rothe, A., and Gee, K. R. (2005) Fluorescent metal ion indicators based on benzoannelated crown systems: a green fluorescent indicator for intracellular sodium ions, *Bioorg. Med. Chem. Lett. 15*, 1851-1855.

17. Kim, M. K., Lim, C. S., Hong, J. T., Han, J. H., Jang, H. Y., Kim, H. M., and Cho, B. R. (2010) Sodium-ion-selective two-photon fluorescent probe for *in vivo* imaging, *Angew. Chem. Int. Ed. 49*, 364-367.

18. Haugland, R. P. (2005) The handbook: a guide to fluorescent probes and labeling technologies, 10^{th} ed., Molecular Probes, Eugene, OR.

19. Johnson, I., and Spence, M. T. Z. (2010) The molecular probes handbook: a guide to fluorescent probes and labeling technologies, 11^{th} ed., Molecular Probes, Eugene, OR.

20. Meier, S. D., Kovalchuk, Y., and Rose, C. R. (2006) Properties of the new fluorescent Na^+ indicator CoroNa Green: comparison with SBFI and confocal Na^+ imaging, *J. Neurosci. Methods 155*, 251-259.

21. Amorino, G. P., and Fox, M. H. (1995) Intracellular Na^+ measurements using sodium green tetraacetate with flow-cytometry, *Cytometry 21*, 248-256.

22. Baron, S., Caplanusi, A., van de Ven, M., Radu, M., Despa, S., Lambrichts, I., Ameloot, M., Steels, P., and Smets, I. (2005) Role of mitochondrial Na^+ concentration, measured by CoroNa Red, in the protection of metabolically inhibited MDCK cells, *J. Am. Soc. Nephrol. 16*, 3490-3497.

23. Jayaraman, S., Joo, N. S., Reitz, B., Wine, J. J., and Verkman, A. S. (2001) Submucosal gland secretions in airways from cystic fibrosis patients have normal Na^+ and pH but elevated viscosity, *Proc. Natl. Acad. Sci. USA 98*, 8119-8123.

24. Jayaraman, S., Song, Y. L., Vetrivel, L., Shankar, L., and Verkman, A. S. (2001) Noninvasive *in vivo* fluorescence measurement of airway-surface liquid depth, salt concentration, and pH, *J. Clin. Invest. 107*, 317-324.

25. Mottram, L. F., Boonyarattanakalin, S., Kovel, R. E., and Peterson, B. R. (2006) The Pennsylvania green fluorophore: a hybrid of Oregon Green and Tokyo Green for the construction of hydrophobic and pH-insensitive molecular probes, *Org. Lett. 8*, 581-584.

26. Egawa, T., Koide, Y., Hanaoka, K., Komatsu, T., Terai, T., and Nagano, T. (2011) Development of a fluorescein analogue, TokyoMagenta, as a novel scaffold for fluorescence probes in red region, *Chem. Commun. 47*, 4162-4164.

27. Yuan, L., Lin, W., Zheng, K., He, L., and Huang, W. (2013) Far-red to near infrared analyte-responsive fluorescent probes based on organic fluorophore platforms for fluorescence imaging, *Chem.Soc. Rev. 42*, 622-661.

28. Gunnlaugsson, T., Gunaratne, H. Q. N., Nieuwenhuyzen, M., and Leonard, J. P. (2002) Synthesis of functionalised macrocyclic compounds as Na^+ and K^+ receptors: a mild and high yielding nitration in water of mono and bis 2-methoxyaniline functionalised crown ethers, *J. Chem. Soc., Perkin Trans 1*, 1954-1962.

29. Callan, J. F., de Silva, A. P., and Magri, D. C. (2005) Luminescent sensors and switches in the early 21st century, *Tetrahedron 61*, 8551-8588.

30. de Silva, A. P., Gunaratne, H. Q. N., Gunnlaugsson, T., Huxley, A. J. M., McCoy, C. P., Rademacher, J. T., and Rice, T. E. (1997) Signaling recognition events with fluorescent sensors and switches, *Chem. Rev. 97*, 1515-1566.

31. Miura, T., Urano, Y., Tanaka, K., Nagano, T., Ohkubo, K., and Fukuzumi, S. (2003) Rational design principle for modulating fluorescence properties of fluorescein-based probes by photoinduced electron transfer, *J. Am. Chem. Soc. 125*, 8666-8671.

32. Urano, Y., Kamiya, M., Kanda, K., Ueno, T., Hirose, K., and Nagano, T. (2005) Evolution of fluorescein as a platform for finely tunable fluorescence probes, *J. Am. Chem. Soc. 127*, 4888-4894.

33. Fu, M., Xiao, Y., Qian, X., Zhao, D., and Xu, Y. (2008) A design concept of long-wavelength

fluorescent analogs of rhodamine dyes: replacement of oxygen with silicon atom, *Chem. Commun. 15*, 1780-1782.

34. Yamaguchi, S., and Tamao, K. (1998) Silole-containing sigma- and pi-conjugated compounds, *J. Chem. Soc., Dalton Trans. 22*, 3693-3702.

35. Benesi, H. A., and Hildebrand, J. H. (1949) A spectrophotometric investigation of the interaction of iodine with aromatic hydrocarbons, *J. Am. Chem. Soc. 71*, 2703-2707.

36. Minta, A., Kao, J. P. Y., and Tsien, R. Y. (1989) Fluorescent indicators for cytosolic calcium based on rhodamine and fluorescein chromophores. *J. Biol. Chem. 264*, 8171-8178.

3.6. Experimental Details of the X-ray Structural Analysis

Empirical formula	$C_{48}\ H_{58}\ Br_2\ N_4\ O_{14}$	
Formula weight	1074.80	
Temperature	100(2) K	
Wavelength	0.71073 Å	
Crystal system	Orthorhombic	
Space group	Pna2(1)	
Unit cell dimensions	a = 15.278 Å	α = 90°.
	b = 14.461 Å	β = 90°.
	c = 22.362 Å	γ = 90°.
	4940.6 $Å^3$	
Z	4	
Density (calculated)	1.445 Mg/m^3	
Absorption coefficient	1.710 mm^{-1}	
F(000)	2224	
Crystal size	0.18 x 0.10 x 0.08 mm^3	
Theta range for data collection	1.68 to 25.70°.	
Index ranges	-18<=h<=18, -16<=k<=17, -22<=l<=27	
Reflections collected	66732	
Independent reflections	8947 [R(int) = 0.0418]	
Completeness to theta = 25.00°	100.0 %	
Absorption correction	Semi-empirical from equivalents	
Max. and min. transmission	0.8753 and 0.7483	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	8947 / 1 / 614	
Goodness-of-fit on F^2	1.046	
Final R indices [I>2sigma(I)]	R1 = 0.0362, wR2 = 0.0838	
R indices (all data)	R1 = 0.0434, wR2 = 0.0870	
Absolute structure parameter	0.452(6)	
Largest diff. peak and hole	1.379 and -0.801 $e.Å^{-3}$	

Chapter 4

Engineering New and Improved Photophysical Properties into Copper Sensors

Abstract

We report the syntheses, spectroscopic properties and live-cell imaging of a new series of Cu^+ sensors, which consist of benzyl linkers between the Cu^+-selective receptor and the BODIPY fluorophore. The new series of copper sensors, CS4, CS5 and CS7, feature robust turn-on in response to Cu^+ and are highly selective towards Cu^+ over other biologically relevant metals. Live-cell imaging experiments establish that CS4, CS5 and CS7, which have apparent dissociation constants spanning 2.5 orders of magnitude *in vitro*, are able to detect different pools of chelatable Cu^+. Having an appropriate K_d of 0.079 pM, CS5 is the first fluorescent Cu^+ indicator that is able to detect both exogenous and endogenous copper levels in HEK293T cells. Moreover, due to its greatly improved photostability compared to previously reported Cu^+ sensors, CS5 is the first Cu^+ sensor that is able to report Cu^+ levels in real time.

4.1. Introduction

Due to its potent redox and catalytic activities, copper is an essential nutrient in all cell types and it is well understood as an enzyme cofactor. High copper levels, however, can cause oxidative stress, which is toxic to cells. Therefore, cells have evolved intricate systems to tightly control the intracellular copper concentration and localization.[1-4] Misregulation of copper has been implicated in serious neurodegenerative disorders.[5-8] In order to understand the roles of biological copper, we need to understand its trafficking, which is not well understood due to the lack of tools.

In addition to the importance of copper as a tightly bound protein cofactor, it also exists in a chelatable pool. Because of the high abundance of high affinity ligands in the cytosol, such as metallotheoneins, metallochaperones and glutathione, the chelatable pool can be thought of as a buffered pool of metal ions. These chelatable metal ions can be readily exchanged between ligand sets.[9,10] Molecular imaging using small molecule fluorescent sensors with selectivity for copper provides an attractive way to study the dynamic copper pool.[11] Because this non-destructive technique is able to detect copper with spatial and temporal resolution, it is an ideal technique to explore copper accumulation, efflux, trafficking and organelle distribution.

Several small molecule fluorescent Cu^+ sensors have been reported, but only a few of them have been used successfully in biological experiments.[12-18] Moreover, none of them are able to detect copper levels in real time. In order to detect copper mobility in real time, the fluorescent sensor needs to be robust and photostability is crucial in obtaining a reliable readout after multiple illumination cycles.[19,20] The C. Chang group has previously developed Coppersensor-3 (CS3), which is the brightest and most sensitive small molecule fluorescent sensor for biological Cu^+ reported in the literature (Figure 4.1a).[12] CS3, however, suffers from rapid photobleaching and thus is only suitable for snapshot imaging. A more photostable sensor would allow kinetic studies of copper trafficking and overcome complications with heterogeneous cell populations. Here, we present the design, synthesis and properties of a series of Cu^+-selective sensors with improved photostability.

4.2. Results and Discussion

4.2.1. Design and Synthesis of Copper-Responsive Fluorescent Sensors

The design of the series of sensors in this work is based on previously published copper sensors, CS1 and CS3, which consist of a BODIPY fluorescence scaffold and an azatetrathia receptor.[12,18] The thioether-rich receptor provides Cu^+-selectivity due to preferred soft-soft

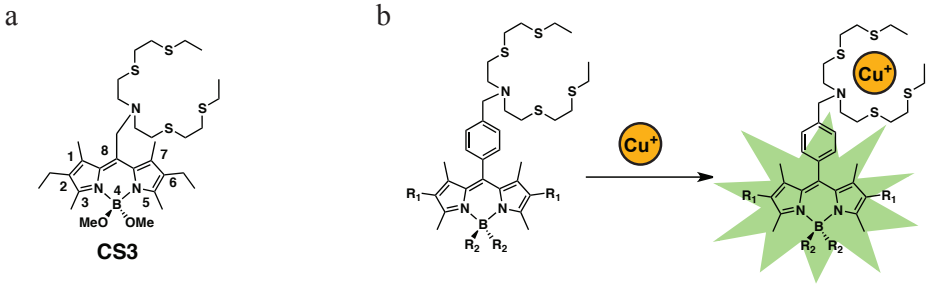


Figure 4.1. a) Structure of CS3 with conventional numbering system for the BODIPY core. b) Scheme for the detection of Cu^+ with benzyl-linked copper sensors.

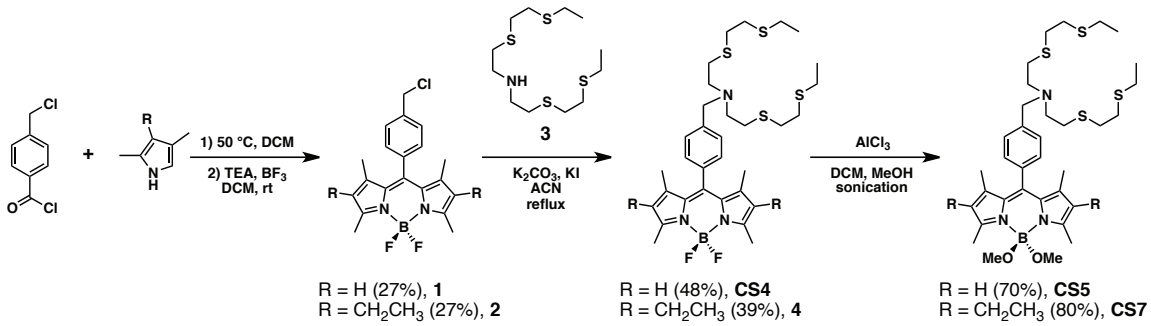


Figure 4.2. Synthetic routes to benzyl-linked copper sensors.

interactions between ligands and metal centers. Fluorescence turn-on of these sensors is suggested to operate through photoinduced electron transfer (PeT) mechanism, which is a widely accepted mechanism for the behavior of many turn-on fluorescence sensors.[21-24]

To improve photostability, we have designed a series of benzyl-linked copper sensors, CS4, CS5 and CS7 (Figure 4.1b and Figure 4.2). Upon excitation of CS3, hydrogen abstraction is likely to occur at the methylene linker, resulting in a radical that is conjugated to the BODIPY fluorophore. The delocalized radical may lead to photodegredation of the sensor. Instead of a methylene linker, a benzyl linker was chosen because the aryl ring is orthogonal to the BODIPY fluorophore.[25] We reasoned that when the fluorophore is excited, the benzyl linker minimizes electronic coupling between the excited fluorophore and the Cu^{+}-selective receptor, which may result in improved photostability. Figure 4.2 outlines the facile synthesis of a series of benzyl-linked copper sensors with various substituents on the BODIPY fluorophore. Electronic properties of sensors can be tuned by installing different substituents on the 2-, 4- and 6- positions of the BODIPY fluorophore (Figure 4.1a). The synthesis of these sensors begin with condensation of 2,4-dimethylpyrrole or 3-ethyl-2,4-dimethylpyrrole with (chloromethyl)benzoyl chloride. This is followed by oxidation and boron insertion, which occur in one-pot to form the BODIPY fluorophores with reactive benzyl chloride handles. Nucleophilic displacement of chloro **1** and **2** with receptor **3** affords CS4 and **4**. Substituting fluoro substituents with methoxy groups on the boron center of BODIPY yields CS5 and CS7.

4.2.2. Spectroscopic Properties and Responses to Copper

Spectroscopic analysis of CS4, CS5 and CS7 in HEPES buffer at pH 7.4, demonstrates that

	R_1	R_2	λ_{abs} / nm		λ_{em} / nm		ε / $cm^{1}M^{1}$		Φ		Fold turn-on	K_d / pM
			Apo	Cu^{+}-bound	Apo	Cu^{+}-bound	Apo	Cu^{+}-bound	Apo	Cu^{+}-bound		
CS4	H	F	525	495	535	508	1.3 x10^4	2.8 x10^4	0.051	0.21	7	12
CS5	H	OMe	525	495	506	506	1.6 x10^4	3.9 x10^4	0.053	0.30	12	0.079
CS7	Ethyl	OMe	545	520	535	532	2.8 x10^4	5.5 x10^4	0.033	0.45	27	1.1
CS3	--	--	550	540	560	548	3.1 x10^4	4.6 x10^4	0.007	0.40	75	0.089

Table 4.1. Properties of benzyl-linked copper sensors (CS4, CS5 and CS7) and CS3. Spectroscopic measurements of CS4, CS5 and CS7 were performed in 20 mM pH 7.4 HEPES. Spectroscopic measurements of CS3 were reported in the literature[12] (see Figure 4.2 for the structures of CS3, CS4, CS5 and CS7). Quantum yields were measured based on fluorescein standard ($\Phi_{standard}$ = 0.79 in 0.1 M NaOH). CS4 and CS5 were excited at 488 nm. CS7 was excited at 520 nm.

these sensors all exhibit robust turn-on in response to Cu^+. Table 4.1 provides a summary of the properties of these copper sensors. Apo-CS4 exhibits maximum absorption peak at 525 nm (ε = 1.3 x 10^4 $cm^{-1}M^{-1}$) with a maximum fluorescence emission at 535 nm (Φ = 0.051). Upon addition of Cu^+, the maximum absorption peak is hypsochromically shifted to 495 nm (ε = 2.8 x 10^4 $cm^{-1}M^{-1}$). The maximum emission peak is also hypsochromically shifted to 508 nm and quantum yield (Φ) is increased to 0.21, thus resulting in a 7-fold fluorescence turn-on (Figure 4.3). Sensor 4 is found to be too hydrophobic to be characterized in aqueous buffer, so it was deemed unsuitable for use in biological systems.

Based on PeT mechanism, increasing electron density of the BODIPY (electron acceptor) will decrease quenching efficiency of the amine lone pair of the receptor (electron donor). With a more electron rich PeT acceptor, the quantum yield of the sensor should be greater and its turn-on response may be greater. To this end, we synthesized and characterized CS5, which is a more electron rich dimethoxy analogue of CS4.[26] Similar to CS4, apo-CS5 has maximum absorption peak at 525 nm (ε = 1.6 x 10^4 $cm^{-1}M^{-1}$) and maximum emission peak at 506 nm (Φ = 0.053). Upon addition of Cu^+, the maximum emission peak remains at 506 nm, but the quantum yield of 0.30

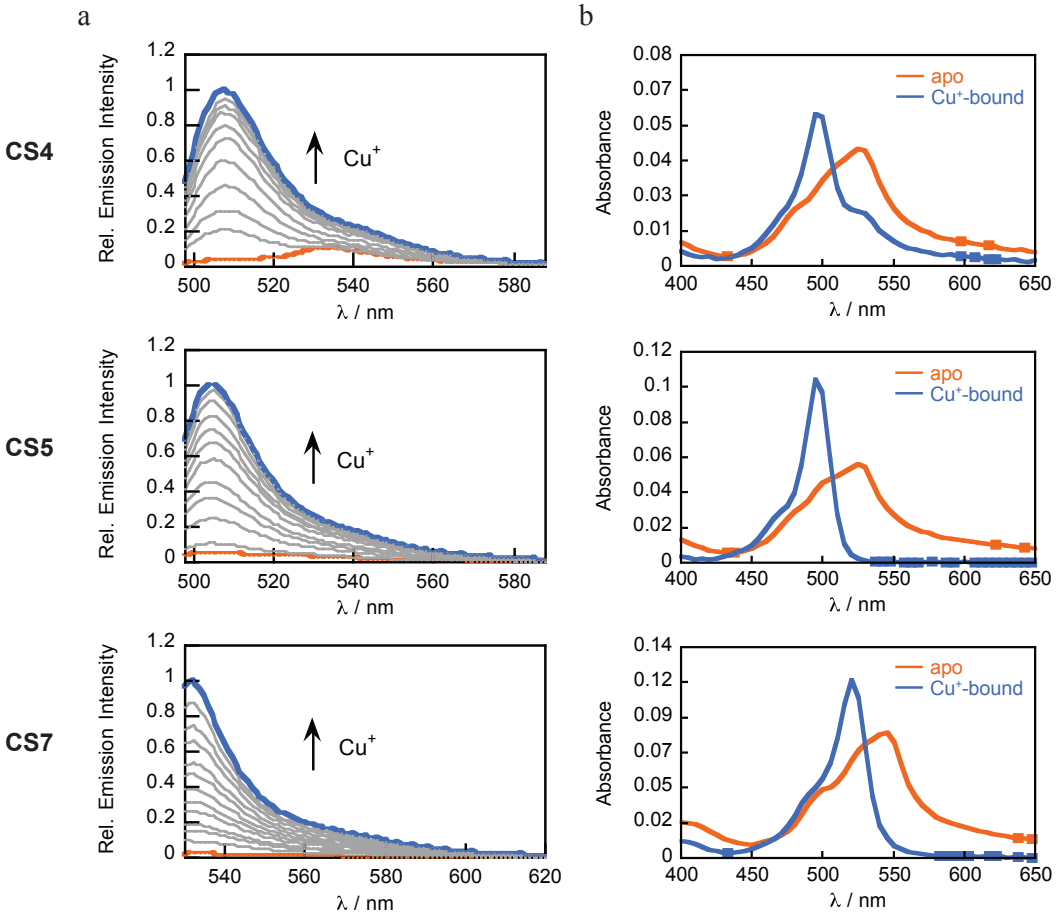


Figure 4.3. Fluorescence response to Cu^+ and absorption spectra of CS4, CS5 and CS7. a) Normalized fluorescence responses of 4 μM CS4, CS5 and CS7 to Cu^+. Spectra shown are for buffered [Cu^+] of 0, 0.4, 0.8, 1.2, 1.6, 2.0, 2.4, 2.8, 3.2, 3.6 and 4.0 μM. Spectra were acquired with excitation at 488 nm for CS4 and CS5, and 520 nm for CS7. b) Absorption spectra of 4 μM CS4, CS5 and CS7. Spectra shown are for buffered [Cu^+] of 0 μM (apo) and 4.0 μM (Cu^+ bound). All spectra were acquired in 20 mM pH 7.4 HEPES.

results in a 12-fold turn-on (Figure 4.3).

CS7 is an even more electron rich-analogue of CS5 because of additional diethyl substituents on the pyrrole. Apo-CS7 has maximum absorption peak at a longer wavelength (545 nm, ε = 2.8 x 10^4 $cm^{-1}M^{-1}$) and maximum emission peak at 535 nm (Φ = 0.033). Upon addition of Cu^+, the maximum emission peak is hypsochromically shifted to 520 nm (ε = 5.5 x 10^4 $cm^{-1}M^{-1}$) and the maximum emission peak is also hypsochromically to 532 nm. The quantum yield is increased to 0.45, which is the greatest in the series. CS7 also has the highest fluorescence turn-on of 27-fold in the series (Figure 4.3). The quantum yields and fluorescence turn-on's for CS4, CS5 and CS7 agree with the hypothesis that the greater the electron density of the fluorophore, the higher the quantum yield and the higher the turn-on with addition of Cu^+. The maximum absorbance and quantum yield can be modulated by different substituents on the pyrrole and boron center, respectively.

Because of their thioether rich receptor, the fluorescence response of CS4, CS5 and CS7 are

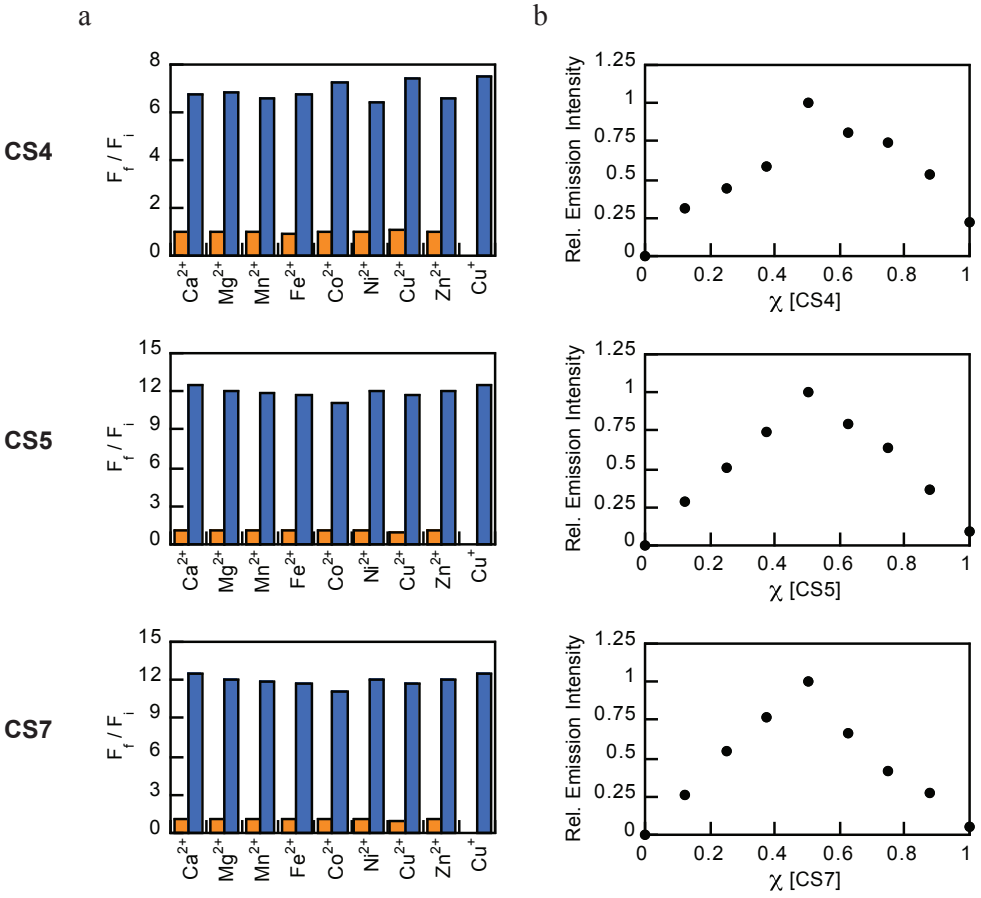
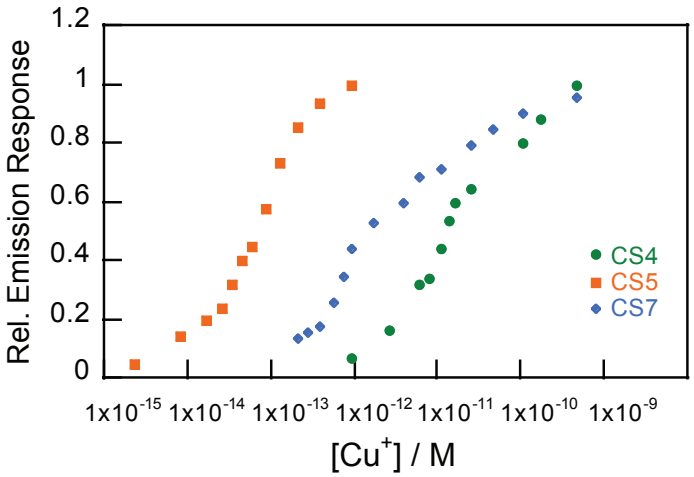


Figure 4.4. Metal selectivity and Job's plot of sensors with Cu^+. (a) Normalized fluorescence responses of CS4, CS5 and CS7 to various metal ions. Bars represent the final integrated fluorescence emission (F_f) over the initial integrated emission (F_i). Orange bars represent the addition of an excess of biologically relevant metal ion (2 mM for Ca^{2+} and Mg^{2+}; 0.5 mM for Zn^{2+}, 50 µM for all other cations) to a 2 µM solution of CS4, CS5 or CS7. Blue bars represent the subsequent addition of 2 µM of Cu^+ to the solution. (b) Job's plot of sensors and Cu^+. The total concentration of sensor and Cu^+ were kept at a constant 4 µM. The maximum fluorescence response at 0.5 mole fraction of CS4 indicates formation of a 1:1 Cu^+: sensor complex. Excitation was provided at 488 nm for CS4 and CS5, and 520 nm for CS7. The collected emission was integrated over 498 to 650 nm for CS4 and CS5, 530 to 650 nm for CS7. All spectra were acquired in 20 mM pH 7.4 HEPES.



	K_d / M
CS4	1.2 x10^{-11}
CS5	7.9 x10^{-14}
CS7	1.1 x10^{-12}

Figure 4.5. Normalized fluorescence response of 2 μM CS4, CS5 and CS7 to thiourea buffered Cu^+ solutions for K_d value determination. For CS4 and CS5, excitation was provided at 488 nm and the collected emission was integrated from 498 nm to 650 nm. For CS7, excitation was provided at 520 nm, and the collected emission was integrated over 530 to 650 nm. Spectra were acquired in 20 mM pH 7.4 HEPES.

selective towards Cu^+ and not interfered by the presence of biologically relevant amounts of other cations, including Cu^{2+} (Figure 4.4a). Job's plots indicate that the binding ratio of Cu^+ and sensor is 1:1 for all three sensors (Figure 4.4b). Despite sharing the same receptor, the apparent dissociation constants (K_d) of CS4, CS5 and CS7 are surprisingly different, spanning 2.5 orders of magnitude (Figure 4.5). Having a panel of probes with different K_d values is useful for detection of copper at different concentrations of chelatable copper, which is demonstrated in the following cell study.

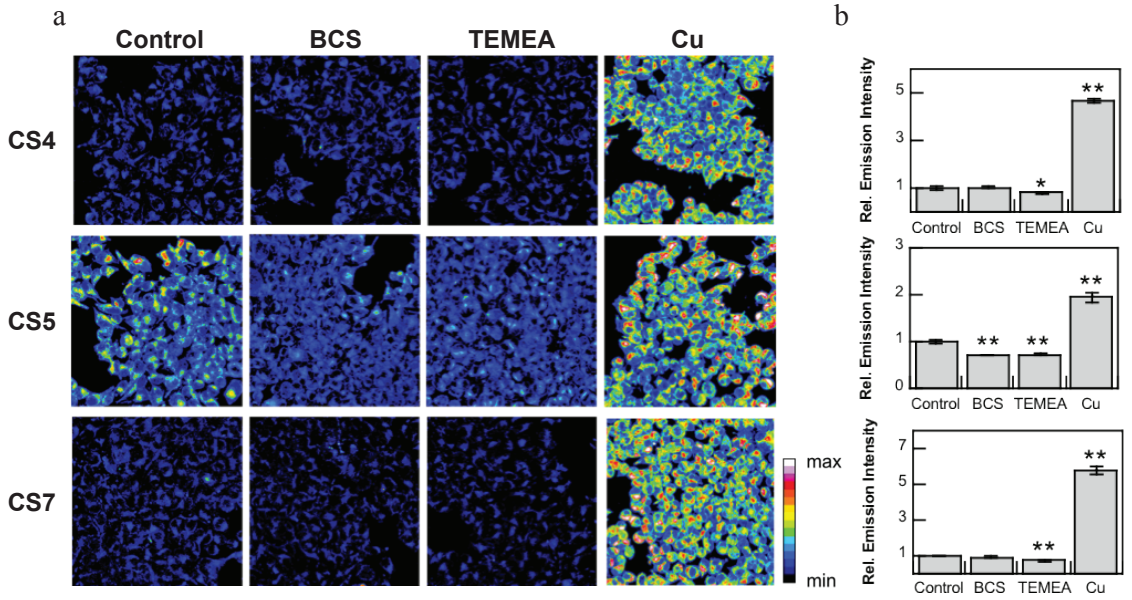


Figure 4.6. Confocal microscopy images (a) and quantifications (b) of human embryonic kidney (HEK293T) cells stained with 2 μM CS4, CS5 or CS7. Cells were grown in basal media (control), in the presence of the extracellular Cu^{2+} chelator bathocuproine disulfonate (BCS, 200 μM) for 19 h, treated with the intracellular Cu^+ chelator TEMEA (100 μM) for 15 min, or in the presence of $CuCl_2$ (100 μM) for 19 hours. Bar graphs show quantification of mean fluorescence intensity of each condition normalized to the control condition ($N = 6$ fields of cells per condition). Bars represent mean ± SEM; * denotes $p < 0.05$ and ** denotes $p < 0.01$ with respect to control cells.

4.2.3. Fluorescence Detection of Copper in Living Cells

Previously reported reversible copper sensors for live-cell Cu^+ detection, CTAP-1 and CS1, can detect copper overload in NIH-3T3 and HEK293T cells, respectively. CTAP-1 and CS1 have K_d of 39.8 pM and 3.6 pM respectively, which are too high to detect depleted copper levels.[17,18] On the other hand, CS3 has a K_d of 0.089 pM, which is low enough to detect copper depletion in HEK293T cells, but too low to detect copper overload.[12] Considering that glutathione has a K_d of 9.1 pM, sensors with K_d in the pM range or higher will not be able to compete with the glutathione for Cu^+. In order to compete with Cu^+-chaperones, such as ATOX1 (K_d of 0.017 pM), for labile Cu^+, the K_d values of sensors have to be comparable to those of chaperones.[9] Therefore, an optimal K_d is required to detect both depletion and increased levels of Cu^+.

Live-cell imaging of HEK293T cells stained with CS4, CS5 or CS7 shows that the probes are able to report Cu^+ levels in cells (Figure 4.6). Furthermore, the copper pools that each of them can detect correlate with their apparent dissociation constants, i.e. loose-binding sensor detects high copper levels while tight-binding sensors detects low copper levels. CS5 has an appropriate K_d of 0.079 pM to detect both exogenous and endogenous copper levels in HEK293T cells.

4.2.4. Photostability of Benzyl-Linked Copper Sensors

We next sought to compare photostability of the series of benyzl-linked copper sensors with CS3 by staining HEK293T cells with the sensors and continuously illuminating each field of cells with a confocal microscope. Because three different laser lines are used, the photobleaching rates of each sensor cannot be directly compared. Even after adjusting laser power and gain of the confocal microscope to lower settings for the CS3, however, CS3 photobleaches much more rapidly than the series of benzyl-linked copper sensors (Figure 4.7a). Notably, CS5 and CS7 have similar photobleaching rates while CS4 is the most photostable, suggesting that the dimethoxy groups on the boron center may contribute to lower photostability.

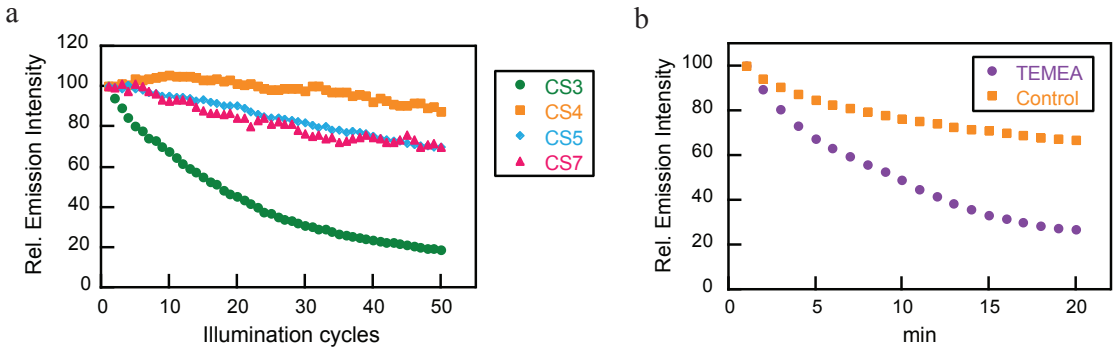


Figure 4.7. (a) Photostabilities of CS3, CS4, CS5 and CS7 are compared by the decay of fluorescence intensity upon continuous illumination of HEK293T cells stained with 2 μM copper sensors. Relative emission intensities are normalized to the initial emission intensities of each copper sensor. Confocal microscope settings were chosen to maximize signals in cells without pixel saturation. Laser power and gain for CS3 were at lower settings compared to those of CS4, CS5 and CS7. Images were taken continuously for 50 cycles. (b) Normalized fluorescence intensities from a real-time confocal imaging session of HEK293T cells stained with 2 μM CS5 were quantified. Cells were treated on stage with the intracellular Cu^+ chelator TEMEA (100 μM) or DMSO vehicle (control) at t = 0. Images were taken every minute for 20 min.

4.2.5. Real-Time Fluorescence Detection of Copper

After establishing that the benzyl-linked copper sensors have improved photostabilities, we next tested whether or not they are suitable for real-time imaging of Cu^+ levels in cells. To this end, we stained HEK293T cells with CS5 and treated them with the intracellular Cu^+ chelator tris((ethylthio)ethyl)amine (TEMEA, $K_{d(Cu^+)} = 2.95 \times 10^{-16}$ M) on stage.[27] As shown in Figure 4.7b, the decrease in fluorescence in TEMEA treated cells is greater than the vehicle control. Therefore, CS5 is able to report Cu^+ levels in real time.

4.3. Conclusions

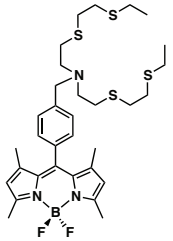
We have presented the synthesis and characterization of a series of benzyl-linked copper sensors *in vitro* and *in cellulo*. We found that the absorbance wavelength can be modulated by substituents on the pyrrole. The electronic properties of sensors can also be tuned by substituents on the pyrrole and boron center: the more electron-rich BODIPY, the less efficient PeT quenching and therefore the brighter the sensor. The series of benzyl-linked copper sensors have K_d values spanning 2.5 orders of magnitude and CS5 was found to be the first copper sensor that has the optimum K_d to detect both exogenous and endogenous copper levels in HEK293T cells. Furthermore, benzyl-linked copper sensors are more photostable than CS3. We were able to use CS5 to monitor in real time the decrease in copper level by addition of copper chelator. This is the first report of real-time imaging of copper levels in cells. The ability of these new Cu^+ sensors to monitor changes in copper levels and localization in real time will allow kinetic studies of copper trafficking.

4.4. Materials and Methods

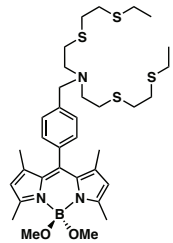
4.4.1. General Procedures and Materials

All reactions utilizing air- or moisture-sensitive reagents were performed in dried glassware under an atmosphere of dry N_2. Silica gel P60 (SiliCycle) was used for column chromatography and SiliCycle 60 F254 silica gel (precoated sheets, 0.25 mm thick) was used for analytical thin layer chromatography and visualized by fluorescence quenching under UV light or by staining with iodine. All chemicals were purchased from Sigma-Aldrich (St. Louis, MO) and used as received. BODIPY 1,[28] BODIPY 2,[25] and receptor 3,[18] and tris((ethylthio)ethyl)amine (TEMEA)[29] were synthesized according to literature procedures. ^1H and ^{13}C NMR spectra for characterization of new compounds were collected in $CDCl_3$ (Cambridge Isotope Laboratories) at 25 °C at the reported frequency at the College of Chemistry NMR Facility at the University of California, Berkeley. All chemical shifts are reported in parts per million and referenced to the residual solvent peak from $CDCl_3$ at 7.26 ppm. Splitting patterns are indicated as follows: br, broad, d, doublet, dd, doublet of doublets, dt, doublet of triplets, m, multiplet, q, quartet, s, singlet, t, triplet. High-resolution electrospray ionization mass spectral (HRMS-ESI) analyses were carried out at the College of Chemistry Mass Spectrometry Facility at the University of California, Berkeley.

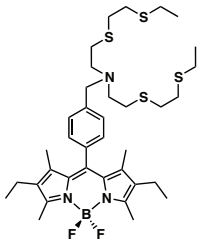
4.4.2. Small Molecule Synthesis



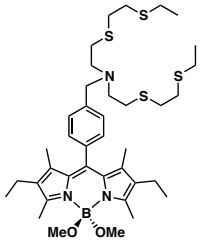
10-(4-((Bis(2-((2-(ethylthio)ethyl)thio)ethyl)amino)methyl)phenyl)-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-4-ium-5-uide (CS4). A dried round-bottom flask was charged with BODIPY 1 (200 mg, 0.54 mmol, 1.0 equiv.), K_2CO_3 (304 mg, 2.2 mmol, 4.0 equiv.) and KI (366 mg, 2.2 mmol, 4.0 equiv.). CH_3CN (20 mL) was added to the round-bottom flask and heated to 45 °C. Receptor **3** (508 mg, 1.62 mmol, 3.0 equiv.) in CH_3CN (10 mL) was then added to the heated solution in one portion. After the reaction mixture was allowed to stir at 45 °C for 5 h, the solvent was removed by rotary evaporation and the residue was taken up in ethyl acetate. The organic layer was then washed with water and brine, dried over Na_2SO_4, and the solvent was removed by rotary evaporation. Purification by flash column chromatography (silica gel, 0%-2% ethyl acetate/dichloromethane) gave CS4 as an orange film (177 mg, 0.27 mmol, 50% yield). ^{1}H NMR (300 MHz, chloroform-*d*) δ 7.47 (d, *J* = 7.8 Hz, 2H), 7.22 (d, *J* = 7.8 Hz, 2H), 5.97 (s, 2H), 3.71 (s, 2H), 2.91 – 2.61 (m, 16H), 2.61 – 2.47 (m, 10H), 1.38 (s, 6H), 1.25 (t, *J* = 7.4 Hz, 6H). HRMS-ESI calculated for $C_{32}H_{47}N_3BF_2S_4$ ($M+H^+$), 650.2708, found, 650.2720.



10-(4-((Bis(2-((2-(ethylthio)ethyl)thio)ethyl)amino)methyl)phenyl)-5,5-dimethoxy-1,3,7,9-tetramethyl-5*H*-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-4-ium-5-uide (CS5). CS4 (70 mg, 0.11 mmol, 1.0 equiv.) was dissolved in 2 mL of dichloromethane in a 20 mL scintillation vial. Anhydrous $AlCl_3$ (73 mg, 0.55 mg, 5.0 equiv.) was then added to the vial and the reaction mixture was sonicated for 5 min. 1 mL of methanol was added to the reaction mixture and stirred for 5 min at room temperature. The reaction mixture was then poured into water and extracted with ethyl acetate (2x). The combined organic layer was washed with water and brine, dried over Na_2SO_4, and the solvent was removed by rotary evaporation. Purification by flash column chromatography (silica gel, 20%-60% ethyl acetate/dichloromethane) gave CS5 as orange/red oil (52 mg, 0.077 mmol, 70% yield). ^{1}H NMR (400 MHz, chloroform-*d*) δ 7.46 (d, *J* = 7.6 Hz, 2H), 7.24 (d, *J* = 7.6 Hz, 2H), 5.95 (s, 2H), 3.72 (s, 2H), 2.94 (s, 6H), 2.81 – 2.61 (m, 16H), 2.56 (q, *J* = 7.4 Hz, 4H), 2.52 (s, 6H), 1.39 (s, 6H), 1.25 (t, *J* = 7.4 Hz, 6H). HRMS-ESI calculated for $C_{34}H_{53}O_2N_3BS_4$ ($M+H^+$), 674.3108, found, 674.3115.



10-(4-((Bis(2-((2-(ethylthio)ethyl)thio)ethyl)amino)methyl)phenyl)-2,8-diethyl-5,5-difluoro-1,3,7,9-tetramethyl-5*H*-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-4-ium-5-uide (4). A dried round-bottom flask was charged with BODIPY **2** (100 mg, 0.23 mmol, 1.0 equiv.), K_2CO_3 (129 mg, 0.93 mmol, 4.0 equiv.) and KI (154 mg, 0.93 mmol, 4.0 equiv.). CH_3CN (10 mL) was added to the round-bottom flask and heated to 50 °C. Receptor **3** (200 mg, 0.64 mmol, 2.8 equiv.) in CH_3CN (5 mL) was then added to the heated solution in one portion. After the reaction mixture was heated to reflux for 2.5 h, the solvent was removed by rotary evaporation and the residue was taken up in ethyl acetate. The organic layer was then washed with water and brine, dried over Na_2SO_4, and the solvent was removed by rotary evaporation. Purification by two consecutive flash column chromatography (silica gel, 7.5% ethyl acetate/hexanes) gave 4 as an orange film (88 mg, 0.12 mmol, 54% yield). ^{1}H NMR (300 MHz, chloroform-*d*) δ 7.46 (d, *J* = 7.2 Hz, 2H), 7.23 (d, *J* = 7.2 Hz, 2H), 3.72 (s, 2H), 2.82 – 2.62 (m, 16H), 2.56 (q, *J* = 7.2 Hz, 4H), 2.53 (s, 6H), 2.29 (q, *J* = 7.5 Hz, 4H), 1.31 – 1.22 (m, 12H), 0.97 (t, *J* = 7.5 Hz, 6H).



10-(4-((Bis(2-((2-(ethylthio)ethyl)thio)ethyl)amino)methyl)phenyl)-2,8-diethyl-5,5-dimethoxy-1,3,7,9-tetramethyl-5*H*-dipyrrolo[1,2-*c*:2',1'-*f*][1,3,2]diazaborinin-4-ium-5-uide (CS7). Product **4** (63 mg, 0.091 mmol, 1.0 equiv.) was dissolved in 2 mL of dichloromethane in a 20 mL scintillation vial. Anhydrous $AlCl_3$ (36 mg, 0.27 mg, 3.0 equiv.) was then added to the vial and the reaction mixture was sonicated for 5 min. Methanol (1 mL) was added to the reaction mixture and stirred for 5 min at room temperature. The reaction mixture was then poured into water and extracted with ethyl acetate (2x). The combined organic layer was washed with water and brine, dried over Na_2SO_4, and the solvent was removed by rotary evaporation. Purification by flash column chromatography (silica gel, 20%-35% ethyl acetate/dichloromethane) gave CS7 as red oil (28 mg, 0.038 mmol, 42% yield). ^{1}H NMR (300 MHz, chloroform-*d*) δ 7.45 (d, *J* = 7.2 Hz, 2H), 7.24 (d, *J* = 7.2 Hz, 2H), 3.73 (s, 2H), 2.91 (s, 6H), 2.83 – 2.62 (m, 16H), 2.56 (q, *J* = 7.5 Hz, 4H), 2.50 (s, 6H), 2.30 (q, *J* = 7.5 Hz, 4H), 1.29 (s, 6H), 1.25 (t, *J* = 7.2 Hz, 6H), 0.98 (t, *J* = 7.5 Hz, 6H). HRMS-ESI calculated for $C_{38}H_{61}O_2N_3BS_4$ ($M+H^+$), 730.3734, found, 730.3740.

4.4.3. Spectroscopic Materials and Methods

Millipore water was used to prepare all aqueous solutions. For all studies, sensors were diluted in aqueous buffers from 2 mM stock in DMSO. All spectroscopic measurements were

performed in 20 mM HEPES buffer, pH 7.4. Fluorescence spectra were obtained using a Quanta Master 4 L-format scanning spectrofluorometer equipped with an LPS-220B 75 W xenon lamp and power supply, A-1010B lamp housing with an integrated igniter, switchable 814 photon-counting/analog photomultiplier detection unit, and MD5020 motor driver (Photon Technology International, Inc.) and UV spectra were acquired using a Cary Bio50 UV spectrophotometer (Varian). Samples for emission and absorption measurements were contained in 1 cm × 0.1 cm quartz cuvette (Starna). Absorption and emission spectra for CS4, CS5, CS7 and fluorescein were obtained over a range of concentrations where a linear correlation between concentration and absorption was observed. Extinction coefficients were obtained by plotting absorbance versus sensor concentration. Quantum yields were determined by reference to fluorescein ($\Phi_{standard}$ = 0.79 in 0.1 M NaOH) and were calculated according to the equation $\Phi_{sample} = \Phi_{standard}$ ($Grad_{sample}$/ $Grad_{standard}$) (ηsample/ηstandard), where Φ is quantum yield, Grad is the slope of the plot of absorbance versus integrated emission intensity, and η is the refractive index of the solvent. The binding affinities of Cu^+ to CS4, CS5 and CS7 were measured according to literature procedure.[12,18] Thiourea was used as a competitive ligand to provide buffered Cu^+ solutions. Stability constants for thiourea binding were taken from the literature: $\beta_{12} = 2.0 \times 10^{12}$, $\beta_{13} = 2.0 \times 10^{14}$ and $\beta_{14} = 3.4 \times 10^{15}$.[30] Cu^+ was added in the form of $[Cu(MeCN)_4][PF_6]$ dissolved in an acetonitrile stock solution (2 mM). The apparent dissociation constants (K_d) were determined using the following equation: $(F - F_{min}) / (F_{max} - F_{min}) = [Cu^+]/(K_d + [Cu^+])$, where F is the observed fluorescence, F_{max} is the fluorescence of the Cu^+:CS complex, and F_{min} is the fluorescence for the apo-sensor.

4.4.4. Preparation and Staining of Cell Culture

HEK 293T cells were maintained in exponential growth as a monolayer in Dulbecco's Modified Eagle Mediaum (DMEM, Invitrogen) supplemented with 10% fetal bovine serum (FBS, Hyclone), and incubated at 37 °C in 5% CO_2. Two days before imaging, the cells were passaged and plated in phenol red-free medium on a 4-well Lab Tek borosilicate chambered coverglass slides (Nunc) and allowed to grow to 60-80% confluence. Before the imaging experiments, cells were incubated with 2 μM sensor (from 2 mM stock in DMSO) in DPBS at 37 °C for 15 min, washed with DPBS (2x), and imaged in DPBS.

4.4.5. Confocal Imaging Experiments

Confocal fluorescence imaging studies were performed with a Zeiss laser scanning microscope 710 with a 40× water objective lens, with Zen 2009 software (Carl Zeiss). CS4 and CS5 were excited at 488 nm with an Argon laser, and emission collected using a META detector between 493 and 630 nm. CS7 was excited at 514 nm with an Argon laser, and emission collected using a META detector between 520 and 658 nm. CS3 was excited at 543 nm with a He-Ne laser, and emission collected using a META detector between 548 and 680 nm. The cells were imaged at 37 °C. 15 × 2 μM z-stacks were collected, ensuring that all of the cellular fluorescence was included within the z-stacks throughout the course of the experiment. Image analysis was performed in ImageJ (National Institute of Health) and Zen 2009 software (Carl Zeiss).

4.5. References

1. Camakaris, J., Voskoboinik, I., and Mercer, J. F. (1999) Molecular mechanisms of copper homeostasis, *Biochem. Biophys. Res. Commun. 261*, 225-232.

2. Kim, B. E., Nevitt, T., and Thiele, D. J. (2008) Mechanisms for copper acquisition, distribution and regulation, *Nat. Chem. Biol. 4*, 176-185.

3. Lutsenko, S. (2010) Human copper homeostasis: a network of interconnected pathways, *Curr. Opin. Chem. Biol. 14*, 211-217.

4. Turski, M. L., and Thiele, D. J. (2009) New roles for copper metabolism in cell proliferation, signaling, and disease, *J. Biol. Chem. 284*, 717-721.

5. Barnham, K. J., Masters, C. L., and Bush, A. I. (2004) Neurodegenerative diseases and oxidative stress, *Nat. Rev. Drug Discov. 3*, 205-214.

6. Madsen, E., and Gitlin, J. D. (2007) Copper and iron disorders of the brain, *Annu. Rev. Neurosci. 30*, 317-337.

7. Que, E. L., Domaille, D. W., and Chang, C. J. (2008) Metals in neurobiology: probing their chemistry and biology with molecular imaging, *Chem. Rev. 108*, 1517-1549.

8. Waggoner, D. J., Bartnikas, T. B., and Gitlin, J. D. (1999) The role of copper in neurodegenerative disease, *Neurobiol. Dis. 6*, 221-230.

9. Banci, L., Bertini, I., Ciofi-Baffoni, S., Kozyreva, T., Zovo, K., and Palumaa, P. (2010) Affinity gradients drive copper to cellular destinations, *Nature 465*, 645-648.

10. Rosenzweig, A. C., and O'Halloran, T. V. (2000) Structure and chemistry of the copper chaperone proteins, *Curr. Opin. Chem. Biol. 4*, 140-147.

11. Domaille, D. W., Que, E. L., and Chang, C. J. (2008) Synthetic fluorescent sensors for studying the cell biology of metals, *Nat. Chem. Biol. 4*, 168-175.

12. Dodani, S. C., Domaille, D. W., Nam, C. I., Miller, E. W., Finney, L. A., Vogt, S., and Chang, C. J. (2011) Calcium-dependent copper redistributions in neuronal cells revealed by a fluorescent copper sensor and X-ray fluorescence microscopy, *Proc. Natl. Acad. Sci. USA 108*, 5980-5985.

13. Dodani, S. C., Leary, S. C., Cobine, P. A., Winge, D. R., and Chang, C. J. (2011) A targetable fluorescent sensor reveals that copper-deficient SCO1 and SCO2 patient cells prioritize mitochondrial copper homeostasis, *J. Am. Chem. Soc. 133*, 8606-8616.

14. Domaille, D. W., Zeng, L., and Chang, C. J. (2010) Visualizing ascorbate-triggered release of labile copper within living cells using a ratiometric fluorescent sensor, *J. Am. Chem. Soc. 132*, 1194-1195.

15. Hirayama, T., Van de Bittner, G. C., Gray, L. W., Lutsenko, S., and Chang, C. J. (2012) Near-infrared fluorescent sensor for *in vivo* copper imaging in a murine Wilson disease model, *Proc.*

Natl. Acad. Sci. USA 109, 2228-2233.

16. Taki, M., Iyoshi, S., Ojida, A., Hamachi, I., and Yamamoto, Y. (2010) Development of highly sensitive fluorescent probes for detection of intracellular copper(I) in living systems, *J. Am. Chem. Soc. 132*, 5938-5939.

17. Yang, L. C., McRae, R., Henary, M. M., Patel, R., Lai, B., Vogt, S., and Fahrni, C. J. (2005) Imaging of the intracellular topography of copper with a fluorescent sensor and by synchrotron x-ray fluorescence microscopy, *Proc. Natl. Acad. Sci. USA 102*, 11179-11184.

18. Zeng, L., Miller, E. W., Pralle, A., Isacoff, E. Y., and Chang, C. J. (2006) A selective turn-on fluorescent sensor for imaging copper in living cells, *J. Am. Chem. Soc. 128*, 10-11.

19. Shaner, N. C., Steinbach, P. A., and Tsien, R. Y. (2005) A guide to choosing fluorescent proteins, *Nat. Methods 2*, 905-909.

20. Shaner, N. C., Lin, M. Z., McKeown, M. R., Steinbach, P. A., Hazelwood, K. L., Davidson, M. W., and Tsien, R. Y. (2008) Improving the photostability of bright monomeric orange and red fluorescent proteins, *Nat. Methods 5*, 545-551.

21. Callan, J. F., de Silva, A. P., and Magri, D. C. (2005) Luminescent sensors and switches in the early 21st century, *Tetrahedron 61*, 8551-8588.

22. De Silva, A. P., Gunaratne, H. Q. N., Gunnlaugsson, T., Huxley, A. J. M., McCoy, C. P., Rademacher, J. T., and Rice, T. E. (1997) Signaling recognition events with fluorescent sensors and switches, *Chem. Rev. 97*, 1515-1566.

23. Miura, T., Urano, Y., Tanaka, K., Nagano, T., Ohkubo, K., and Fukuzumi, S. (2003) Rational design principle for modulating fluorescence properties of fluorescein-based probes by photoinduced electron transfer, *J. Am. Chem. Soc. 125*, 8666-8671.

24. Urano, Y., Kamiya, M., Kanda, K., Ueno, T., Hirose, K., and Nagano, T. (2005) Evolution of fluorescein as a platform for finely tunable fluorescence probes, *J. Am. Chem. Soc. 127*, 4888-4894.

25. Garcia-Moreno, I., Costela, A., Campo, L., Sastre, R., Amat-Guerri, F., Liras, M., Arbeloa, F. L., Prieto, J. B., and Arbeloa, I. L. (2004) 8-phenyl-substituted dipyrromethene center·BF_2 complexes as highly efficient and photostable laser dyes, *J. Phys. Chem. A 108*, 3315-3323.

26. Gabe, Y., Ueno, T., Urano, Y., Kojima, H., and Nagano, T. (2006) Tunable design strategy for fluorescence probes based on 4-substituted BODIPY chromophore: improvement of highly sensitive fluorescence probe for nitric oxide, *Anal. Bioanal. Chem. 386*, 621-626.

27. Ambundo, E. A., Deydier, M. V., Grall, A. J., Aguera-Vega, N., Dressel, L. T., Cooper, T. H., Heeg, M. J., Ochrymowycz, L. A., and Rorabacher, D. B. (1999) Influence of coordination geometry upon copper(II/I) redox potentials. Physical parameters for twelve copper tripodal ligand complexes, *Inorg. Chem. 38*, 4233-4242.

28. Michel, B. W., Lippert, A. R., and Chang, C. J. (2012) A reaction-based fluorescent probe for

selective imaging of carbon monoxide in living cells using a palladium-mediated carbonylation, *J. Am. Chem. Soc. 134*, 15668-15671.

29. Cooper, T. H., Mayer, M. J., Leung, K. H., Ochrymowycz, L. A., and Rorabacher, D. B. (1992) Kinetic and thermodynamic measurements on branched amino polythiaether ligands - a family of complexing agents analogous to EDTA and NTA exhibiting enhanced selectivity for copper(II), *Inorg. Chem. 31*, 3796-3804.

30. Martell, A. E., and Smith, R. M. (1989) Critical stability constants. Plenum Press, New York.

Appendix

Preparation of a Library of PEGylated Intrinsically Disordered Proteins

The studies described in this appendix were done in collaboration with Alexander C. Hoepker, Ioana L. Aanei and Jenna M. L. Bernard.

A.1. Introduction

Protein-based nanomaterials have emerged as a promising class of drug delivery agents.[1-3] Their advantages include the following: 1) monodisperse protein carriers can be recombinantly produced, unlike synthetic polymers, 2) biophysical properties can be tuned by protein engineering, 3) unique conjugation handles on proteins can be inserted for site-specific and/or high cargo loading, and 4) inherent biocompatibility. Current efforts have focused on self-assembled proteins and globular proteins.[2] Even though there is ongoing research on self-assembling intrinsically disordered proteins and peptides (IDP) as drug delivery scaffolds,[4] IDP monomers have not been explored.

Unlike self-assembled and globular proteins, IDPs lack canonical secondary structures and are flexible (see Section 2.1.1 for more information on IDPs). A key parameter in the success of drug delivery is the ability of the agent to diffuse from the circulation by crossing the openings, called fenestrations, of the epithelial barrier and subsequent biodistribution in different organs.[5] Even though agent size has been shown to be a critical factor in determining its biodistribution,[6] other factors, such as flexibility, are much less studied and understood. A few reports have implicated that the flexibility of filomicelles influences their biodistributions.[7,8] We envision that we can take advantage of the flexibility of IDPs, which may facilitate diffusion through fenestrations, for drug delivery applications.

Herein, we utilized a sequence from the neurofilament heavy side-arm (NFH-SA), which is a fully disordered IDP described in Section 2.1.1. Due to its lack of hydrophobic core, chemical modification

Full length IDP construct

SEFTSMSTHIKVKSEEKIKVVEKSEKETVIVEEQTEEIQVTEEVTEEEDKEAQGEEEEEAEEGGEEAATTSPPAEEAASPEKETKSPVKEEAKSPAE
AKSPAEAKSPAEAKSPAEVKSPAVAKSPAEVKSPAEVKSPAEAKSPAEAKSPAEVKSPATVKSPGEAKSPAEAKSPAEVKSPVEAKSPAEAKSPA
SVKSPGEAKSPAEAKSPAEVKSPATVKSPVEAKSPAEVKSPVTVKSPAEAKSPVEVKSPASVKSPSEAKSPAGAKSPAEAKSPVVAKSPAEAKSP
AEAKPPAEAKSPAEAKSPAEAKSPAEAKSPAEAKSPVEVKSPEKAKSPVKEGAKSLAEAKSPEKAKSPVKEEIKPPAEVKSPEKAKSPMKEEAK
SPEKAKTLDVKSPEAKTPAKEEAKRPADIRSPEQVKSPAKEEAKSPEKEETRTEKVAPKKEEVKSPVEEVKAKEPPKKVEEEKTPATPKTEVKESK
KDEAPKEAQKPKAEEKEPLTEKPKDSPGEAKKEEAKEKKAAAPEEETPAKLGVKEEAKPKEKAEDAKAKEPSKPSEKEKPKKEEVPAAPEKKD
TKEEKTTESKKPEEKPKMEAKAKEEDKGLPQEPSKPKTEKAEKSSSTDQKDSQPSEKAPEDKLYWCLVPRGSWSHPQFEKASHHHHHH

Truncated IDP construct

AWRGSPWAEAKSPAEAKSPAEVKSPAVAKSPAEVKSPAEVKSPAEAKSPAEAKSPAEVKSPATVKSPGEAKSPAEAKSPAEVKSPVEAKSPAEA
KSPASVKSPGEAKSPAEAKSPAEVKSPATVKSPVEAKSPAEVKSPVTVKSPAEAKSPVEVKSPYWCA

Figure A.1. Engineered intrinsically disordered protein (IDP) sequences derived from disordered C-terminal sidearm domain of the heavy subunit of the neurofilament complex. The full length IDP construct was described in section 2.2.1, and the truncated sequence is highlighted in green. An N-terminal alanine (grey), a tyrosine (orange), a tryptophan (orange) and a cysteine (light blue) were introduced in the truncated construct used in this chapter. This truncated construct contains 25 lysines.

of NFH-SA along the entire protein is possible. Poly(ethylene glycol) (PEG) has been extensively used to passivate drug delivery agents to increase circulating lifetimes in blood. The size and surface density of PEG have been found to be an important parameter in reducing the rate of clearance.[9-12] Therefore, our goal is to conjugate NFH-SA with various numbers of PEG chains of different sizes, and characterize the biophysical properties of these modified IDPs for future use in drug delivery systems.

A.2. Results and Discussion

A.2.1. Truncated IDP Construct Design

We have selected a 150 residue highly repetitive sequence from NFH-SA (Figure A.1). This sequence has a theoretical length of ca. 50 nm when stretched linearly, and it presents 25 lysine residues as sites for protein modification. We have introduced a unique cysteine near the C-terminus as a site-specific conjugation handle. Since this sequence lacks aromatic amino acid residues, a tyrosine and a tryptophan were also added for the determination of protein concentration by absorbance at 280 nm, resulting in the sequence shown in the bottom of Figure A.1.

Because IDPs are often degraded in live cells, it is difficult to produce a large amount of IDPs.[13] As described in Section 2.4.3, the yield of the full length NFH-SA is quite low. With the aims of increasing overexpression levels in *Escherichia coli* and facilitating protein purification, we

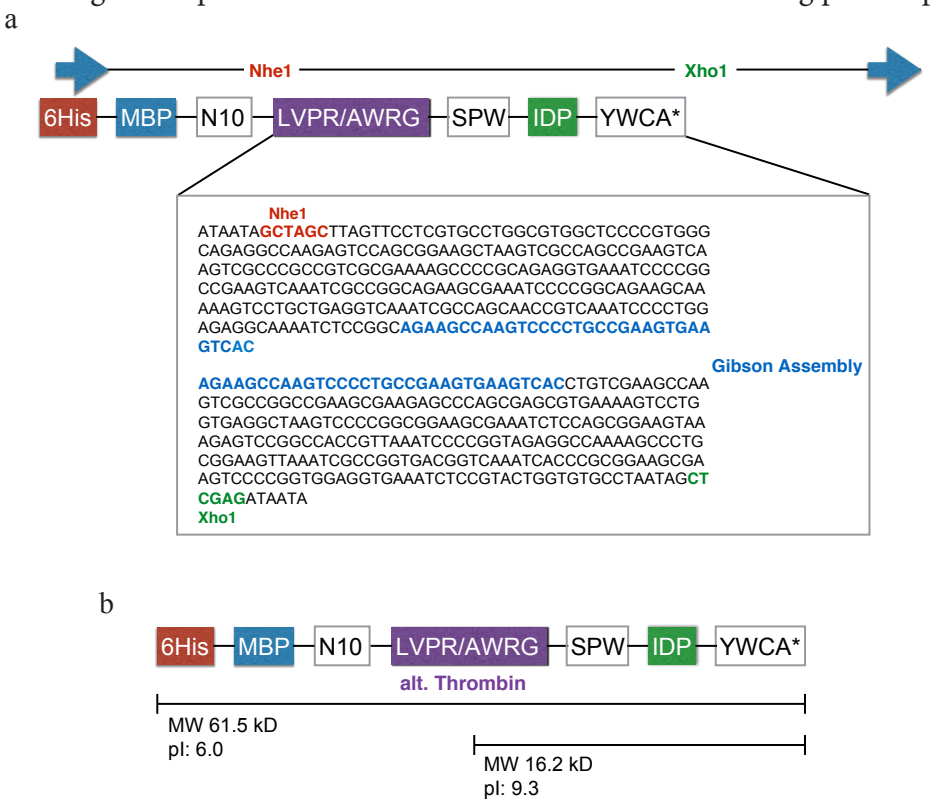


Figure A.2. Truncated IDP expression construct containing 6xHistidine-tag (6His, red), maltose-binding protein (MBP, blue), 10xAsparagine (N10) linker and an alternative thrombin cleavage site (purple). (a) The gene for MBP-IDP expression was constructed by Gibson assembly of two gBlocks gene fragments. Overlapping gene sequence for Gibson assembly is shown in blue. (b) Molecular weights and pIs of MBP-IDP and IDP after thrombin cleavage are shown. * denotes stop codon.

used maltose-binding protein (MBP) in conjunction with 6xHis as fusion tags at the N-terminus of IDP.[14-16] A thrombin cleavage site was introduced to cleave the fusion tags from IDP (Figure A.2). The most commonly used thrombin cleavage site has the sequence Leu-Val-Pro-Arg-Gly-Ser, and the cleavage occurs between the arginine and the glycine residues. Glycine, however, is not an ideal handle for N-terminal modifications developed in the Francis lab. Therefore, we used an alternative thrombin cleavage site, Leu-Val-Pro-Arg-Ala-Trp, which will result in an N-terminal alanine after thrombin cleavage.[17] The cleaved target IDP has a molecular weight of 16.2 kDa and a theoretical pI of 9.3, while the uncleaved protein (MBP-IDP) has a molecular weight of 61.5 kDa and a theoretical pI of 6.0 (Figure A.2b). The details of cloning and overexpression of the MBP-IDP

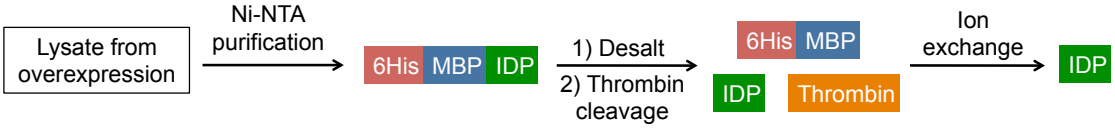


Figure A.3. Scheme for the purification of IDP.

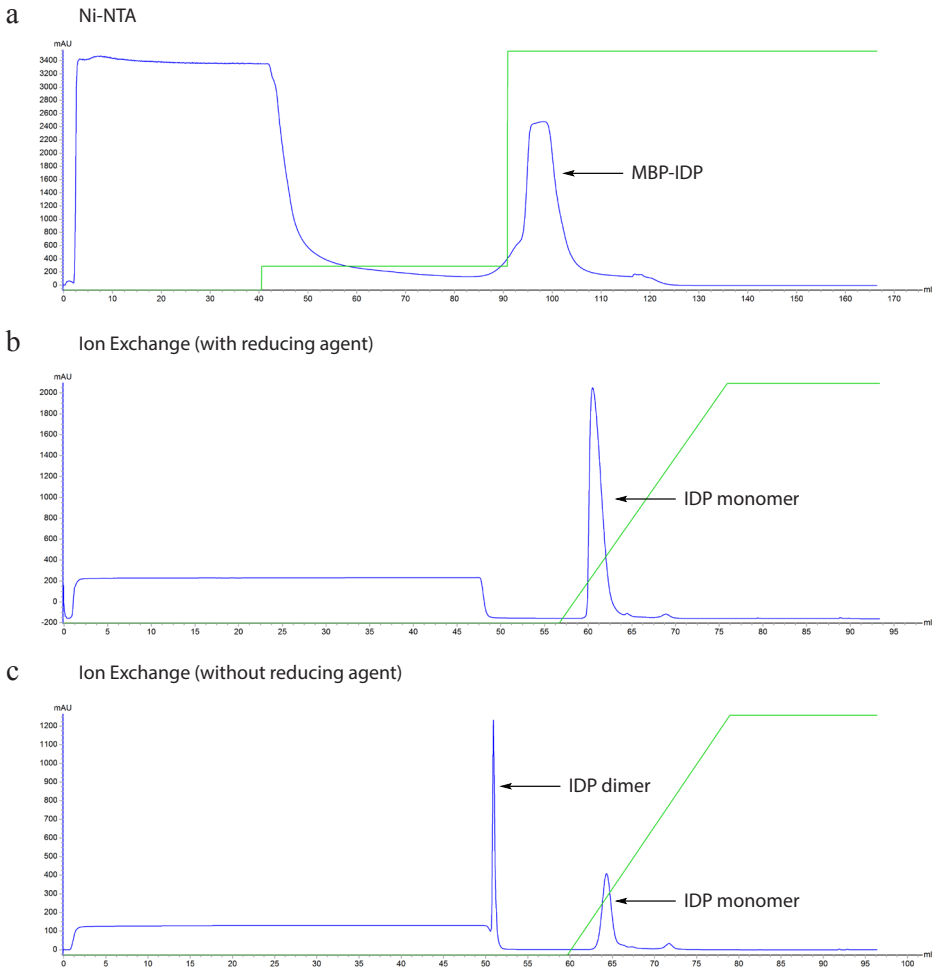


Figure A.4. Fast protein liquid chromatography (FPLC) chromatograms of IDP purification. (a) Ni-NTA purification of MBP-IDP is shown (see Figure A.4a, top, for LCMS and Figure A.4b, lane 2, for Coomassie-stained SDS-PAGE of the Ni-NTA purified MBP-IDP). Ion-exchange purification of IDP with (b) and without (c) 10 mM of β-mercaptoethanol as a reducing agent (see Figure A.4a, middle, for LCMS). Addition of reducing agent eliminates the IDP dimer peak. Absorbance at 280 nm (blue traces, y-axes) and conductivity (green traces) are shown.

fusion are described in Section A.4.2 and A.4.3 respectively.

A.2.2. Purification and Characterization of IDP

The IDP purification protocol includes Ni-NTA affinity chromatography to isolate MBP-IDP from cell lysates, thrombin cleavage and ion exchange chromatography to purify target IDP from MBP and thrombin (Figure A.3). After Ni-NTA purification (Figure A.4a), the isolation of MBP-IDP fusion was confirmed by mass spectrometry and SDS-PAGE (Figure A.5a, top, and A.5b, lane 2). Imidazole was removed using a desalting column prior to thrombin cleavage. In the final ion exchange purification, monomeric IDP was eluted at around 200 mM NaCl in the presence of β-mercaptoethanol (Figure A.4b). Without addition of reducing agent, we observed significant dimeric formation that elutes with little retention at the tail end of the loading step (Figure A.4c). IDP dimerization was also observed in SDS-PAGE when no reducing agent was added to the protein sample (Figure A.5b, lanes 1 and 3). The reversibility of the dimerization using reducing agents supports the hypothesis that IDP is dimerized through a disulfide bond formed between the cysteine residues introduced in the sequence. A yield of 13 mg IDP per liter culture can be obtained using the described method.

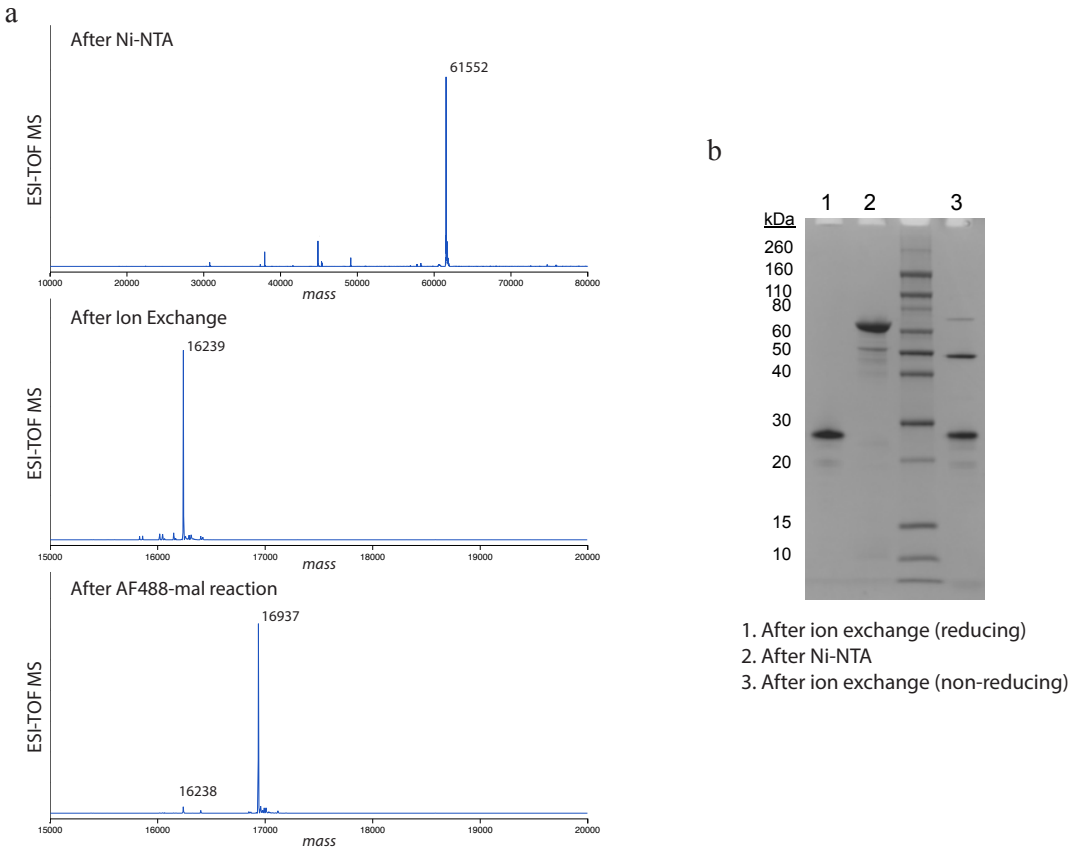


Figure A.5. Deconvoluted ESI-TOF MS spectra (a) and Coomassie-stained SDS-PAGE (b) of purified MBP-IDP and IDP. Ni-NTA purified MBP-IDP (a, top, and b, lane 2) and ion-exchanged purified IDP with reducing agent (a, middle, and b, lanes 1 and 3) are shown. The same ion-exchanged purified IDP samples stored for over a week at 4 ˚C were used in lane 1 and lane 3, but the SDS-PAGE loading buffer of lane 1 contained DTT while that of lane 3 did not contain any reducing agents. (a, bottom) Deconvoluted ESI-TOF MS spectra of IDP conjugation with Alexa Fluor 488 maleimide (AF488-mal).

A.2.3. Modification of IDP

With the goal of investigating the effect of PEGylation on the conformation of IDP, we constructed an IDP library with four different amine-reactive PEG reagents: a monodisperse 0.5 kDa PEG NHS-ester, a 1.2 kDa PEG NHS-ester, a 4 kDa branched PEG TFP-ester, and a polydisperse 5 kDa PEG-NHS ester. Since many biophysical characterization techniques, such as fluorescence anisotropy and fluorescence correlation spectroscopy (FCS), require controlled fluorescence labeling, the unique cysteine near the C-terminus of IDP was alkylated with Alexa Fluor 488 maleimide to excellent yield as confirmed by mass spectrometry (Figure A.5a, bottom). We then modified the fluorescently labeled IDP with various concentrations of PEGylating agents in order to produce a library with different extents of PEG addition. As shown in the mass spectra and SDS-PAGE of the

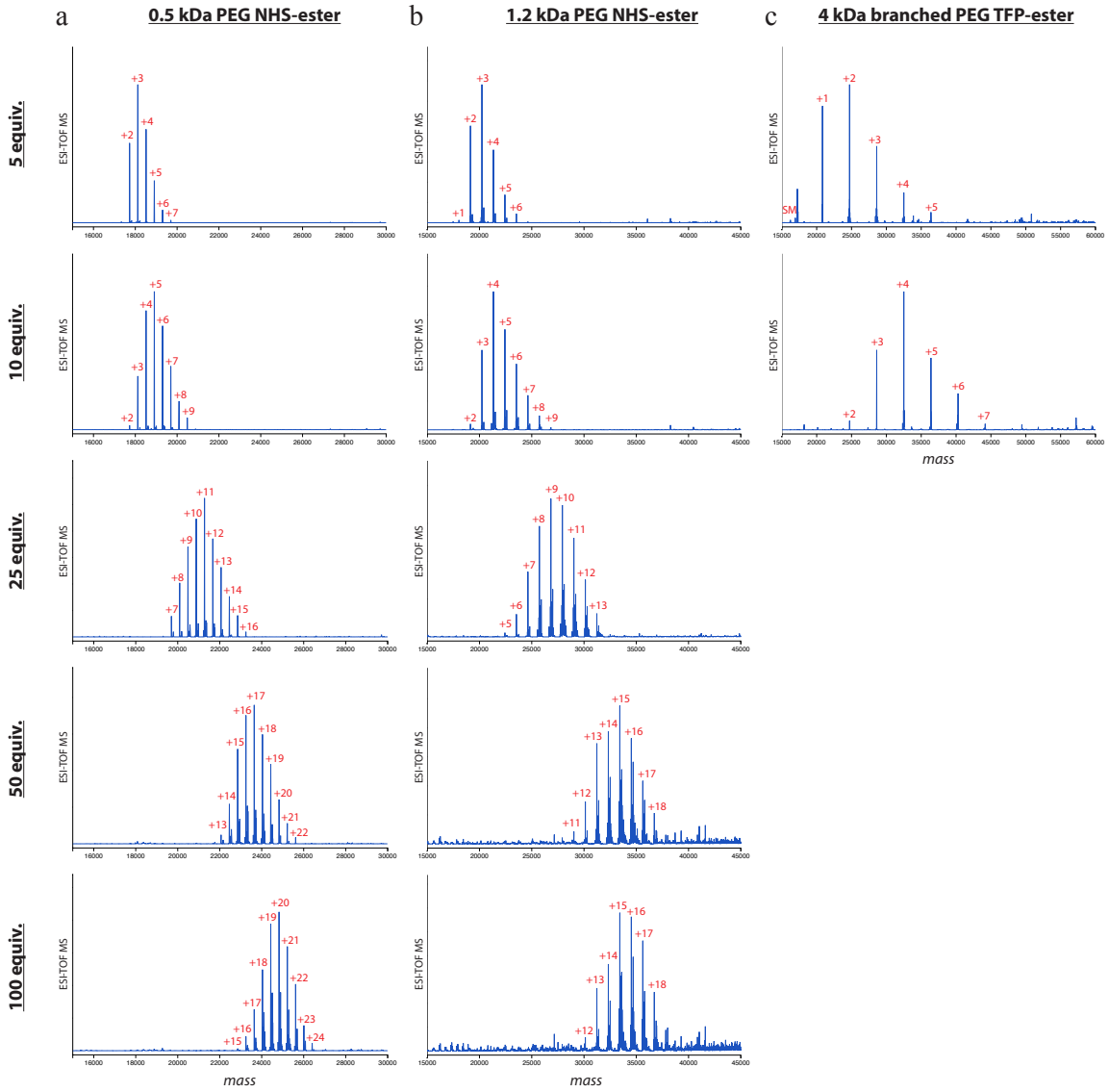


Figure A.6. Deconvoluted ESI-TOF MS spectra of IDP samples modified with various equivalents of monodisperse 0.5 kDa PEG NHS-ester, 1.2 kDa PEG NHS-ester, and 4 kDa branched PEG TFP-ester with respect to IDP. The number of PEG chains added is indicated in red. The MS spectra of IDP modified with 25, 50 amd 100 equivalents of 4 kDa branched PEG TFP-ester could not be deconvoluted.

library (Figures A.6 and A.7), each reaction condition led to a statistical distribution of products because any number of the 25 lysines and the N-terminus can be modified. The distributions in the mass spectra can be fitted with Gaussian curves to quantify the average number of PEG chains added and the effective protein charge (Z_{eff}, Figure A.8 and Table A.1). As expected, the numbers of PEG chains added increase with increasing concentrations of monodisperse PEGylation reagents. Comparing between the three different reagents, fewer PEG chains were conjugated onto IDP as the size of the PEG chain increases because of steric bulk. With the smallest 0.5 kDa PEG, up to 24 PEG chain addition could be observed with mass spectrometry (Figure A.6a), indicating that our library spans a wide extent of PEGylation for future biophysical investigations of how PEGylation affects IDP conformation.

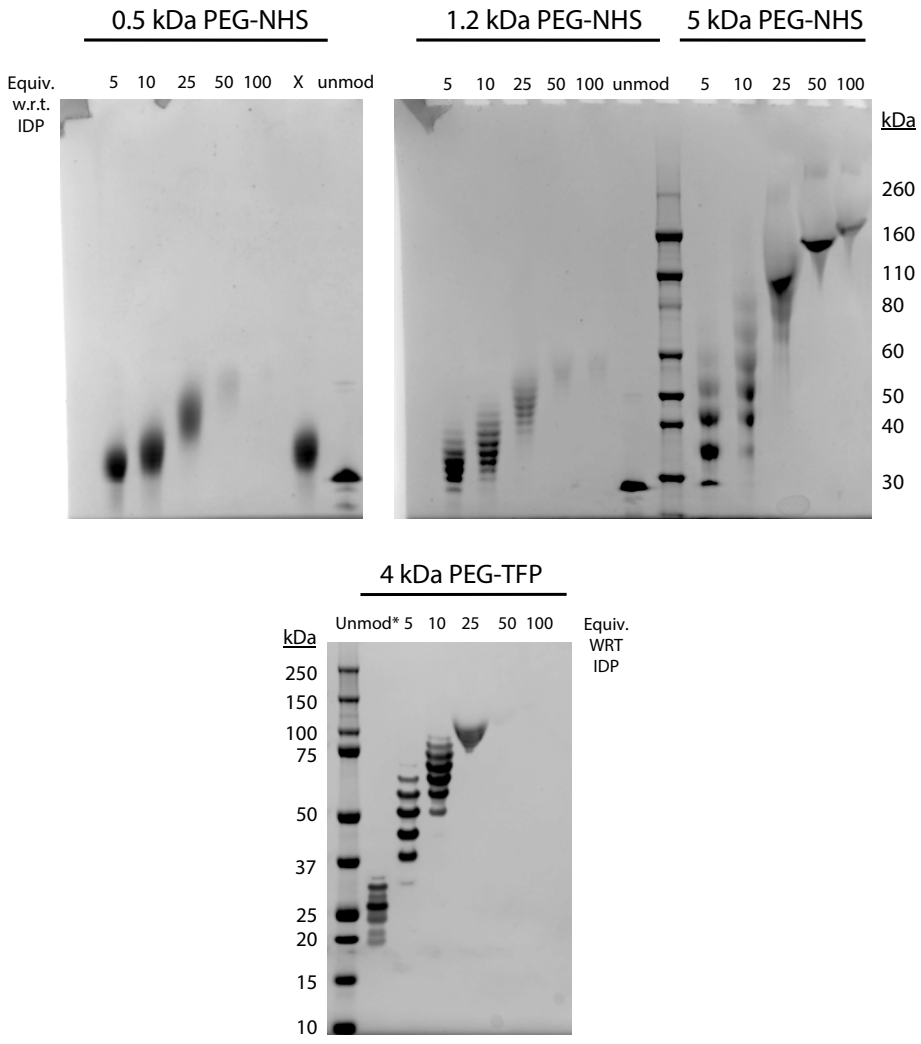


Figure A.7. Coomassie-stained SDS-PAGE of IDP samples modified with various equivalents of monodisperse 0.5 kDa PEG NHS-ester, 1.2 kDa PEG NHS-ester, 4 kDa branched PEG TFP-ester, and a polydisperse 5 kDa PEG NHS-ester. The IDP sample modified with 5 equivalents of 1.2 kDa PEG-NHS (top left gel) was loaded in the same lane as the ladder. The unmodified IDP used in bottom gel (unmod*) was an old sample, which showed significant degradation, but was not used for PEGylation reactions. The faint staining and warped shape of bands with high PEG equivalents are likely due to the presence of large amounts of conjugated PEG and free PEG that were not removed before gel analysis.

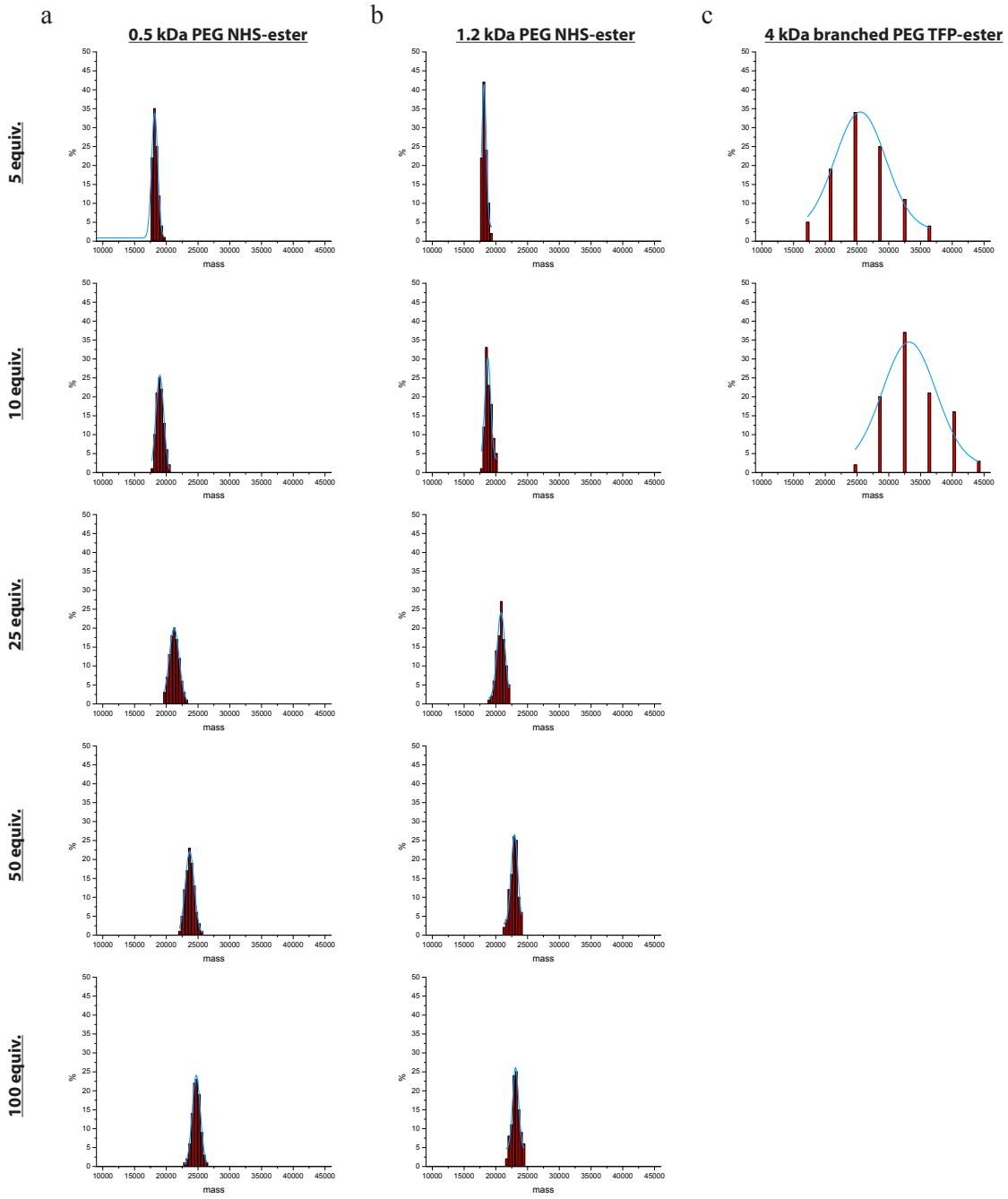


Figure A.8. Gaussian fit for the distribution of IDP samples conjugated with various equivalents of monodisperse 0.5 kDa PEG NHS-ester, 1.2 kDa PEG NHS-ester, and 4 kDa branched PEG TFP-ester with respect to IDP. Peaks from ESI-MS spectra (Figure A.5) were integrated and their areas were used to construct column plots of the distribution of PEG modifications. A Gaussian single peak fit (Gauss fit in Origin, blue curve) was applied to the plots using Origin 8.0 Software. The center and width of the peak were used in calculations of radius of gyration (*vide infra*) and for obtaining the average number of PEG modification for each set of reaction conditions (Table A.1).

0.5 kDa PEG	Avg # of conjugated PEG	Z_{eff}
5 equiv.	4.9	+0.7
10 equiv.	6.9	-1.3
25 equiv.	12.6	-7.0
50 equiv.	18.8	-13.2
100 equiv.	21.4	-15.8
1.2 kDa PEG	**Avg # of conjugated PEG**	**Z_{eff}**
5 equiv.	2.7	+2.9
10 equiv.	4.3	+1.3
25 equiv.	9.5	-3.9
50 equiv.	14.8	-9.2
100 equiv.	15.3	-9.7
4 kDa branched PEG	**Avg # of conjugated PEG**	**Z_{eff}**
5 equiv.	2.4	+3.2
10 equiv.	4.6	+1.0

Table A.1. The average number of conjugated PEG chains and the effective protein charge (Z_{eff}) of IDP conjugated with various equivalents of monodisperse 0.5 kDa PEG NHS-ester, 1.2 kDa PEG NHS-ester, and 4 kDa branched PEG TFP-ester. The average number of PEG chains added was obtained using a Gaussian single-peak fit (Figure A.6). The Z_{eff} was calculated by subtracting the average number of conjugated PEG chains from the theoretical protein charge at pH 7.5, which is 5.6.

A.3. Conclusions and Future Directions

We have successfully cloned, expressed and purified a 16 kDa IDP, and constructed a PEGylated IDP library with different sizes of PEG chains. Biophysical characterization of the library, including fluorescence anisotropy, fluorescence correlation spectroscopy (FCS), and Förster resonance energy transfer (FRET), will uncover the effect of size and density PEG chains on the conformation of IDP. This information will shed light on the potential of using IDPs as drug delivery agents.

A.4. Materials and Methods

A.4.1. General Procedures and Materials

Unless otherwise noted, all reagents and solvents used were of analytical grade and were used as received from commercial sources. For the determination of protein concentrations, UV-visible spectrometry was performed using a NanoDrop 1000 (Thermo Scientific). Protein mass spectrometer data were collected on an Agilent 6224 TOF LC/MS equipped with a Turbospray ion source and connected to an Agilent 1260 Infinity Series HPLC system and a Proswift RP-4H LC column (Thermo Fisher, cat# 069477). Because of the hydrophilicity of IDP, we used the following MeCN:H_2O gradient containing 0.1% formic acid.

Minutes	% MeCN
0	5
2	10
5.5	100
6	100

Protein mass reconstruction was performed on the charge ladder with Mass Hunter software (Agilent Technologies).

Sodium dodecyl sulfate-poly(acrylamide) gel electrophoresis (SDS-PAGE) analysis was performed using either 12% or 4 - 12% NuPAGE Bis-Tris gels (Thermo Fisher) with MOPS or MES buffer. SDS-PAGE loading buffer containing β-mercaptoethanol or dithiothreitol (DTT) was added to all electrophoresis samples (10-15 μL of 20 μM IDP), except for non-reducing gel samples. Coomassie Brilliant Blue R-250 (Bio-Rad) and Gel Doc EZ System (Bio-Rad) were used to stain and image SDS-PAGE gels.

A.4.2. Cloning

The inherent repetitive sequence of IDP made the synthesis of one contiguous gene block difficult. We therefore used an alternative strategy involving two gene blocks (gBlocks: IDT Technologies) with a 32 bp consensus sequence that allows for Gibson assembly (see Figure A.2a for the two gene block sequences). gBlocks (100 ng each) and 10 μl of 2x Gibson master mix[18] (a gift from the M. Chang lab) was adjusted with water to a volume of 20 μl and was incubated at 50 °C for 60 min. After DNA cleanup with QiaQuick (Qiagen), the assembly product was PCR amplified (VENT polymerase from NEB, T_m = 61 °C) with forward and reverse primers: 5'- ATA ATA GCT AGC TTA GTT CCT CGT GCC TGG CGT G -3' and 5'- TAT TAT CTC GAG CTA TTA GGC ACA CCA GTA CGG AGA TTT C -3'. The insert containing NheI and XhoI restriction sites was double digested. After the restriction enzymes were heat inactivated at 80 °C for 10 min, the double digested insert was ligated (QuickLigase, NEB) with the pSKB3 vector containing MBP. Plating on Kanamycin agar plates yielded individual colonies that were cultured, DNA purified (NucleoSpin, Macherey-Nagel) and sequenced (Quintara BioSciences). The complete sequence is located at the end of this appendix.

A.4.3. Protein Expression

Plasmids were transformed into *E. coli* BL21 (DE3) competent cells. Starter cultures (20 mL of LB, 50 mg/L Kanamycin) were grown from single colonies, grown overnight at 37 °C, and used to inoculate 1 L of TB media (50 mg/L Kanamycin). Cultures were grown to an optical density (OD) of ca. 0.5, cooled for 20 min at 25 °C, induced with 0.5 mM IPTG, and expressed overnight (ca. 18 h) at 25 °C. Cells were harvested by centrifugation for 15 min at 4,000 rcf at 4 °C.

A.4.4. Protein Purification

The cell pellet was transferred to a 50 mL Falcon tube in PBS buffer, and spun down for 10 min at 4,000 rcf. The resulting pellet (ca. 5 g) was lysed in 30 mL of buffer A (20 mM pH 7.5 HEPES, 300 mM NaCl, 10 mM imidazole) supplemented with one tablet of EDTA-free SigmaFast Protease Inhibitor (Sigma Aldrich), 2 mM PMSF, and 10 mg lysozyme. The resuspended sample was lysed with an Avestin C3 homogenizer followed by 20 min of centrifugation at 24,000 rcf at 4 °C. The supernatant was filtered through a 40 mm Steriflip filter (Millipore), and loaded onto a 5 mL Ni-NTA column (Protino, Machery Nagel) connected to an AKTA FPLC system (GE Healthcare Life Sciences) that was pre-equilibrated with buffer A. The column was washed with 50 mL (10 CV) of buffer A containing 10 mM β-mercaptoethanol. The protein was eluted with 20 mM pH 7.5 HEPES, 300 mM NaCl, 250 mM imidazole, 10 mM β-mercaptoethanol. Imidazole was removed by exchanging against 20 mM HEPES (pH=7.5), 100 mM NaCl with a 10DG desalting column (BioRad), and subsequently digested with 1 mg of thrombin protease (high purity from Bovine, MP Biomedicals). Complete digestion was achieved at room temperature after 1 h as confirmed by LC/MS. The protein mixture was diluted with salt-free 20 mM pH 7.5 HEPES buffer to 50

mL ([NaCl] = ca. 5 mM) before loading it onto a 1 mL HiTrap SP HP cation exchange column connected to an FPLC. The column was pre-equilibrated with 20 mM pH 7.5 HEPES containing 10 mM β-mercaptoethanol, washed with 10 mL (10 CV) of the same buffer and eluted with a gradient from 0-1 M NaCl (50 mL total volume). The final IDP sample was obtained with a final desalting column (10DG, BioRad) to obtain the protein at a final concentration of 270 μM in 20 mM pH 7.5 HEPES, 50 mM NaCl. The protein was >95% pure by SDS-PAGE and LCMS. The purified protein was flash frozen with liquid N_2 in 20 μl aliquots and stored at -80 °C until use. We observed dimer formation upon storage at room temperature, but it can be reversed using reducing agents such as β-mercaptoethanol, Tris(2-carboxyethyl)phosphine (TCEP) or DTT.

A.4.5. Protein Modification

Procedure for cysteine alkylation with Alexa Fluor 488-maleimide. An 800 μL solution of IDP (270 μM) in 20 mM pH 7.5 HEPES, 50 mM NaCl, was incubated with 4 equiv. of Alexa Fluor 488-maleimide (AF488-mal, Thermo Fisher A10254) in the presence of 5.6 mM TCEP (0.5 M, Sigma) for 30 min at room temperature. Excess dye was removed using a desalting column (10DG, BioRad) and 20 mM pH 7.5 HEPES, 50 mM NaCl.

Procedure for PEGylation. Monodisperse 0.5 kDa PEG NHS-ester (cat# 10260), 1.2 kDa PEG NHS-ester (cat# 10304) and 4 kDa branched PEG TFP-ester (cat# 10458) were purchased from Quanta BioDesign. Polydisperse 5 kDa PEG-NHS ester (Fmoc-PEG-SVA, MW 5000) was purchased from Laysan Bio. NHS- or TFP-ester reagents (20 mM stock solutions in H_2O) were made fresh immediately before PEGylation reactions. PEG-NHS- or TFP- ester reagent (5-100 equiv.) was added to a 50 μL solution of AF488 labeled-IDP (152 μM) in 20 mM pH 7.5 HEPES, 50 mM NaCl. This buffer was also added to each reaction to reach a final volume of 150 μL. The reactions were incubated at room temperature in the dark for 12 – 18 h before purification and buffer exchange using spin concentrators with a 10 kDa molecular weight cut-off (Amicon Ultra, EMD Millipore). The resulting PEGylated IDP was analyzed by SDS-PAGE (all samples) and LC/MS (only samples with monodisperse PEG).

A.5. References

1. Yoo, J. W., Irvine, D. J., Discher, D. E., and Mitragotri, S. (2011) Bio-inspired, bioengineered and biomimetic drug delivery carriers, *Nat. Rev. Drug Discov. 10*, 521-535.

2. Elzoghby, A. O., Samy, W. M., and Elgindy, N. A. (2012) Protein-based nanocarriers as promising drug and gene delivery systems, *J. Control. Release 161*, 38-49.

3. Vazquez, E., and Villaverde, A. (2010) Engineering building blocks for self-assembling protein nanoparticles, *Microb. Cell Fact. 9*, 101.

4. Srinivasan, N., and Kumar, S. (2012) Ordered and disordered proteins as nanomaterial building blocks, *WIRES Nanomed Nanobi 4*, 204-218.

5. Gaumet, M., Vargas, A., Gurny, R., and Delie, F. (2008) Nanoparticles for drug delivery: The need for precision in reporting particle size parameters, *Eur. J. Pharm. Biopharm. 69*, 1-9.

6. De Jong, W. H., and Borm, P. J. A. (2008) Drug delivery and nanoparticles: applications and hazards, *Int. J. Nanomed. 3*, 133-149.

7. Geng, Y., Dalhaimer, P., Cai, S. S., Tsai, R., Tewari, M., Minko, T., and Discher, D. E. (2007) Shape effects of filaments versus spherical particles in flow and drug delivery, *Nat. Nanotechnol. 2*, 249-255.

8. Oltra, N. S., Swift, J., Mahmud, A., Rajagopal, K., Loverde, S. M., and Discher, D. E. (2013) Filomicelles in nanomedicine - from flexible, fragmentable, and ligand-targetable drug carrier designs to combination therapy for brain tumors, *J. Mater. Chem. B 1*, 5177-5185.

9. Bazile, D., Prudhomme, C., Bassoullet, M. T., Marlard, M., Spenlehauer, G., and Veillard, M. (1995) Stealth Me.PEG-PLA nanoparticles avoid uptake by the mononuclear phagocytes system, *J. Pharm. Sci 84*, 493-498.

10. Soppimath, K. S., Aminabhavi, T. M., Kulkarni, A. R., and Rudzinski, W. E. (2001) Biodegradable polymeric nanoparticles as drug delivery devices, *J. Control. Release 70*, 1-20.

11. Perrault, S. D., Walkey, C., Jennings, T., Fischer, H. C., and Chan, W. C. W. (2009) Mediating tumor targeting efficiency of nanoparticles through design, *Nano Lett. 9*, 1909-1915.

12. Gref, R., Domb, A., Quellec, P., Blunk, T., Muller, R. H., Verbavatz, J. M., and Langer, R. (2012) The controlled intravenous delivery of drugs using PEG-coated sterically stabilized nanospheres, *Adv. Drug Deliver. Rev. 64*, 316-326.

13. Hwang, P. M., Pan, J. S., and Sykes, B. D. (2012) A PagP fusion protein system for the expression of intrinsically disordered proteins in *Escherichia coli*, *Protein Express. Purif. 85*, 148-151.

14. Diguan, C., Li, P., Riggs, P. D., and Inouye, H. (1988) Vectors That Facilitate the Expression and Purification of Foreign Peptides in *Escherichia coli* by Fusion to Maltose-Binding Protein, *Gene 67*, 21-30.

15. Terpe, K. (2003) Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems, *Appl. Microbiol. Biotechnol. 60*, 523-533.

16. Pryor, K. D., and Leiting, B. (1997) High-level expression of soluble protein in Escherichia coli using a His(6)-tag and maltose-binding-protein double-affinity fusion system, *Protein Express. Purif. 10*, 309-319.

17. Gallwitz, M., Enoksson, M., Thorpe, M., and Hellman, L. (2012) The extended cleavage specificity of human thrombin, *PLOS ONE 7*.

18. Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases, *Nat. Methods 6*, 343-U341.

A.6. Open Reading Frame Sequence of MBP-IDP Encoding pSKB3 Vector

CC**ATG**GCCAGCAGCCATCATCATCATCATCACGATTACGATATCCCAACGAC-
CGAAAACCTTTACTTCCAGGGATCCGAAAACCTTTACTTCCAGGGATCCAAAATCGAA-
GAAGGTAAACTGGTAATCTGGATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTC-
GGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCACCGTTGAGCATCCGGATAAACT-
GGAAGAGAAATTCCCACAGGTTGCGGCAACTGGCGATGGCCCTGACATTATCTTCTGG-
GCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCCGG-
ACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATGCCGTACGTTACAACGG-
CAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCT-
GCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAG-
CGAAAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCT-
GATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGCAAGTACGACATTAAAGAC-
GTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTGACCTGATTA-
AAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTTTAATAAAGG-
CGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACACCAGCAAAGT-
GAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCATCCAAACCGTTCGTTGGC-
GTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAAAGAGTTCCTC-
GAAAACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAACCGCTGGGT-
GCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGTGAAAGATCCACGTGTTGCCGCCAC-
TATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCGCTTTCT-
GGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCGAT-
GAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCGAACAACAACAACAATAACAATAA-
CAACAACCTCGGGGCTAGCTTAGTTCCTCGTGCCTGGCGTGGCTCCCCGTGGGCAGAG-
GCCAAGAGTCCAGCGGAAGCTAAGTCGCCAGCCGAAGTCAAGTCGCCCGCCGTCGCGA-
AAAGCCCCGCAGAGGTGAAATCCCCGGCCGAAGTCAAATCGCCGGCAGAAGCGAAATC-
CCCGGCAGAAGCAAAAAGTCCTGCTGAGGTCAAATCGCCAGCAACCGTCAAATCCCCT-
GGAGAGGCAAAATCTCCGGCAGAAGCCAAGTCCCCTGCCGAAGTGAAGTCACCTGTC-
GAAGCCAAGTCGCCGGCCGAAGCGAAGAGCCCAGCGAGCGTGAAAAGTCCTGGTGAG-
GCTAAGTCCCCGGCGGAAGCGAAATCTCCAGCGGAAGTAAAGAGTCCGGCCACCGTTA-
AATCCCCGGTAGAGGCCAAAAGCCCTGCGGAAGTTAAATCGCCGGTGACGGTCAAATCAC-
CCGCGGAAGCGAAGTCCCCGGTGGAGGTGAAATCTCCGTACTGGTGTGCC**TAA**TAGCTC-
GAG

Underlined: restriction sites in order of sequence NcoI/BamHI/NheI/XhoI

Bold: Start and Stop codons

Green: 6xHis affinity tag encoding sequence

Grey: TEV encoding sequence

Maltose Binding Protein encoding sequence

N_{10} linker encoding sequence

Thrombin encoding sequence

IDP encoding sequence

Cysteine encoding sequence