Title
Design, Engineering, and Characterization of a Photoswitchable Cadherin

Author
Ritterson, Ryan Scott

Publication Date
2013

Peer reviewed|Thesis/dissertation

Design, Engineering, and Characterization of a Photoswitchable Cadherin

by

Ryan S. Ritterson

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

[redacted]

Dedication and Acknowledgements

Although Ph.D. theses are written as though they were the work of a single individual, in reality this is very rarely the case, and this thesis is no exception. Without the help of many individuals, directly or indirectly, with contributions large or small, this work would not have been possible.

I am indebted to my advisor, Prof. Tanja Kortemme. She was willing to take an enormous leap and trust a complicated, difficult experimental project to a student with virtually no experimental lab experience. Further yet, she supported me and my work through the years that it took, through times both difficult and good. I am grateful to her for the enthusiastic scientific discussions, the continuous words of encouragement, and for providing a place for me to find a professional home.

I am also very grateful for the help of my thesis committee, Profs. David Agard, Joe DeRisi, and Keith Mostov. On many an occasion, their guidance was instrumental in the direction of my Ph.D. and my post-graduate plans.

Scientifically, many others contributed to my work. I am grateful to (in loose chronological order) Dr. Matthew Banghart, Prof. Dirk Trauner, Prof. Andrew Wooley, Michael Michalik, Alec Nielsen, Flora Yi, Prof. Nicolas Borghi, Prof. James Nelson, Kris Kuchenbecker, and Justin Farlow. For at least one point in this project, each one of them contributed their time and scientific resources to the project, and some part of the credit for this work is owed to them. When appropriate, I have credited them directly at the beginning of the chapters where they contributed.

I appreciated the help of all of the members of the Kortemme Lab and their insights over the many years this project took. I would also like to specifically acknowledge Dr. Greg Kapp and Cristina Melero for patiently teaching me the basics of experimental biology even when it meant my mistakes interrupted their own work. I would also like to thank the two of them plus Debbie Jeon and Dr. Brianna Burden for making the daily work in the lab fun and happy.

Finally, I am deeply indebted to my loving partner Elizabeth. Without her love and ceaseless support, I would not and could not have been successful in my pursuit of a Ph.D. She was there to pick me up when times were tough, celebrate with me when things were well, help me figure things out, and never feared to argue or debate some esoteric scientific detail even when weary.

Abstract

There is a growing interest in engineering proteins whose function can be controlled with the spatial and temporal precision of light. Here, I present a novel example of a functional light-triggered switch in the calcium-dependent cell-cell adhesion protein E-cadherin, created using a mechanism-based design strategy. I report an 18-fold change in apparent calcium binding affinity upon illumination. This work includes a detailed description of the design methodology used, as well as a detailed examination of functional switching via linked changes in Ca^{2+} binding and cadherin dimerization. This design opens avenues towards controllable tools that could be applied to many long-standing questions about cadherin's biological function in cell–cell adhesion and downstream signaling. It also presents a potential generalizable strategy for creating additional photoswitchable proteins.

Table of Contents

List of Tables

List of Equations

List of Figures

Chapter 1: Introduction

There has been considerable interest in light-based control of biological systems,[1] and successful applications include light-modulation of neuronal ion channels[2], light-switchable signaling proteins[3-5] and light-controlled protein targeting[6]. Light-based methods offer titratable, precise spatial, and temporal regulation that has been demonstrated *in vitro*[7], in cell culture[6,8], and in whole animals.[9] Most examples of light-based control fall into one of two categories: (a) those that are genetically encoded using a recombinantly-produced protein borrowed from nature,[6] and (b) those created *via* targeted insertion of amino acids into a protein sequence and subsequent reaction with them of an exogenously introduced photoisomerizable small molecule, typically azobenzene based.[10] Azobenzene and related molecules undergo a reversible *cis–trans* isomerization when exposed to specific wavelengths of light, and this change in molecular shape can be coupled to changes in protein function. While in (a) the functional design is already provided naturally, one is both limited by the function (e.g., modulation of protein–protein binding, tuning fluorescence intensity) already encoded by the natural gene, and by the requirement to fuse the natural protein with the protein to be modulated. In contrast, the designs in (b) allow many types of functional modulation, such as changes in agonist binding, protein–protein binding, and protein folding, limited only by the cleverness of the designer, though they must be designed and engineered for each new target. In this work, I used an azobenezene-based (type b) model to design and engineer a new photoswitchable protein: the

cell-cell adhesion protein E-cadherin. I used a new strategy where isomerization produces changes in protein–ion affinity, in this case calcium, which then couple to changes protein dimerization (**Figures 1 & 2**).

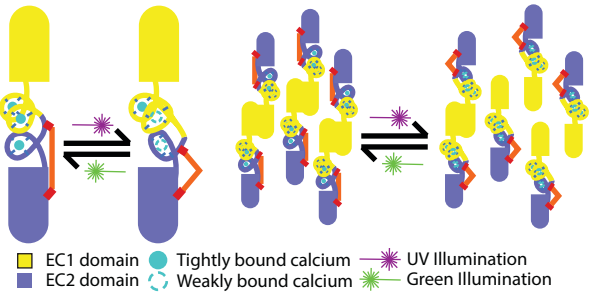

Figure 1 – A cartoon showing the basis of the design. As designed, the photoswitch reduces calcium binding affinity, which, in turn, reduces cadherin homodimer affinity.

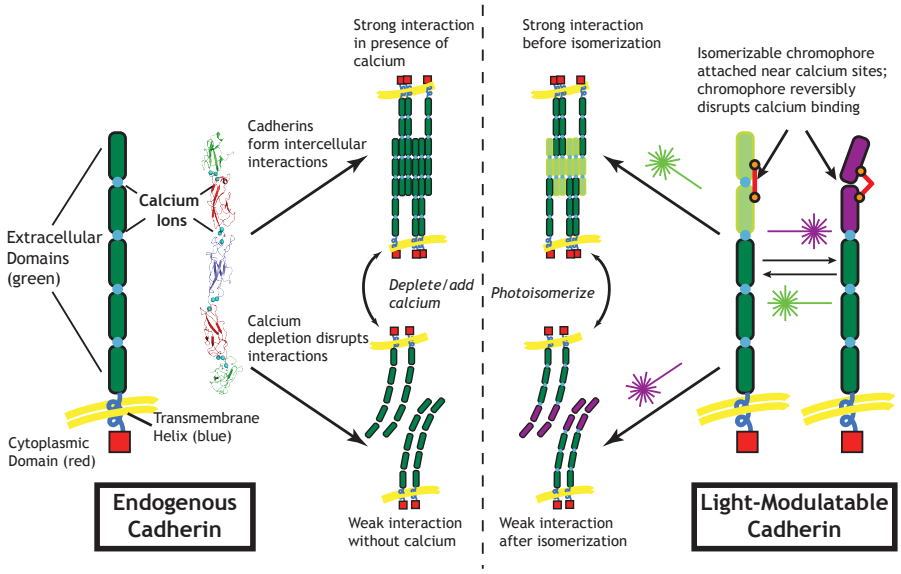

Figure 2 – Cartoon showing a comparison of endogenous cadherin (left) and the light modulatable cadherin (right). The designed photoswitchable cadherin uses the calcium-dependence of endogenous cadherin to accomplish its effect.

Cadherins are a key family of calcium-dependent cell-cell adhesion proteins, and are divided into several subtypes, including the most commonly studied subtype, the classical cadherins. Classical cadherins, which include E-, N-, P-, R-, and C-cadherin[11], are composed of an intracellular domain, a single transmembrane

helix, and five, repeated, immunoglobulin-like extracellular domains labeled EC1 (N-terminal, membrane distal) to EC5 (C-terminal, membrane proximal), along with three calcium binding sites present in the loops at each extracellular domain boundary (**Figure 2-left**).[12,13] Calcium binding is required for cadherin function, as depletion of calcium disrupts cadherin-mediated cell adhesion;[14] the presence of calcium is suggested to rigidify the cadherin structure, allowing it to multimerize.[15] Depletion of cadherin significantly slows cell-cell adhesion[16] and E-cadherin knockout mice do not develop past the 32-cell state[17]. Additionally, in a classic experiment, cadherin-free, non-adherent cells transfected with cadherin acquire morphological similarities to naturally adherent cells, indicating the critical role cadherin plays in controlling and enabling cell-cell adhesion.[18]

Cadherins form multimers in cellular junctions and play critical roles in tissue morphogenesis during development, and tumor invasiveness in cancer. Open questions about cadherin function include: (i) whether cadherin directly participates in outside-in transmembrane signaling; (ii) whether spatially localized changes in cadherin adhesiveness induce spatially constrained or globally transduced signals; and (iii) whether, in remodeling of cell-cell adhesion occurring in embryogenesis or wound healing, changes in cadherin adhesiveness act early as a primary signal transducer or late as a response to other, earlier, signals. My objective in developing a light-controlled cadherin is to enable the study of these questions. A key advantage of my approach over existing methods to modulate cell adhesion, such as calcium depletion (**Figure 2-left**), is the high spatial and temporal resolution achievable with light. This approach, if integrated with cultured cells

would thus enable one to derive information on cell adhesion and its spatiotemporal integration with signaling and developmental processes currently unobtainable. In addition to advancing fundamental understanding of these processes, light control would provide a mechanism to precisely control adhesion in engineering applications.

My approach to creating a light-switchable cadherin aimed to modulate its calcium-binding affinity. Because calcium binding is essential for cadherin multimerization, I reasoned that reversibly changing calcium binding affinity would be an effective way to also modulate cadherin adhesive function (**Figures 1 & 2**). I engineered cysteine residues into the protein to serve as conjugation sites for an azobenzene-based photoisomerizable chromophore, BSBCA (**Figure 3**). BSBCA has been used in previous applications,[7,19-21] demonstrating reversible switching between the *cis* and *trans* states when exposed to 370 nm (near UV) and 550 nm (green) light, respectively.[7,19] My strategy involved conjugating both ends of the chromophore to the calcium-binding loops between cadherin domains EC1 & EC2, as these calcium sites have previously been shown to be most critical for function.[13]

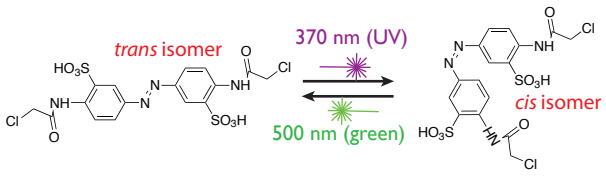

Figure 3 – The photoisomerizable molecule BSBCA.

In addition, because the calcium binding sites are located in loop regions, and bind calcium with relatively weak affinities near 20 μM[22], I reasoned it would be comparatively easier to induce conformational changes affecting calcium binding

there than in more rigid secondary structural elements or well-packed core regions of the protein.

Because BSBCA spontaneously crosslinks cysteine residues,[23] the design challenge can be generalized as the problem of finding the best pair of residues to mutate to cysteine. In practice, however, an enormous number of pairs are possible, the overwhelming majority of which are likely to be non-functional. I took a sequential and computational approach to identifying likely functional pairs that resulted in a relatively small library of only 11 constructs to test, one of which I showed to be functional. This compares favorably to a high-throughput screen of a large library of double mutants, both in cost and in time expended.

This work presents the development of a photoswitchable cadherin from the beginning of the design process through the *in vitro* characterization of the functional photoswitchable molecule. This *in vitro* characterization focused on the first two extracellular domains of E-cadherin (EC12) because they contain the homodimeric binding interfaces[24] and they are the specificity determining domains,[25] making them most principally responsible for cadherin's function. Additionally, the shortened EC12 construct can be readily produced in high yields in *E.coli*. EC12 contains a single native cysteine residue, which I mutated to alanine (C9A) in order to avoid undesired chromophore conjugation to that residue; this mutation has been previously shown not to affect cadherin function.[26]

The chapters of this work are in loose chronological order of experiments. Chapter 2 presents initial control experiments done early in the project, showing that treating cells with azobenzene-based molecules is feasible. Chapters 3 and 4

cover the computational process used to create the library and the basic measurements used to identify mutant 129/138 as the most promising candidate. Chapter 5 covers the optimization of the biochemistry involved in purifying conjugated 129/138, necessary to produce the quantities of protein used in the assays to characterize the mutant in detail. Chapters 6 through 8 present the assays used to characterize 129/138. Although this work could have been presented in fewer, longer chapters, I have chosen to break it up into smaller pieces so that the interested reader can more quickly find the topic of interest, particularly Chapters 3 and 4, which may be of use to others seeking to make light-modulatable proteins.

Chapter 2: Preliminary Control Experiments

Introduction

At this point in the project, the intended use of the photoswitchable cadherin was to look at cell genomic expression profiles pre- and post- illumination, in order to investigate the signaling cascade triggered by cadherin adhesion or disruption. I had not yet been made aware of the coated coverslip assays (**see Chapters 1 & 9**), and so assumed that I would have to introduce the engineered cadherin directly into cells. As such, I believed I would have to use the full-length cadherin molecule and not just the EC12 subsection, and that I'd have to use a different modified azobenzene, bis-MTS-sulf-azo **(Figure 4)**, which contains a different reactive end that is much more reactive than BSBCA, but that hydrolyzes in water. It also forms a disulfide bond with the protein, instead of the much more stable C-S bond formed with BSBCA, necessitating extra care in handling to avoid reducing the chromophore-protein bond in the presence of, e.g., DTT or TCEP.

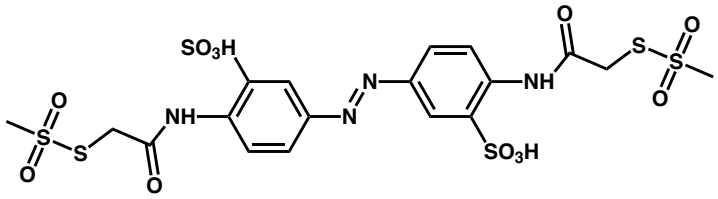

Figure 4 – The structure of bis-MTS-sulf-azo

Prior to designing the light-modulatable cadherin molecule, I wanted to do two control experiments to ensure that it would be functional when used with cells. First, I wanted to ensure that the small molecule would not disturb the two disfulfide bonds in cadherin's EC5 domain and react with them. These disulfides

have been shown to be important for cadherin function, so I needed to avoid disturbing them. Second, I wanted to know that the small molecule would react specifically with cadherin and not other proteins in the cell, to avoid triggering genomic changes unrelated to cadherin. To answer the first question, I used a series of protein gels, as well as Ellman's test[27]; to answer the second, I treated cells with varying concentrations of the small molecule and looked at mRNA expression levels at two different times post-treatment. The results showed that the small molecule did not interfere with either, providing support for moving forward with the design strategy.

Methods

EC5 Purification

The EC5 domain was PCR subcloned from a plasmid containing the full-length mouse E-Cadherin sequence using 5' oligonucleotide CACCATGAATGACAACGCTCCCATCCCAGAACCTCGAAA and 3' oligonucleotide AACTTGCAATCCTGCTGCCACGATTCCCGCCT. The resulting PCR product was inserted into vector pENTR/D-TOPO (Life Technologies K2400-20) using the manufacturer's provided protocol, after which the insert sequence was verified via DNA sequencing. Next, the insert was cloned into vector pDEST-17 (Life Technologies #11803-012) using Gateway technology (Life Technologies #11791-019). pDEST-17 contains a C-terminal 6xHis tag, necessitating the use of an insert without a stop codon.

The EC5-containing pDEST-17 vector was transformed into BL21 (DE3) cells and grown in LB culture medium (BD #244610) containing 100 μg/mL carbenicillin.

Cultures were grown until OD = 0.6, at which protein expression was induced with 500 μM IPTG (Research Products International #156000). Cultures were grown for 3-4 hours post-induction, then spun at 6000x*g* for 15 minutes to pellet cells, and pellets were stored at -80° C until purification.

EC5-containing lysates were lysed, cleared, and purified using a protocol described previously[28], resulting in ~10 mg of protein per liter of culture volume. After purification, protein purity and monomeric state were verified using a HiLoad 16/600 Superdex 75 column (GE #28-9893-33) attached to an FPLC. Protein concentrations were measured using A_{280} as measured with a Cary 50 Bio UV/Vis spectrophotometer (Varian), with $\varepsilon_{280} = 13200$ as predicted by the ExPASY online protein parameter tool.[29]

EC5 Disulfide Reactivity

To verify disulfide formation using Ellman's test[27], 200 μL, 163 μM EC5 in TBS (25 mM Tris, pH 7.5, 150 mM NaCl) was added to 800 μL of 1x Ellman's reagent buffer (20 mM $Na_xH_yPO_4$, pH 7.5, 150 mM NaCl, 150 μM DTNB). In addition, 200 μL of TBS buffer containing no protein was added to 800 μL of 1x Ellman's reagent buffer to act as control. After waiting 5 minutes, the absorbance at 412 nm was measured for each sample, also using a Cary 50 Bio spectrophotometer. To verify DNTB freshness, 40 μL of 1 mM dithiothreitol (DTT) was added to 800 μL of 1x Ellman's reagent buffer.

To verify that the protein remained monomeric after chromophore addition, bis-MTS-sulf-azo was added to a final concentration of 50 μM to 200 μL, 163 μM EC5 in TBS (25 mM Tris, pH 7.5, 150 mM NaCl) and the solution was left to react for one

hour. After reacting, protein was loaded into a HiLoad 16/60 Superdex 75 column attached to an FPLC.

Cell Exposure Experiments & cDNA Construction

L cells[30] (ATCC #CCL-1) growing at 37° C in D-MEM[31] High Glucose (UCSF Cell Culture Facility #CCFAA005) containing 10% v/v fetal bovine serum (JR Scientific #43603-100) were exposed to bis-MTS-sulf-azo at concentrations of 100 nM, 1 μM and 10 μM. Cells were harvested 20 minutes and 2 hours after exposure. Plates of cells unexposed to chromophore were used as background controls. Cells were harvested using a 1% solution containing 5 mM EDTA (UCSF Cell Culture Facility #CCFGP003). Total RNA was purified from cells using an RNeasy Mini kit (Qiagen #74104) using the manufacturer's instructions. Messenger RNAs were purified from total RNA using an Oligotex Mini kit (Qiagen #70022) according to the manufacturer's instructions. After purification, mRNAs were converted into a labeled cDNA library using a SuperScript Plus Direct cDNA Labeling System kit (Life Technologies #L1015-06) following the manufacturer's instructions.

Microarrays

Experiments used Mouse Exonic Evidence Based Oligonucleotide (MEEBO) arrays[32] provided as a gift from the Keith Yamamoto lab. Microarrays were post-processed using protocols available from the Joe DeRisi lab[33]. Microarrays were hybridized using MAUI mixer/hybridizer chambers (BioMicro Systems) following a protocol provided by the Keith Yamamoto lab[34]. A single microarray was used for each concentration/time pair (6 arrays total). After hybridization, microarrays were scanned using a GenePix 4000 Scanner (Molecular Devices).

After scanning, data were normalized such that the mean ratio of medians between channels of all spots was 1.0. Afterward, spots with raw signals < 300 were removed, and remaining genes were clustered using average linkage hierarchical clustering for genes present on 75% or more of the arrays and channels with $|max\ signal| > 1.2$. After control and viral spots, as well as any spots that showed an expression pattern uncorrelated with concentration (i.e. likely statistical noise) were removed from the analysis, approximately 300 genes that showed significant expression changes were left to analyze in detail.

Results

EC5 Disulfide Analysis

The Ellman's tests showed that virtually no free cysteines were available in purified EC5. The increase in A_{412} of samples containing protein *versus* buffer was .0016 AU, close to the noise threshold of the instrument and statistically insignificant. This indicates that purified protein came with disulfides preformed. Insight provided to us by Dirk Trauner[35] indicated that bis-MTS-sulf-azo was unable to undergo exchange with protein disulfide bonds, and thus the chromophore would be unlikely to interfere with the disulfides present in EC5.

Protein exposed to chromophore aggregated after exposure, as measured by FPLC (**Figure 5**), though a significant portion stayed monomeric. To determine whether the chromophore was crosslinking proteins into higher order species or whether the protein was forming non-covalent aggregates, I ran chromophore-exposed EC5 in a non-reducing SDS-PAGE gel. This gel, due to the presence of SDS, should break apart any non-covalent protein aggregates, but not reduce the

chromophore-protein or protein-protein disulfide bonds. All protein appeared monomeric in the gel, further confirming that the chromophore is not crosslinking proteins together (data not shown).

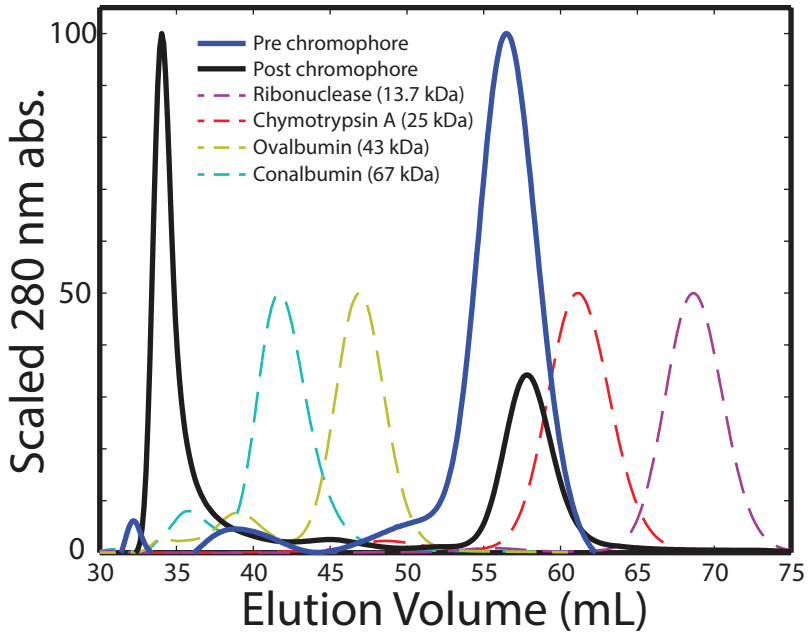

Figure 5 – Size exclusion chromatography of EC5 pre- and post- treatment with bis-MTS-sulf-azo. Also shown are calibration standards for comparison. EC5 peaks have been scaled such that the maximum peak height is 100. Size standards were scaled to a maximum height of 50 for visual convenience.

On multiple occasions, I observed that the chromophore appears to aggregate EC5 monomers. I am unsure whether the conditions in which I did the reaction were incompatible with the protein, or whether the chromophore itself was inducing aggregation. As chromophore was added directly from a highly-concentrated stock, the protein buffer only changed slightly. This observation, combined with the observation that BSBCA induces EC12 aggregation when exposed to certain agarose matrices, indicates that the chromophore is likely directly inducing aggregation. However, given that the protein concentrations used here

were many orders of magnitude higher than those likely to be present on the cell surface, it may not be an issue in *in vivo* experiments.

Between the lack of available disulfides to react, as well as the lack of protein crosslinking, I was satisfied that the chromophore was unlikely to affect the EC5 disulfides. However, in retrospect, these conclusions come with multiple caveats. First, while it is extremely unlikely given the known chemistry, it is possible, strictly speaking, that BSBCA could react with fully formed disulfide bonds where bis-MTS-sulf-azo cannot. BSBCA would be a poor choice to use to label cells given its very slow reactivity, but if it were to be used, it would require new testing. Second, a lack of multimerization or disulfide availability doesn't exclude the possibility that the chromophore is intramolecularly crosslinking a single monomer, as was the intent and was accomplished with BSBCA. One additional test that should have been run was exposure of EC5 to chromophore, followed by the removal of excess chromophore *via* desalting, then measurements of A_{370}. Because the protein has very little absorbance at 370 nm but the chromophore absorbs extensively at this wavelength, measurements of 370 nm absorbance after exposure would indicate whether any chromophore remains attached to the protein. If an increase in A_{370} was observed, I could have removed the chromophore via DTT, desalted the protein a second time, then remeasured the absorbance. If chromophore were specifically attached, I'd expect it to be removed following DTT exposure, resulting in a decrease in absorbance. These experiments did not come to mind at the time, owing to my relative inexperience in the lab. I did not go back and investigate EC5 disulfides in more detail, because the focus of the project turned to EC12 and *in vitro* studies, but

were I to attempt to label cells, I would reexamine these results and perhaps do further experiments to be more sure.

Microarray Experiments

The vast majority of genes on the arrays showed virtually no difference in expression after exposure compared to those unexposed (**Figure 6**). In the cluster of approximately 300 genes that remained after the filtering, none showed an obvious correlation to any cell process. Instead, they appeared to be random collections of genes from random processes, very few of which would be expected if the chromophore were activating cell signaling or toxicity pathways. The limited number of genes and the lack of correlation between them strongly suggest that the effect on cells of mere exposure to the chromophore is minimal. This conclusion is backed by the widespread use of this chromophore in living systems, ranging from cells[8] to whole organisms[9], without any reported adverse effects.

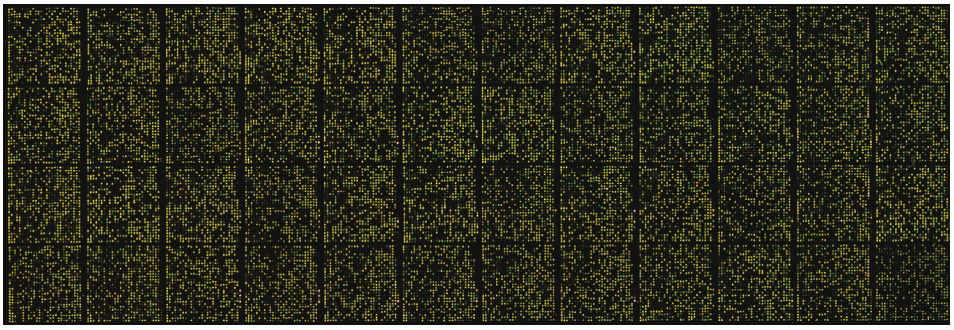

Figure 6 – Image of a microarray used to determine the effect of bis-MTS-sulf-azo on cells. This microarray represents 2 hours after treatment with 1 μM chromophore. The extensive number of yellow spots indicates very little changes between untreated and treated cells.

Chapter 3: Double Mutant Site Selection & Library Creation

Introduction

Because BSBCA spontaneously crosslinks cysteine residues,[23] the design challenge presented in this work can be generalized as the problem of finding the best pair of residues to mutate to cysteine. In practice, however, an enormous number of pairs are possible, the overwhelming majority of which are likely to be non-functional. I took a sequential and computational approach to identifying likely functional pairs (**Figure 7**).

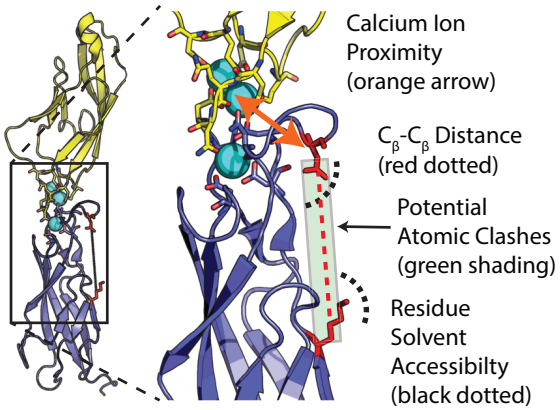

Figure 7 – EC12 structure showing the region targeted for photoswitchability. Labels indicate factors considered during the design process. EC1, gold; EC2, blue; calcium ions, cyan spheres.

Methods & Results

My design process involved first finding residues likely to be mutatable using the computational design program Rosetta.[36] I used a previously described method[37,38] to predict the change in fold stability of mutating, one by one, every residue in four representative E-cadherin structures (PDB identifiers 1FF5, 1EDH, 2O72 and 1Q1P) to alanine, the simplest mutation. Residues that had a mean predicted change in fold stability > 1 kcal/mol upon alanine mutation were removed

from consideration. I also removed all residues that directly bound calcium from consideration, reasoning that mutating them would likely cause a significant change in cadherin's calcium affinity. Next, I calculated all pair-wise C_β-C_β distances between the residues remaining using the 1FF5 structure, and kept those pairs that fell in the range 17-20 Å (appropriate for the BSBCA *trans* isomer). These first two steps reduced the potential double mutants from >20000 pairs to 1513.

Following these initial steps, I further restricted the number of pairs via a series of additional structural criteria (**Figure 7**). First, I restricted the remaining pairs to those that had at least one endpoint within 20 Å of a calcium ion as measured by the molecular graphics program PyMol[39], reasoning that pairs with both ends distant from the calcium binding sites would be unlikely to have an effect on calcium binding. This further reduced the number of pairs to 1120. Next, to only use residues that were accessible for chromophore conjugation, I restricted pairs to those where both endpoints were in residues that had >30 $Å^2$ of solvent accessible surface area (SASA) when the protein was in the monomeric state, again measured using PyMol. This reduced the set of pairs to 272. Finally, I plotted the remaining pairs on the 1FF5 PDB structure and eliminated those where the addition of the chromophore would sterically interfere or clash with the native protein structure, as estimated by using the surface representation of the protein structure and eliminating those pairs whose path of shortest distance penetrated the surface. This step eliminated a large part of the remaining set, leaving 28 pairs of conjugation sites. As a final step, I manually examined the 28 pairs and chose a set of 10, seeking to build a diverse set of conjugation sites to maximize the chances of finding a

functional photoswitch. One mutant pair in the library, 5/137, did not satisfy the structural criteria; it was chosen manually in an attempt to construct a photoswitch coupled to β-strand one, based on the previously observed functional importance of the strand to cadherin dimerization.[24,40] A flow chart summarizing the approach is presented in **Figure 8**.

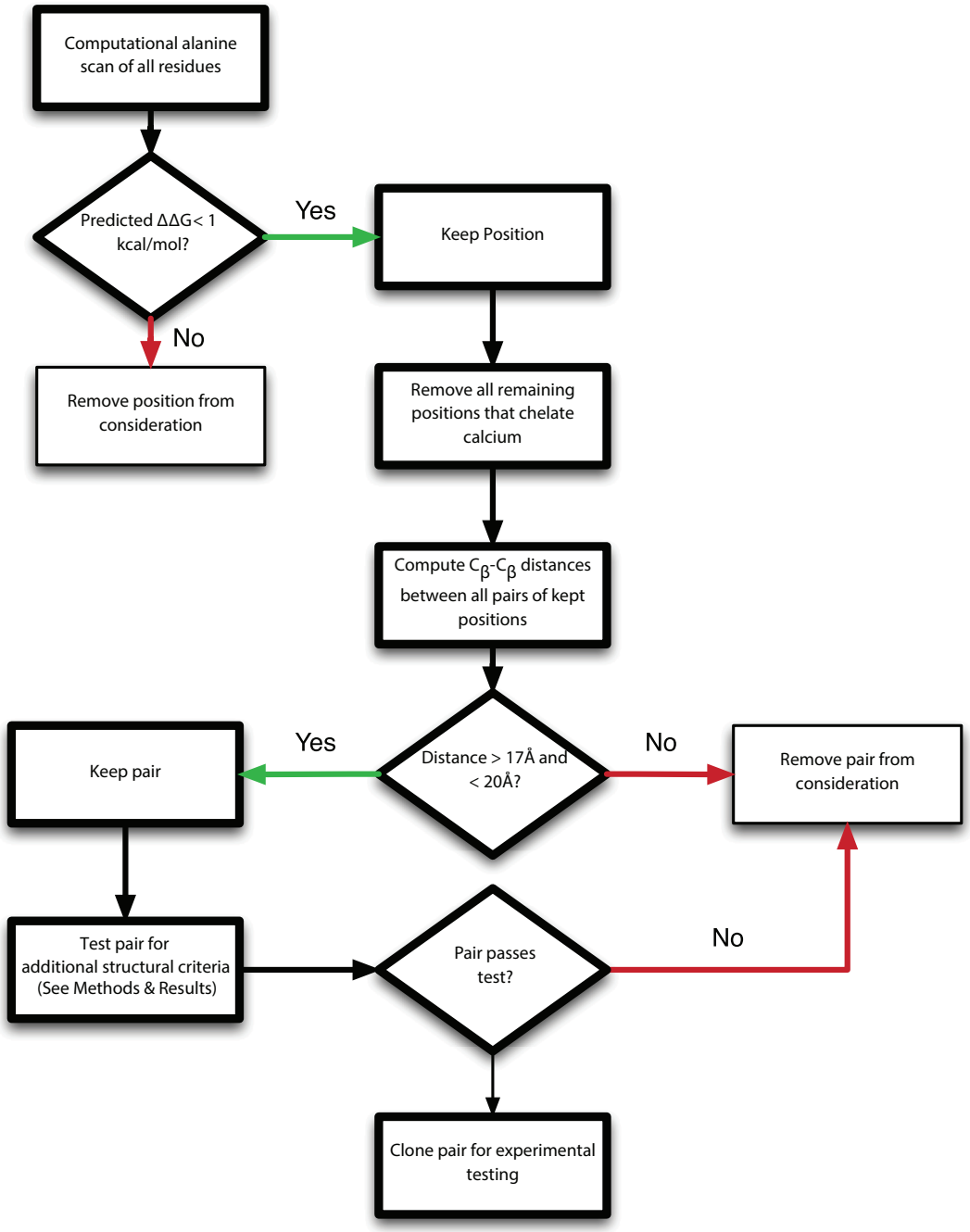

Figure 8 – Flow chart showing the computational methodology used to create the mutant library.

Chapter 4: Preliminary Characterization of Library Mutants

Introduction

After cloning the 11 double mutants (see **Appendix A**), I sought to do a preliminary characterization of each of them, in order to verify that they were stable after purification, that they conjugated the small molecule, and that they switched as expected. Instead of directly measuring stability, I instead monitored it by observing the extent of protein aggregation during the course of other measurements and assays. In comparison, conjugability and switchability were measured directly. Conjugability was estimated by looking at the relative heights of conjugated *versus* unconjugated peaks in mass spectrometry and computing the relative fraction that was conjugated, assuming equal ionizability. Switchability was measured by comparing the extinction coefficients of the thermally relaxed *trans* protein state to the *cis/trans* mixture after illumination and comparing the ratio to that of a known pure *trans* and pure *cis* mixture. After completing the analysis on all 11 mutants, it became clear that some were too unstable to be practical, while others conjugated or switched poorly. The data allowed me to choose my top candidates for further study, based on predictions of how functional they were likely to be.

Methods

Conjugability Measurements

Post conjugation (see **Chapter 5**), conjugation reactions were diluted to a final protein concentration of 1 μM in pure water, and then analyzed *via* mass

spectrometry using an LCT Premier (Waters). I observed the peak ratio between the unconjugated and conjugated (+453 Da) peaks, and, assuming equal ionizability for each species, made an estimate of the fraction conjugated. Highly conjugatable mutants showed only trace remaining unconjugated proteins, whereas poorly conjugatable mutants showed as little as 10% estimated conjugation.

Switchability Measurements

For each cadherin double mutant, I determined whether the BSBCA conjugated to the protein was photoisomerizable by illuminating *trans*-relaxed conjugated protein with UV light. *Trans* BSBCA contains a characteristic absorbance peak near 370 nm, and, upon illumination at that wavelength, the peak amplitude decreases as the small molecule isomerizes into the *cis* state.[10] Data provided to me by Dr. Andrew Woolley,[41] containing extinction coefficients for conjugated proteins separated by HPLC, containing isolated *trans* or *cis* isomers, allowed us to estimate the fraction of protein that switched to *cis*. By assuming the observed 370 nm absorbance of any mixture of *cis* and *trans* can be described by the simple sum of the independent absorbances of the underlying *cis* and *trans* states, switching percentages can be calculated by

$$Frac = \frac{\varepsilon_{370nm,trans} - \varepsilon_{370nm,mix}}{\varepsilon_{370nm,trans} - \mathrm{R} * \varepsilon_{370nm,trans}}$$

Equation 1 – Computation of protein fraction switching to *cis* after illumination

where R is the *cis/trans* extinction coefficient ratio computed from the provided data (.0541), $\varepsilon_{370nm,trans}$ is the measured extinction coefficient at 370 nm for the thermodynamically equilibrated, 100% *trans* state, and $\varepsilon_{370nm,mix}$ is the measured extinction coefficient at 370 nm for the photostable, UV illuminated state containing

a mixture of *cis* and *trans*. Using this methodology, I computed a switchability for the free chromophore of 86% *cis,* close to published estimates of maximum switchability,[10] verifying the accuracy and utility of the calculations.

Isomerization Reversibility Measurements

Conjugated 129/138 at 12 μM in TBS (25 mM Tris pH 7.5, 150 mM NaCl) was illuminated using the same UV LED for 2 minutes and 370 nm absorbance was measured. After measurement, protein was illuminated with a 1 W LED (Sparkfun, emission maximum 455 nm, with residual intensity in the 500-550nm range) for 2 minutes, 370 nm absorbance was measured, and the process was repeated for 10 cycles.

Half Life Measurements

Conjugated proteins in TBS (25 mM Tris pH 7.5, 150mM NaCl) at 12 μM were illuminated with a 1 W UV LED (emission maximum 365 nm) for 4 minutes and the absorbance at 370 nm was monitored immediately after illumination and then every 20 minutes thereafter for a total of 180 minutes (**Figure 11**). Between measurements, protein was kept in the dark. The absorbances were blank subtracted and then each absorbance was subtracted from 1.0 and the combined numbers were fit to a single exponential decay function of the form

$$y(t) = A * e^{-bt} + c$$

Equation 2 – Generalized form of a single exponential decay

where t is the time since illumination and $y(t)$ is the absorbance at time t. Curves were fit using the curve fitting toolbox in MATLAB (The MathWorks). R^2 values were typically >.995. Half-lives, λ, were computed as $\lambda = {\ln(2)}/{b}$.

Results

Conjugatability

The 11 mutants showed a large degree of variation in their stabilities. The most stable showed little to no aggregation after conjugation, and remained stable for several weeks afterward, while the least stable precipitated visibly during conjugation and, in the worst case, conjugability could not even be measured due to the lack of monomeric protein to inject into the instrument. (Note: the purification and conjugation protocol was not fully optimized [**Chapter 5**] until after my work had focused exclusively on the mutant 129/138. Therefore, the stabilities and conjugatabilities of the other proteins could be underestimated due to non-optimal handling). An example of a stable and highly conjugatable protein, 129/138, is shown in **Figure 9**. In this example, unconjugated protein (red) was run separately from the conjugated protein (green) to provide a mass standard. Virtually no signal at the unconjugated mass is present in the conjugated signal, indicating that this mutant is highly conjugatable. In comparison, mutant 6/90 conjugates only moderately well, with approximately 60% of protein remaining unconjugated after 3 days (**Figure 10**). The conjugatabilities of all mutants are summarized in **Table 1**.

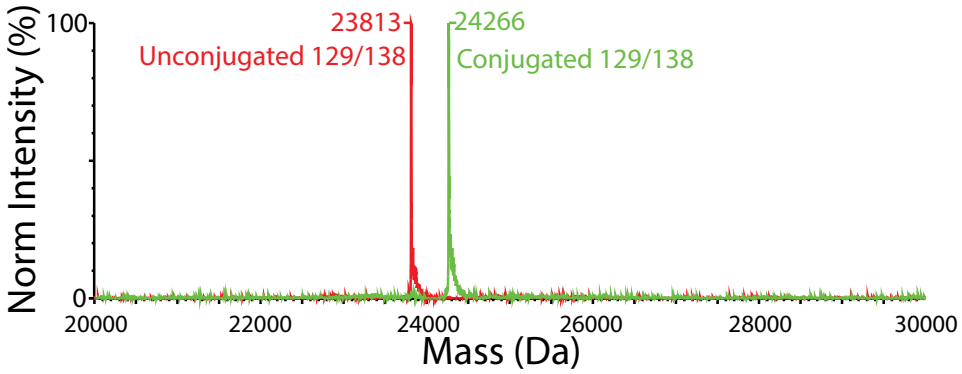

Figure 9 – An example of one of the highly conjugatable mutants, 129/138. Shown are two spectra overlaid to show both unconjugated and conjugated peaks.

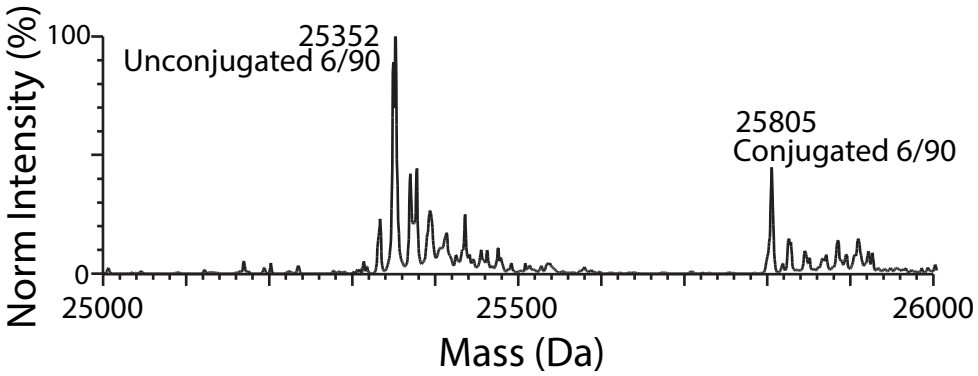

Figure 10 – An example of one of the less conjugatable mutants, 6/90. Shown is a single spectrum of a sample containing both conjugated and significant remaining unconjugated protein.

Switchability

The mutants also showed highly variable switchabilities, ranging from just under 50% to 100%. An example of how the basic switchability of one mutant, 129/138, was measured in shown in **Figure 11**, where illumination with UV light significantly changes the absorbance at 370 nm. The switchabilities of all 11 mutants are summarized in **Table 1**.

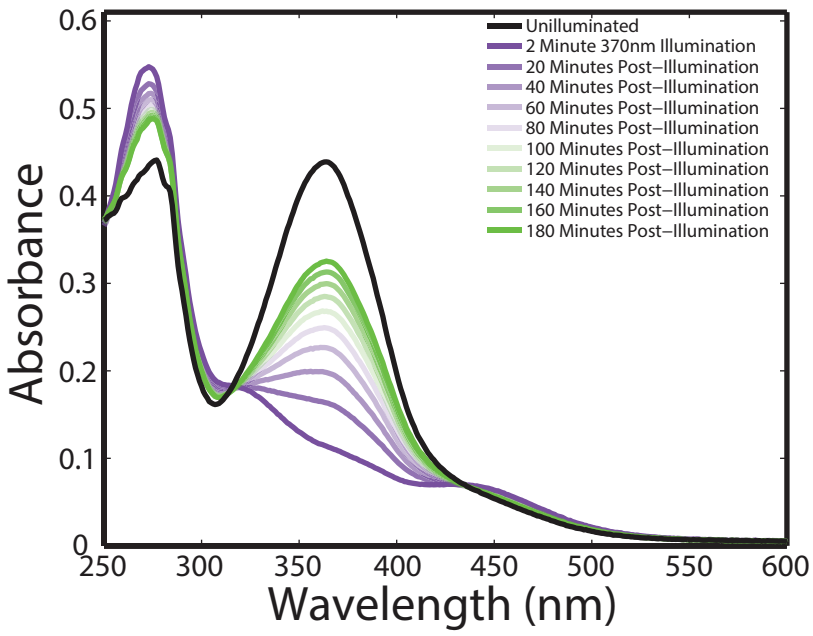

Figure 11 – Illumination and subsequence relaxation of 129/138 used to measure switchability and half-life.

For 129/138, a mutant that was characterized in further detail, I measured two additional properties: the extent to which photoisomerization was reversible and repeatable, and the extent to which it was titratable via short illumination times. The reverse isomerization (pan-visual illumination, including 500-550 nm bands) behaved as expected, leading to a reappearance of the absorbance band of the *trans* state. I illuminated 129/138 for 10 complete UV-green illumination cycles without any apparent loss of absorbance or switchability. The switchability was titratable via shorter illumination times, also as expected (**Figure 13**).

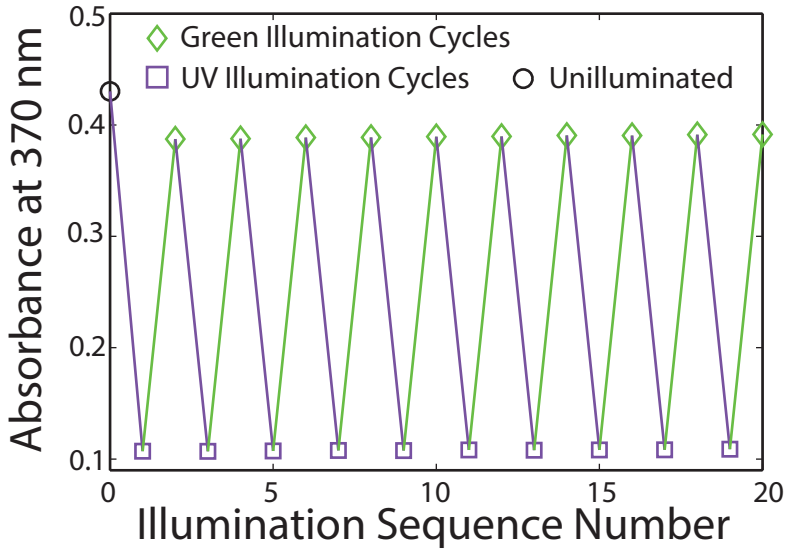

Figure 12 – The absorbance of *cis* and *trans* 129/138 after repeated illumination cycles.

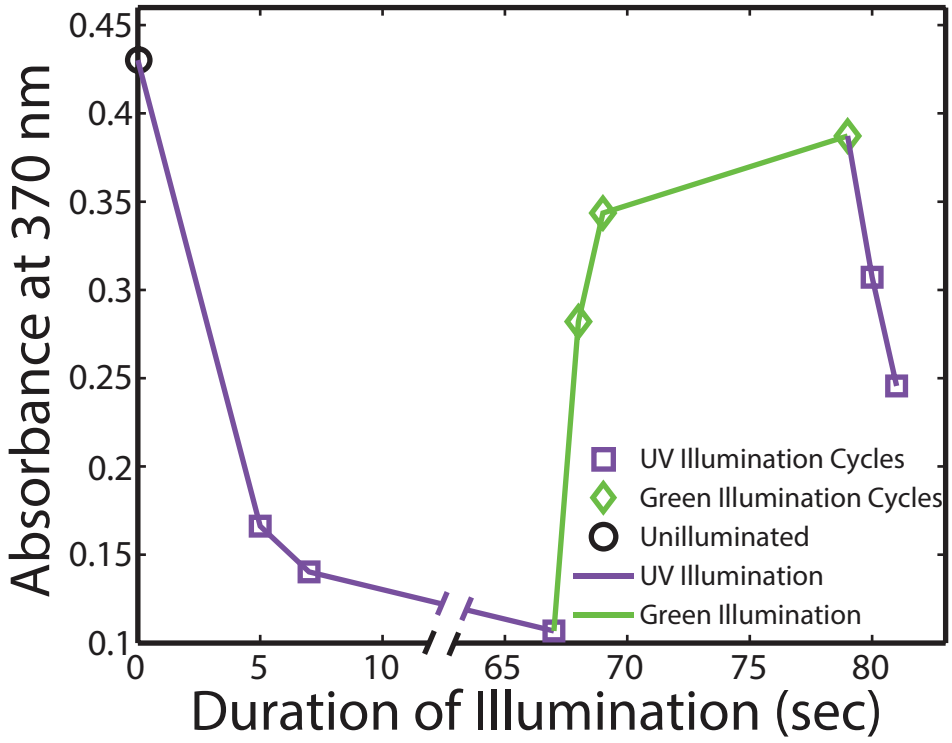

Figure 13 – The absorbance of *cis* and *trans* 129/138 after illumination cycles of variable length.

Half-Life Determination

The mutants that underwent further examination by SPR (**Chapter 6**), were additionally tested to measure their half lives, in order to determine how frequently

reillumination would be necessary to maintain *cis* state proteins. Half-lives varied from 36-72 minutes (**Table 1**), with the half-life of 129/138 significantly longer than the others. Without more detail, it is not possible to determine the reason for the differences in half life, but it is notable that the mutant with the longest half life was the mutant with the greatest functionality. A long half-life suggests the activation energy barrier to thermal relaxation is relatively high. In addition, 129/138 also has relatively high switchability; this could be caused by the *cis* state having a relatively low energy in the ensemble of states, perhaps mimicking one in the native state ensemble and causing it to be relatively favorable as the chromophore returns to the ground state after photon excitation.

Mutant	Stability	Conjugability	Switchability	Half Life (min)
129/138	High	High	80%	72
70/14	High	Poor	51%	36
70/107	High	High	100%	56
5/137	Poor	High	68%	44
6/90†	Poor	Moderate	51%	ND
111/135†	Moderate	High	62%	ND
16/57†	Poor	High	45%	ND
70/105†	Poor	ND	ND	ND
70/133†	Poor	Poor	ND	ND
95/139†	Moderate	Very Poor	ND	ND
70/138†	Moderate	Poor	ND	ND

Table 1 – Summary of preliminary characterization data on all 11 mutants tested. ND: not determined. †These mutants were tested with an earlier version of the EC12 construct that contains a non-native Met residue at the N-terminus.

Chapter 5: Optimization of Conjugation & Cleavage

Introduction

Initially, I focused my efforts on doing initial tests of all members of the library so that I could quickly identify which mutants were most likely to be switchable and functional. After that initial screen identified 129/138 as the best candidate and I turned to SPR to characterize the mutant in more detail, I began to be limited by my ability to produce protein. Because of the weak interactions between cadherin monomers, substantial amounts of protein are needed for SPR (10s of mg). Although I was able to express that much protein, the multiple steps involved in the purification caused extensive loss, such that my final yields were less than 10% of my initial starting mass. Through a careful, step-by-step optimization of the entire purification protocol, I was able to eventually recover >90% of my starting protein, substantially improving my yields and limiting the number of purification preps I needed to do. The complete, optimized, protocol for protein production is in **Appendix A.** What follows in this chapter is the series of optimizations that I did and how that improved the purification.

Methods & Results

Motivation for Construct Design

The initial constructs I used for the preliminary assays were different from the constructs I used for the detailed characterization of 129/138. The first experiments, including those on the WT, used an EC12 construct that contained an alanine removing the single native cysteine (C9A), shown previously not to affect

cadherin's function[26], a non-natural ATG start codon, and a C-terminal 6xHis tag using a cloning process previously described.[12] This construct also contained a double arginine insert between the protein sequence and the introduced 6xHis tag, allowing removal of the tag using trypsin *via* a protocol from the same work.[12] The protein resulting from these constructs tended to be relatively stable, and because trypsin was highly efficient, reactions could be performed quickly, which allowed the protein to remain stable during the cleavage reaction.

In addition to the altered protein sequence, I initially used a more complex conjugation protocol that used additional purification steps to separate the remaining unconjugated protein from the fraction that conjugated (see **Appendix C**). These additional steps were required for two reasons. First, although 129/138 conjugated well, other proteins did not; this necessitated the removal of the substantial fraction that did not conjugate in order to analyze the properties of the fraction that did. Second, because the conjugation buffer conditions had not yet been optimized, even highly conjugatable proteins such as 129/138 still required additional purification steps.

However, after I realized that the presence of the non-native methionine at the N-terminus of the protein was likely to interfere with protein function, I switched to a different construct. Natural E-cadherin contains a pro-peptide cleaved during protein maturation, leaving an N-terminal aspartate residue instead of a methionine. This aspartate forms an important intramolecular salt bridge, resulting in recombinantly produced proteins containing an N-terminal methionine having altered function.[42] In order to remove this methionine, I moved the 6xHis tag to the

N-terminus of the protein and added a TEV protease cleavage site immediately after it, allowing its removal without any non-native residues left in the protein sequence. The C9A mutation was left in place. I also replaced the double arginine cleavage motif with a stop codon, to remove the previously present C-terminal tag, avoiding a need for double-cleavage. While other proteases have been shown to work on EC12's N-terminus, including factor Xa[22], the majority of these proteases are incompatible with the presence of reducing agents. As I needed to cleave the proteins prior to conjugation for reasons explained below, the use of reducing agents was a necessity in order to prevent cysteine oxidation and subsequent non-reactivity with the chromophore. I attempted to use factor Xa in the presence of TCEP, but was unsuccessful (data not shown).

The TEV cleavage site N-terminal extension substantially reduced protein stability, both directly, and because cleavage required protein to be exposed to elevated temperatures for several days. As a result, yields dropped sharply, and protein purification began to take the vast majority of my time. I, therefore, sought to re-optimize the cleavage, and conjugation protocol. Because the yields of protein directly from lysates were high, I did not further optimize that section of the protocol.

Optimization of Cleavage/Conjugation Order

In principle, conjugation could come before or after cleavage of the N-terminus, because the TEV and chromophore do not interact or interfere with each other. Initially, I tried both options. However, when I attempted to cleave protein after conjugating, I discovered something unexpected: conjugated proteins

aggregated and precipitated in the presence of the Ni-NTA beads used to remove uncleaved protein and contaminants, causing a dramatic loss in yield. Initially, I suspected that a different resin or buffer condition might alleviate that problem. BSBCA has moderate hydrophobicity, and I suspected that perhaps this hydrophobicity was inducing aggregation in the presence of salt. However, I tried several different brands of Ni-NTA and similar 6xHis affinity resins, combined with many different buffers, but had similar results with all of them (**Table 2**). The best buffer/resin combination resulted in a loss of approximately 35% of the protein, while the worst resulted in a loss of nearly 70%. Because I wished to be able to use Ni-NTA resin to remove residual uncut protein and TEV protease in one step, I decided to cleave proteins prior to conjugation, so that conjugated proteins would not require exposure to the offending resin.

Buffer	Resin	% Monomeric
5	GE	64.3
4	GE	61.3
6	GE	61.2
7	Qiagen	58.5
9	Qiagen	57.6
7	None	53.8
5	Qiagen	51.1
4	Qiagen	50
1	GE	50
6	Qiagen	49.4
6	None	48.3
2	Qiagen	47.7
7	Talon	47.2
5	None	46.6
9	None	44.8
3	Qiagen	42.9
8	Qiagen	40.9
4	None	40.8

Buffer	Resin	% Monomeric
1	GE	50
1	Ni60	30.8
1	None	34.2
1	Qiagen	40.8
1	Talon	39.2
2	Ni60	31.3
2	None	35.9
2	Qiagen	47.7
3	Ni60	31.1
3	None	38.5
3	Qiagan	42.9
4	GE	61.3
4	None	40.8
4	Qiagen	50
5	GE	64.3
5	None	46.6
5	Qiagen	51.1
6	GE	61.2

1	Qiagen	40.8
1	Talon	39.2
3	None	38.5
2	None	35.9
8	Talon	35.7
1	None	34.2
9	Ni60	33.3
8	None	32.4
2	Ni60	31.3
3	Ni60	31.1
1	Ni60	30.8

6	None	48.3
6	Qiagen	49.4
7	None	53.8
7	Qiagen	58.5
7	Talon	47.2
8	None	32.4
8	Qiagen	40.9
8	Talon	35.7
9	Ni60	33.3
9	None	44.8
9	Qiagen	57.6

Table 2 – The percent of protein that remains monomeric following repurification in various buffers. (left) Results ordered by monomeric percentage. (right) Results ordered by buffer and resin type.

Buffer #	Buffer Composition
1	25mM Tris pH 8.0, 400 mM NaCl, 10 mM Imidazole
2	25mM Tris **pH 7.5**, 400 mM NaCl, 10 mM Imidazole
3	25mM Tris **pH 8.5**, 400 mM NaCl, 10 mM Imidazole
4	25mM Tris pH 8.0, 400 mM NaCl, **20 mM** Imidazole
5	25mM Tris pH 8.0, 400 mM NaCl, **30 mM** Imidazole
6	25mM Tris pH 8.0, 400 mM NaCl, **50 mM** Imidazole
7	25mM Tris pH 8.0, **200 mM** NaCl, 10 mM Imidazole
8	25mM Tris pH 8.0, **800 mM** NaCl, 10 mM Imidazole
9	25mM Tris pH 8.0, 400 mM NaCl, 10 mM Imidazole, **0.1% Tween-20**

Table 3 – Buffers used in experiments presented in **Table 2**. Component in bold indicates the change from buffer #1.

Optimization of Conjugation Conditions

After determining the relative order of conjugation *versus* cleavage, I sought to optimize my conjugation buffer. Prior to any optimization, I had been using PBS, pH 8.5 with Dulbecco's salts (D-PBS) and added TCEP (25 mM $Na_xH_yPO_4$, 150mM NaCl, 3 mM KCl, 500 µM $MgCl_2$, 500 µM $CaCl_2$, 3 mM TCEP). The pH values and TCEP concentrations were chosen due to the higher reaction rate of BSBCA at elevated pH, and the need to keep cysteines reduced for reactivity.[23] D-PBS was chosen due to the salinity matching that of the extracellular space where cadherin is endogenously found. However, after recognizing the relatively weak buffer capacity of phosphate

at pH 8.5, I switched to 25 mM Tris pH 8.5 instead. In addition, prior to any optimization, conjugation reactions were run at 33° C, picked because it provided a rate enhancement over room temperature but remained several degrees below the unfolding temperature for one of the mutants, 70/14 as measured by CD (**Figure 14**, Note: although other mutants could have had lower unfolding temperatures resulting in protein instability at 33° C, this was not tested). Finally, three days was picked as the reaction time due to it being a balance between chromophore reactivity and protein stability.

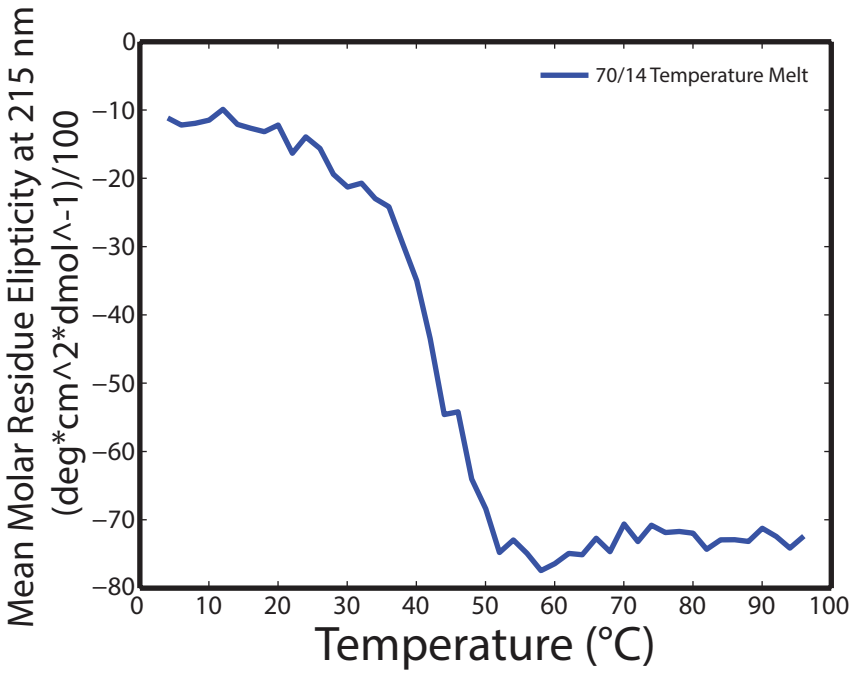

Figure 14 – CD melt of 70/14 used to determine initial conjugation reaction temperature. A temperature of 33 degrees was chosen, below the apparent unfolding temperature of approximately 40 degrees.

When considering factors to optimize, I chose to optimize pH, salt concentration, calcium/EDTA concentration, temperature, and reaction time. A summary of all of the conditions tested, as well as their optimal values, is summarized in **Table 4**. After examining all the data, I chose as a conjugation buffer

TBS pH 8.5 containing 25 mM Tris, 400 mM NaCl, 3mM KCl, 500 μM $MgCl_2$, 1 mM EDTA, and 3 mM TCEP, and I ran reactions for 3 days at 25° C.

The newly optimized buffer significantly increased the rate of the reaction, resulting in nearly 100% conjugation, while simultaneously increasing protein stability (**Figure 16**). I speculate that adding EDTA chelated any residual calcium, which prevented any transient dimers from forming and slowing the reaction. In addition, the increased salt concentration likely stabilized the protein, and, given the hydrophobicity of the BSBCA, likely favored the reaction. Lowering the temperature to 25° C additionally stabilized the protein without an obvious decrease in reaction rate, though the loss of rate was likely more than made up *via* buffer optimizations.

Condition	Values Tested	Optimal Value
pH	7.5, 8.5	8.5
[NaCl]	150, 300, 400 mM	400
[Ca2+/EDTA]	500 μM Ca^{2+}, 0, 1 mM EDTA	1 mM EDTA
Temperature	33, 30, 25 °C	25° C
Reaction Duration	1, 2, 3 days	3 days

Table 4 – Conjugation buffer conditions tested.

Optimization of Cleavage Conditions

TEV protease is active over a wide variety of temperatures (4-30 °C) and pH values (6-8.5),[43] which allowed me flexibility in the cleavage conditions. Because BSBCA is most reactive at higher pH values (>8.0), I had been using a buffer at pH 8.5 for conjugation and knew that the protein was stable at that pH. Additionally, some early experiments were run at a test pH of 6.5, which caused EC12 instability (data not shown). These two factors led me to focus on the pH range 7.0-8.5 for cleavage.

The first buffer I used contained 1X TEV buffer as provided by Invitrogen's protocol.[43] TEV protease, although active from 0-500 mM NaCl, is most active at lower salt concentrations, and thus the provided buffer contains no salt. However, when I attempted to mix EC12 with protease in this buffer, I saw an immediate and dramatic precipitation. This motivated me to seek other buffer conditions where the proteins might remain more stable.

Next, I attempted to cleave in TBS pH 7.5 with 400mM NaCl (25 mM Tris pH 7.5, 400 mM NaCl, 1mM EDTA, 1mM DTT), using a cutting temperature of 16 degrees for 60 hours based on a previously established protocol for cutting cadherin with factor Xa protease.[22,44] While this buffer prevented precipitation, it resulted only 50% cleavage, with the remainder left uncut, as estimated *via* protein concentration before and after removal of uncut protein and TEV via Ni-NTA (data not shown).

Given the increase in conjugation efficiency I saw with the optimized conjugation buffer, I tested whether this buffer may also be suitable for TEV protease cleavage. This buffer both kept the protein monomeric as measured by size exclusion (**Figure 15**) and allowed for nearly complete cleavage as measured by mass spectrometry and estimations of total protein before and after repurification (**Figure 16**). Based on these results, all subsequent cleavage reactions were done in this buffer. This buffer also allowed me to reduce the protease:cadherin mass ratio in the reaction from 1:4 to 1:8 without a loss of cleavage completeness or significant appearance of protein aggregation (**Figure 15**). I did not test ratios lower than 1:8.

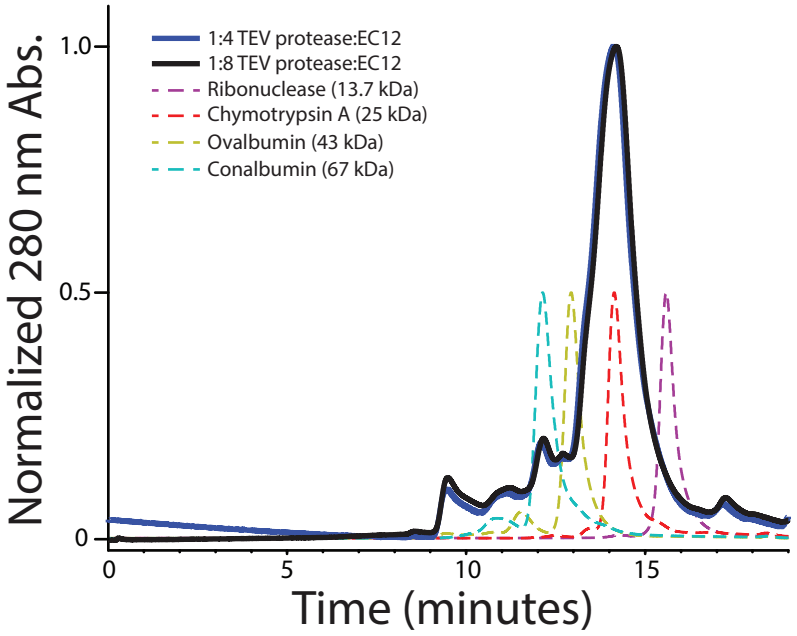

Figure 15 – Size exclusion chromatography of 129/138 in optimized cleavage buffer. Protein remains monomeric in this buffer at both TEV protease concentrations. Absorbances for 129/138 were normalized to the maximum peak height, for visual convenience. Also shown are size standards for comparison, normalized to a maximum peak height of 0.5.

After optimizing the cleavage buffer, all that remained was to optimize the repurification of the protein after cleavage, to remove residual uncut protein and TEV protease. I again used the optimized conjugation buffer, which gave high yields and low aggregation, and I continued to use it without further optimization.

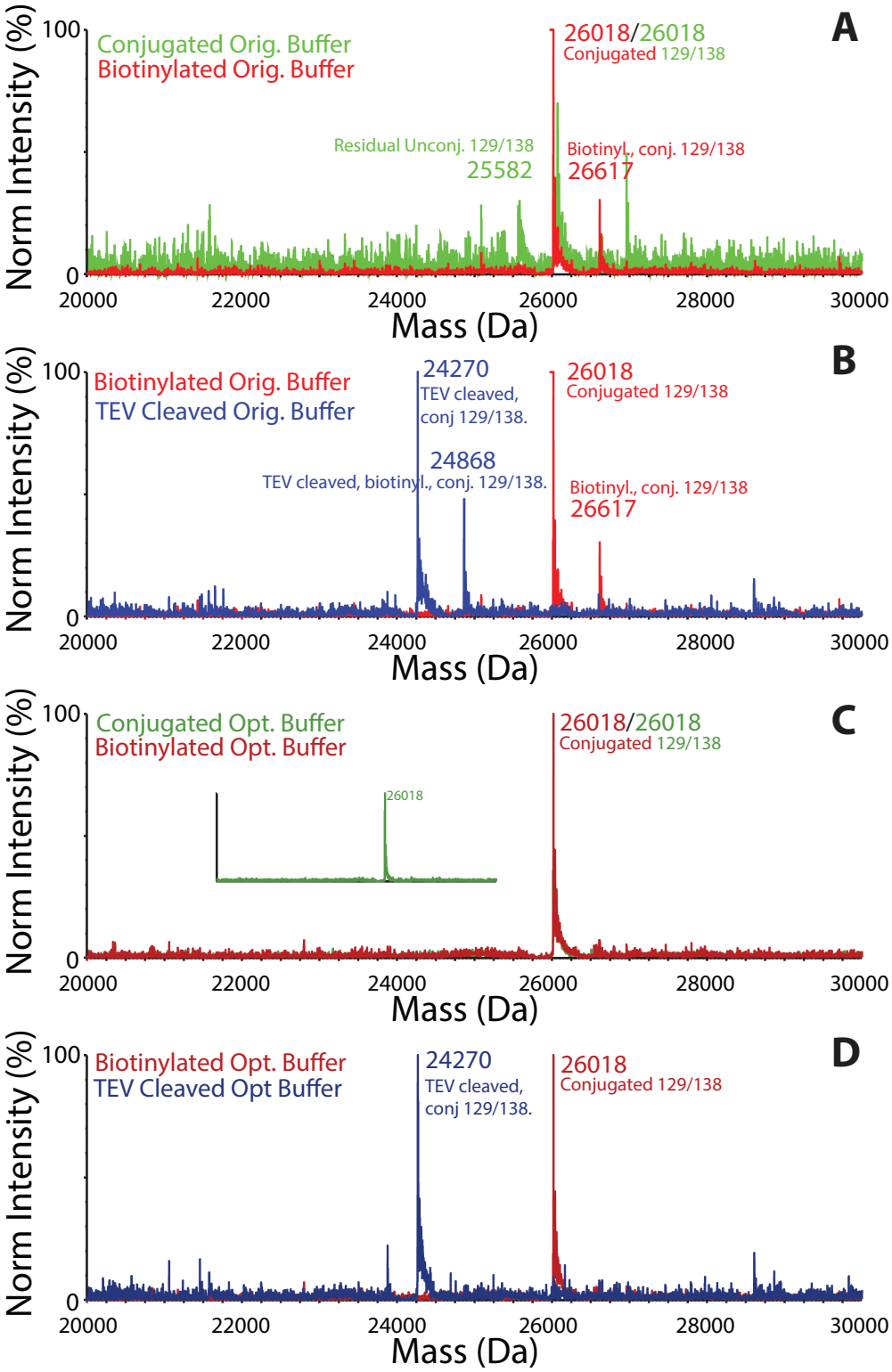

Figure 16 – Mass spectra in the original and optimized conditions. Shown are results from an earlier protocol which included protein biotinylation (see **Appendix C**) (A) Conj *vs.* bio 129/138 in the original buffer. Conj. does not go to completion. (B) Bio. *vs.* TEV protease cleaved in the original buffer. Cleaved spectra are post repurification, removing residual uncut protein. (C) Conj. *vs.* bio. in the optimized buffer. Conj. goes to completion, resulting in no peak for the biotinylated species. Inset shows conjugated protein alone, for visual clarity. (D) Bio. *vs.* TEV in the optimized buffer. Cleaved spectra are post repurification.

Conclusion

After fully optimizing the conjugation, cleavage, and repurification conditions, overall protein yields improved from 10% of total protein to 90% of total protein for unconjugated protein and 60% for conjugates. Shown below are example size exclusion chromatography traces prior to and after protocol optimization, showing a dramatic gain in monomeric fraction (**Figure 17**). Monomeric protein fractions collected remained monomeric when re-run in HPLC-based size exclusion chromatography, and showed dimerization when exposed to calcium (**Appendix D**).

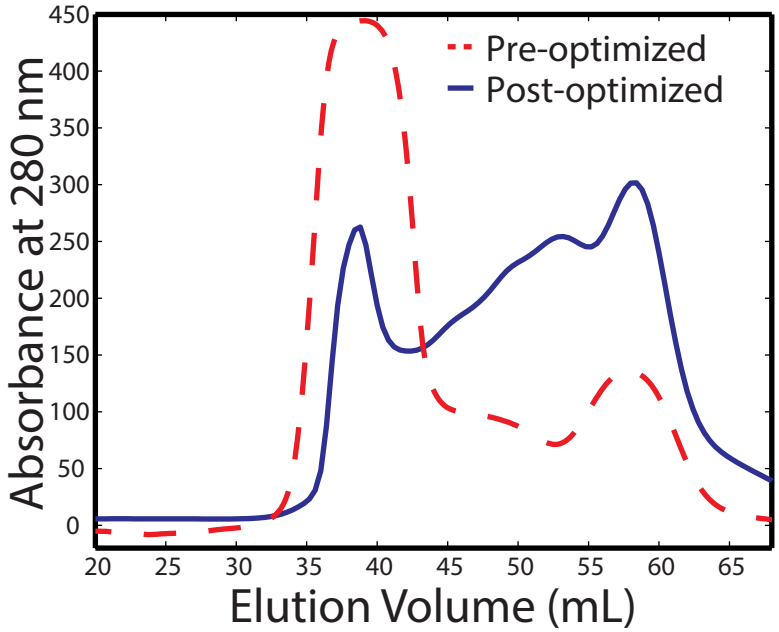

Figure 17 – Size exclusion chromatography of conjugated 129/138 before and after optimization.

Chapter 6: SPR Measurements of Calcium & Protein-Protein Binding

Introduction

The work presented in this chapter was done in collaboration with Kris Kuchenbecker. Kris was responsible for conditioning and functionalizing the surface of the chips we used, as well as operating the machine and assessing data quality. He also provided extensive advice in analyzing the data. I am very grateful for his help.

After completing the measurements in **Chapter 4**, and knowing that several of the mutants appeared to be switchable, I sought an experimental method that would more directly demonstrate changes in calcium binding as well as show whether these change in calcium binding also cause the expected changes in protein binding activity. I used surface plasmon resonance (SPR) to measure protein homodimerization as a function of calcium concentration. In this assay, similar to that of Harrison and colleagues[24], biotinylated WT C9A cadherin was immobilized to the SPR chip and WT C9A, *trans* or *cis* mutants were flowed over it. These experiments were first done on 129/138, and showed promise, resulting in a more detailed characterization. I also briefly examined 3 other mutants, though none showed any change in calcium binding affinity. The characterization of 129/138 confirmed that I could change protein-protein affinities by modulating calcium binding, confirming the photoswitch operated as designed.

Methods

Data Acquisition

Matrix-free, flat, carboxymethylated gold surfaces (GE Healthcare Sensor Chip C1) were used in all SPR experiments. Individual flow cells were prepared with the following protocol: (1) 50 µl injection of 1-Ethyl-3- (3-dimethylaminopropyl)carbodiimide / N-hydroxysuccinimide (0.5M:0.2M); (2) 30 µL injection of 0.25 mg ml-1 ImmunoPure Streptavidin (Pierce #21122) in sodium acetate buffer (pH 5.0) to a total amount of ~500 response units (RU) for all four flow cells; (3) 60 µl injection of 1 M ethanolamine. Cadherin was captured on the active flow cells by manual injection of 25 nM protein to immobilization levels between 250-450 RU. No blocking procedures were performed on the reference flow cell. Prior to each SPR experiment, protein samples were dialyzed against 25 mM Tris, 150 mM NaCl, and 500 µM TCEP. Following dialysis, Tween-20 detergent (Sigma-Aldrich #P9416) was added to the protein solution and the dialysis buffer to achieve a final concentration of 0.05% (v/v). The dialysis buffer was then used as the assay buffer for the SPR measurements.

Dose response titrations were prepared by serial dilutions of the highest concentration into assay buffer. Each sample response was subtracted by a reference response (containing no calcium for calcium titrations, and no protein for protein titrations). The responses of flow cells 2 and 3 were scaled to match flow cell 4 *via* a least squares minimization resulting in a single scalar multiplier for each flow cell.

For the detailed characterization of the WT and 129/138, additional steps were taken to maximize data quality: (1) While measuring *cis* 129/138, a maximum of three data points per illumination cycle were used, to maximize *cis* fraction and

minimize thermal relaxation. To minimize systematic error due to some fraction of the protein reverting to *trans* during the experiment, concentrations were injected in a random order. (2) Every sample injection was followed by a control injection of 20 μM WT C9A cadherin containing 1mM $CaCl_2$ to monitor degradation of chip response over time. Gathered signals were then corrected in magnitude using the magnitude of the control injection.

Solution Homodimer Mitigation

Direct measurements of both calcium affinity and homodimeric protein affinity in SPR are difficult due to solution homodimers competing with those on the surface. I presumed that surface homodimers would not further interact with the surface to form multimers, as cadherin has been shown not to form higher-order species at the concentrations used here. I further assumed that the time a particular molecule spent in the flow cell was too short to allow for solution dimer dissociation and surface dimer reassociation. Thus, the surface dimers reduced the effective protein concentration injected. To minimize solution homodimerization, I used a protein concentration (40 μM) below the K_d for homodimerization of WT EC12 cadherin.[24]

Data Fitting

For calcium titrations, each series was fit using the curve fitting toolbox in MATLAB to an equation of the form:

$$y = \left({bx^{N_h}}/{c^{N_h} + x^{N_h}} \right) + d$$

Equation 3 – General Hill equation

where c is the EC_{50} for calcium-dependent protein binding, N^h is the Hill coefficient, x is the calcium concentration, and y is the response measured. Due to chip-to-chip variation, fits were normalized such that the maximum response at infinite concentration ($x = \infty$) predicted by the fit lines was equal to one. During calcium titrations, I observed an EC_{50}-dependent, non-specific interaction at higher calcium concentrations, with a stronger (lower) EC_{50} leading to a larger magnitude effect (i.e. WT C9A had the strongest effect, *cis* 129/138 had no apparent effect at concentrations tested). Points dominated by this effect were not used in the fits, but are shown as faded markers in **Figure 19**.

Protein Titrations

I injected WT C9A, *trans* and *cis* proteins at a fixed background concentration of 1mM calcium and protein concentrations ranging from 0-300 μM using the same data acquisition methodology as I used for the calcium titrations. At concentrations of protein higher than 300 μM, I observed significant non-specific effects that prevented measurement. In these experiments, the chip is covered with WT C9A cadherins for all injections. As a result, for non-WT injections, solution homodimers and surface heterodimers are in competition, which may have significant differences in affinities. Due to this competition, quantitative determination of binding affinities is difficult, and fits were not determined as a result.

Bootstrap Analysis

For the calcium titrations for 129/138, I used a bootstrapping technique in order to obtain additional verification that the fit values were robust. Data points for each flow cell and concentration were grouped, and then data points were drawn at

random, with replacement, in a number equal to the number of actual data points. These randomly drawn data points were then refit to **Equation 3** and EC_{50} and N_h values were stored. This process was repeated 500 times for each protein state. The fits were then sorted by minimum sum of squared error (SSE), and the average and ±2 SD values were computed for the top 100 fits for each protein (100 out of 500 fits were used to avoid outlier random trials from affecting fit values).

Reversibility Analysis

In the reversibility experiments, 40 μM conjugated 129/138 was alternatingly illuminated with UV (emission maximum 365 nm) and blue LEDs (emission maximum 455 nm, with residual intensity at 500-550 nm), 1mM Ca^{2+} was added, and responses measured. The resulting responses were subject to the same reference and control subtraction, as well as flow cell scaling used in the calcium titrations.

Dissociation Analysis

For 129/138 and WT, I fit each protein variant's SPR dissociation traces to a single exponential of the form:

$$response(t) = a_x * e^{-k*t} + b_x$$

Equation 4 – Determination of cadherin off rates observed in SPR.

where *response(t)* is the response as a function of time, *t*, and *k* is the off rate. For each protein, *a* and *b* were allowed to vary per response, while all responses were simultaneously fit to a single shared off rate, *k*, that minimized the sum of squared errors (SSE) of the responses. Calcium concentrations less than 10 μM were

removed from the fits, as their responses were not significantly above the noise threshold of the instrument.

Results

Preliminary Calcium Binding of 70/14, 70/107 & 5/137

In order to determine whether any other mutants beyond 129/138 may have been functional, I investigated whether illumination of 70/14, 70/107 and 5/137 changed their apparent calcium binding affinities in SPR. The mutant 5/137 showed no specific binding to the WT cadherin on the SPR chip and data was not analyzed any further (data not shown). The other two mutants 70/14, and 70/107, each showed some binding, though neither showed significant changes in calcium binding after illumination (**Figure 18**). The first, 70/14, showed binding in *trans* though the flow cells did not show similar line shapes with flow cell 4 diverging significantly from flow cells 2 and 3, so no overall fit was determined. In *cis*, 70/14 showed binding that saturated, but showed a significant amount of noise, and thus also was not fit. These results could have been due to partial protein instability or aggregation; the mutant was not investigated further. In comparison, 70/107 showed binding with good fit quality, but no difference between pre- and post-illumination. *Trans* 70/107 had EC_{50} = 246.28 ± 16.71 μM and N_h = 1.44 ± 0.11, while *cis* showed EC_{50} = 216.36 ± 83.15 μM and N_h = 1.10 ± 0.47. Although all three of these mutants were tested prior to the full optimization of the SPR assay, they were not tested in more detail when I had results indicating the functionality of 129/138. Mutant 5/137 often showed instability and aggregation, which was unlikely to change with SPR optimization. The data were of high enough quality to

conclude that 70/107, although conjugatable, was not functionally switchable. Finally, 70/14 appeared to be more stable than 5/137, but still showed evidence of aggregation. Notably, prior work has showed that mutations at position 14 in EC12 ablate binding in SPR assays[24], which may have been the source of the problem.

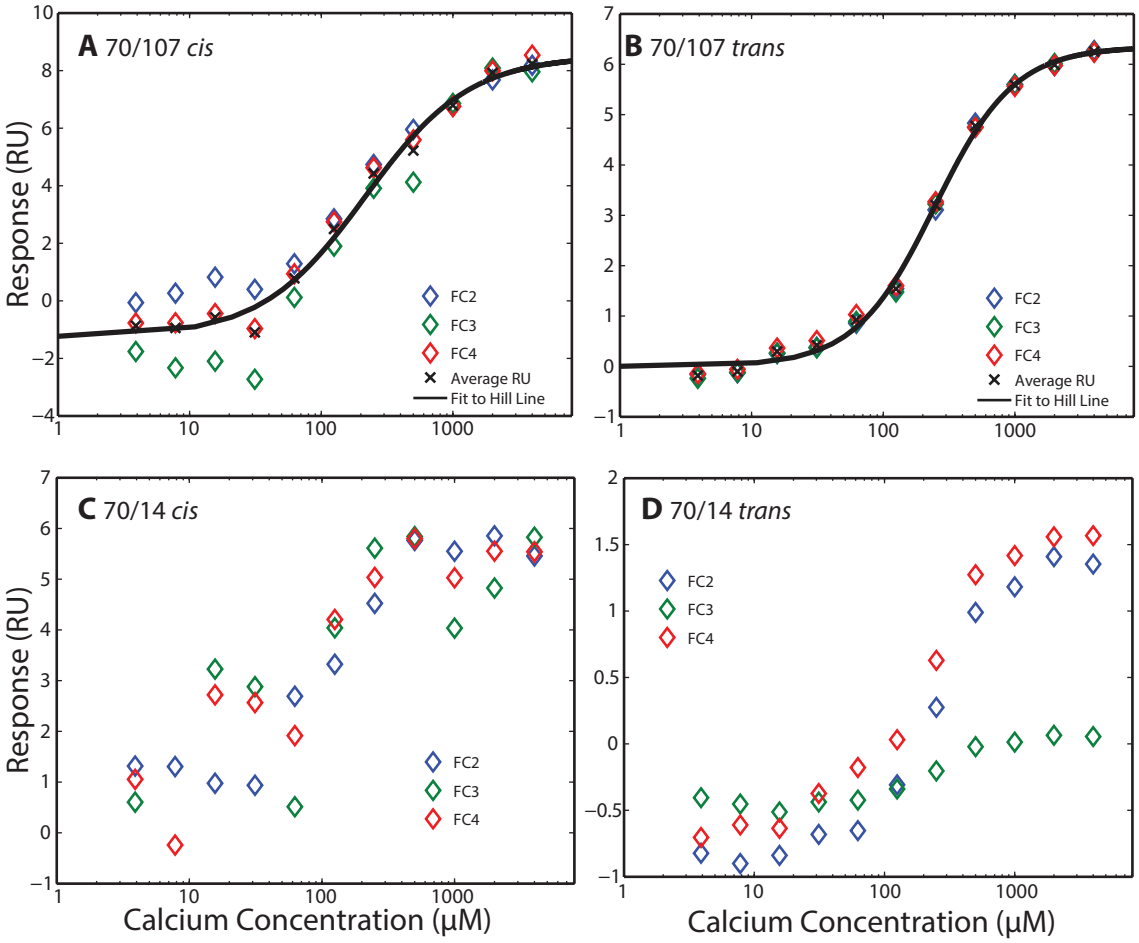

Figure 18 – Data from mutants 70/107 and 70/14 pre- and post-illumination in SPR. Mutant 5/137 did not show any measurable binding (data not shown).

Detailed Calcium Binding of 129/138

In comparison to the other mutants, 129/138 showed significant functional switching upon illumination. I investigated the 129/138 mutant, along with WT C9A as control (**Figure 19**, with raw traces in **Figure 20**). I observed a calcium binding EC_{50} for C9A cadherin, as measured by a Hill fit to a single species, of 72.0 μM, with

mean fit values and ±2 SD error, as measured by a bootstrapping analysis of the data, of 71.2 ± 14 μM. I also measured a Hill coefficient (N_h) of 2.24 (2.45 ± 1.7). Unconjugated 129/138 had an EC_{50} = 104 ± 27 μM and N_h = 1.37 ± 0.47.

In comparison, *trans* 129/138 had an EC_{50} of 156 μM (170 ± 33 μM), with N_h = 1.38 (1.28 ± 0.28). C*is* 129/138 showed substantially weakened binding, with EC_{50} = 619 μM (611 ± 180 μM) and N_h = 0.76 (0.77 ± 0.15), demonstrating a nearly 4-fold change in apparently calcium affinity after illumination under these conditions.

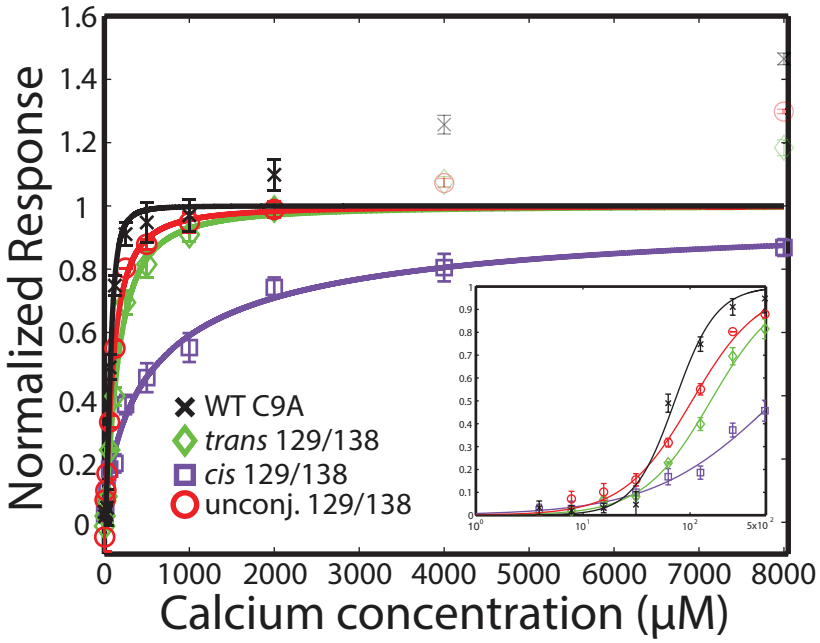

Figure 19 – Fits to Hill equations of the responses measured in SPR. Dimmer points were dominated by non-specific binding and were not used in the fits, but are shown for completeness.

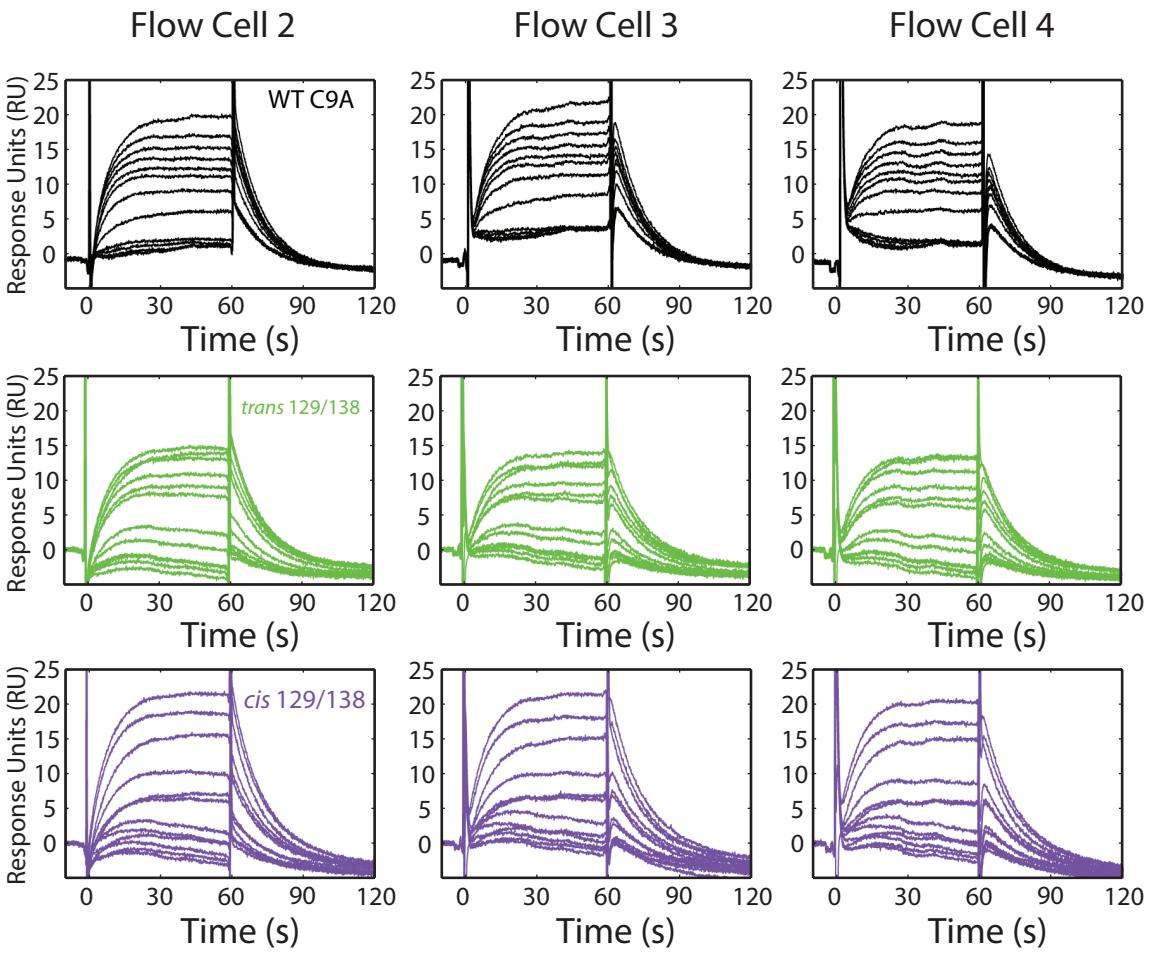

Figure 20 – Raw traces for data shown in **Figure 19.**

These EC_{50} values are different from those measured in the mass spectrometry and half-life analysis described below (**Chapters 7 & 8**); these differences are due to cadherin binding multiple calcium ions to function. Because of the need to bind multiple calcium ions, any measured EC_{50} for calcium binding must be, at a minimum, a multiple of the protein concentration used (with the multiplicity required depending on the extent of cooperativity). In these assay, I used a protein concentration of 40 μM. As such, the EC_{50} for WT C9A being approximately two-fold higher than the protein concentration indicates that it binds with extensive cooperativity, as otherwise at 72 μM calcium most proteins would have only 2 calcium ions bound. Similarly, the four-fold higher EC_{50} for *trans* 129/138 compared to the protein concentration indicates it is substantially less cooperative, a finding confirmed in the data. However, the approximately 15-fold higher EC_{50} for *cis* 129/138 indicates that calcium binding is weak enough that the protein concentration used does not substantially change the measured EC_{50}, and the number measured is accurate.

For *cis* 129/138, the calcium binding EC_{50} is lower (binding is stronger) than that measured in the absorbance assays of **Chapter 8**. This discrepancy is likely due to the lack of complete photoisomerization upon illumination, as well as partial thermal relaxation to *trans* prior to measurement. The assay in **Chapter 8** is measured more quickly after illumination, and only monitors changes in *cis* state.

An alternative explanation for the observed decrease in the SPR signal upon isomerization to the *cis* state could be an increase in *cis* homodimerization in solution, effectively reducing the concentration of cadherin monomers available to

bind to the WT cadherin immobilized on the chip. To exclude that possibility, I analyzed conjugated 129/138 homodimerization via size exclusion chromatography (**Appendix D**). The observed decrease in the dimer/monomer ratio after UV illumination in this assay additionally confirms the expected weaker *cis* homodimerization upon illumination.

Treatment of Strand-Swapped Dimers

Questions remain about the structure of the functional cadherin multimers, including evidence that cadherin forms strand-swapped dimers.[24,45-47] In the SPR assay, I was unable to directly determine the structure of the interacting species formed in the experiments. However, I fit each set of SPR traces for a given cadherin variant to a single off rate returning to baseline levels, even at higher calcium concentrations (**Figure 21**). For WT, the fit-determined off rate was k_{wt} = 0.091 sec^{-1}, for *trans* X-EC12, k_{trans} = 0.075 sec^{-1}, and for *cis* X-EC12, k_{cis} = 0.072 sec^{-1}. Because multiple different dimer structures would likely have different off-rates, especially if one were strand-swapped, this ability to fit to a single rate is consistent with a single dimer type formed for each assayed cadherin variant.

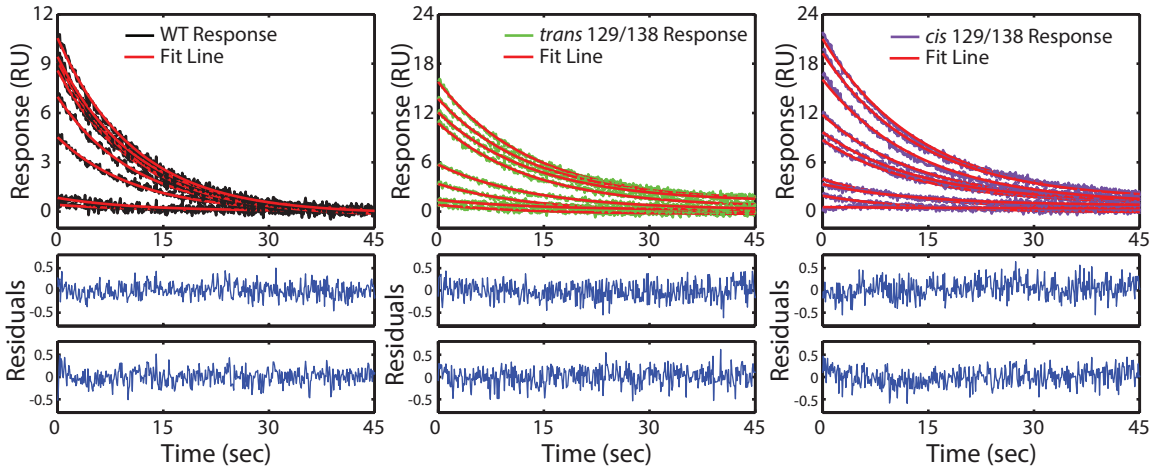

Figure 21 – Fits to single off rates to SPR dissociation of WT as well as *trans* and *cis* 129/138.(top) Fits for WT, *trans* 129/138 and *cis* 129/138. (bottom) Residuals to fits, shown for [Ca^{2+}] = 62.5 μM (upper) and 1000 μM (lower). Reponses are from flow cell 2.

Bootstrapping Analysis

I performed a bootstrapping analysis of the fit values in the calcium titrations in order to verify they were robust (**Figures 22 & 23**). Computed values were all near the values reported for the single best fit using the gathered data, which indicates the data values describe the system well and the fit values are robust. The largest variance in the values came from the Hill coefficient of the WT and the EC_{50} of *cis* 129/138.

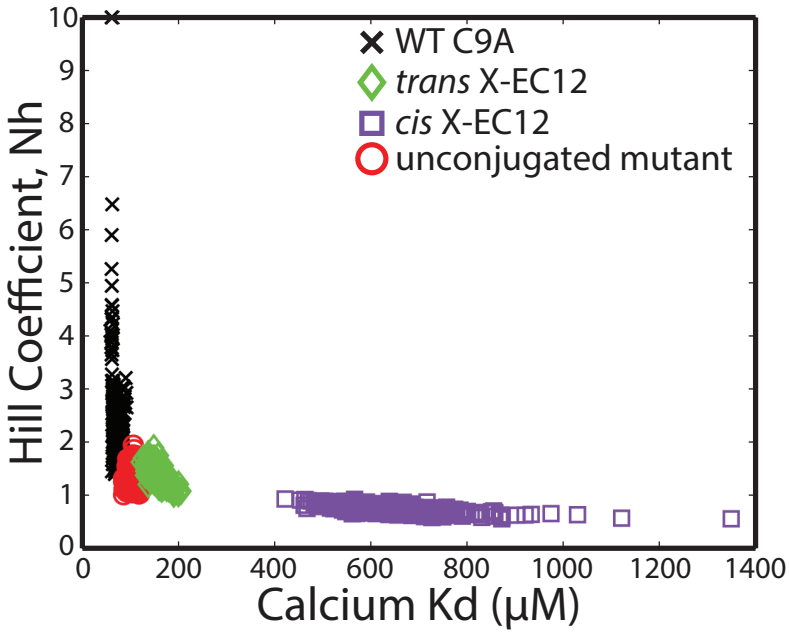

Figure 22 – Bootstrapping analysis of 129/138 calcium titrations. All 500 trials shown.

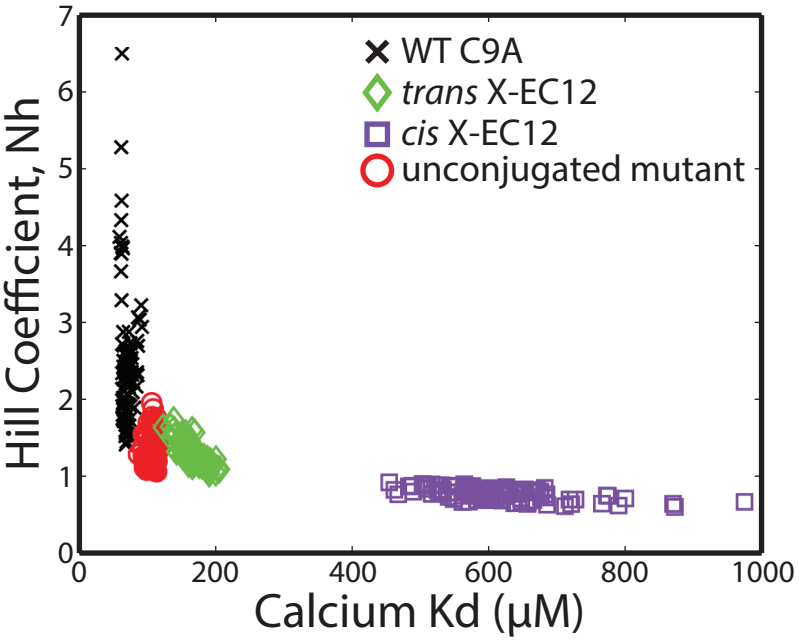

Figure 23 – Bootstrapping as in **Figure 22**, with only the top 100 points by minimum sum of squared error (SSE) shown. For each protein, these values were used to compute the mean and ±2 SD values.

Protein Titrations of 129/138

In order to more directly measure changes in protein-protein binding affinity, I injected WT C9A, *trans* and *cis* proteins at a fixed background concentration of 1mM calcium and protein concentrations ranging from 0-300uM (**Figure 24**). As mentioned in the methods section, quantitative determination of thermodynamic parameters was not possible due to complicating factors, though qualitative comparisons can still be made. The *trans* EC12 construct binds more weakly than WT, indicating a combination of the cysteine mutations and addition of BSBCA weakened either calcium or protein-protein binding. Nevertheless, at all injected concentrations, *cis* state shows a weaker response than *trans,* indicating that it has an apparent weaker protein-protein binding affinity, confirming that isomerization affects protein-protein binding affinity. This measured difference in protein-protein binding after illumination was reversible as measured over multiple illumination cycles with 40 μM protein and 1mM Ca^{2+} (**Figure 25**).

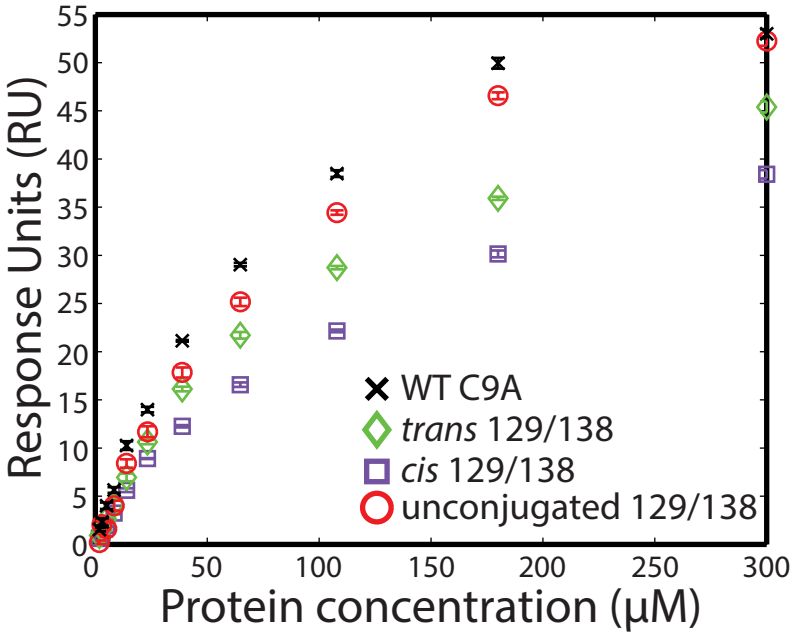

Figure 24 – Protein titrations of 129/138 at concentrations of 0-300 μM protein in a background of 1 mM Ca^{2+}. Due to complicating factors, fits were not determined.

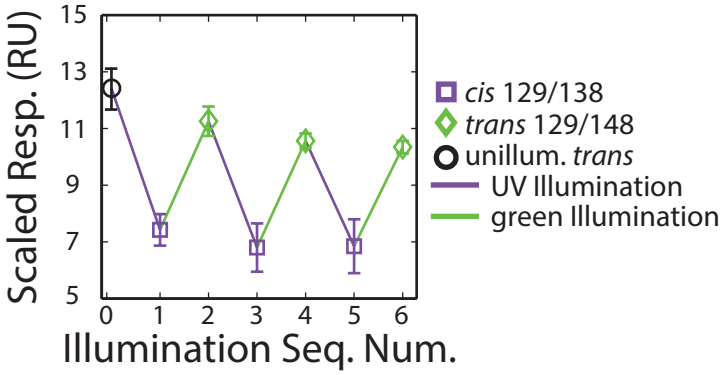

Figure 25 – Reversibility of changes in SPR response.

Chapter 7: MS Measurement of Calcium Binding Affinity

Introduction

After seeing a difference in calcium binding affinity after illumination in SPR (**Chapter 6**) but being unsatisfied with the indirectness of the SPR measurements, I sought a more direct measurement of WT and 128/138 calcium binding affinity. To do so, I relied on a previously described mass spectrometry-based assay[22] to directly measure calcium binding affinity of WT C9A as well as *trans* and *cis* 129/138. If isomerization alters calcium binding, the *cis* 129/138 should have weaker affinity than *trans*. In addition, because EC12 binds 3 calcium ions, any of which could be interfered with during isomerization, a decrease in apparent cooperativity as measured by the Hill coefficient[48] (N_h) would be expected. I observed WT calcium binding K_ds in line with previous results, as well as *trans* calcium binding affinity near WT. In contrast, *cis* bound so weakly that calcium binding K_ds could not be quantitated, as binding was outside the dynamic range of the instrument. Perplexingly, unconjugated 129/138 also showed limited specific binding of calcium, even though in the SPR assays it appears to dimerize equally well as WT at the same calcium concentrations. Reasons for that discrepancy are discussed, along with the results, below.

Methods

Protein samples were prepared as previously described[22], with the modification that samples were diluted to 2 pmol/µL (2 µM). Samples were injected

at 5 μL/min into a QTRAP4000 instrument (Agilent). *Cis* 129/138 was illuminated with a 1 W UV LED (Advancemart, emission maximum 365 nm) for 4 minutes immediately prior to injection. The quantity of bound calcium ions was obtained by determining Ca^{2+} binding occupancies, and assuming that calcium-free and calcium-bound molecules have the same ionizability.[49] By comparing peak areas, fractions of molecules binding 0 to 9 calcium ions were computed for each calcium concentration. For determination of K_d, molecules binding more than three calcium ions were assumed to bind three ions specifically, and these fractions were added to the 3-ion fraction. The average number of calcium ions bound was computed, and the resulting numbers were fit to a Hill equation of the same form as **Equation 3**, where here x is the calcium concentration, y is the number of calcium ions bound, and c is the calcium binding K_d. In order to subtract non-specific binding from *cis*, the following equation was used for each calcium concentration, *i,* and occupancy number, *c*:

$$F_{c,i} = \dot{F}_{c,i} - \left(\sum_{a=1}^{c} \dot{F_{c,\iota}} * x_{a,i}\right) + \left(\sum_{a=1}^{\min(9-c,6)} \dot{F}_{c+a,i} * x_{a,i}\right)$$

Equation 5 – Calculation of the true number of specific calcium ions bound

where $F_{c,i}$ is the true fraction of the molecules binding *c* specific calcium ions at concentration *i*, $\dot{F}_{c,i}$ is the apparent fraction binding *c* calciums at concentration *i*, and $x_{a,i}$ is the fraction of the molecules that bind *a* non-specific calciums at concentration *i*. The first summation in the equation subtracts from the apparent fraction contributions due to non-specific binding, whereas the second term adds from the other calcium occupancy states their non-specific fractions that actually

bind *c* calciums. Note that the maximum of the second sum is not *c* because, discounting the 3 specific sites, there are only 6 non-specific occupancies measurable, with a maximum of 9 total calciums bound. For *c* = 8, the only addition from other occupancies can be from the 9 occupancy state, which if it actually bound 8 specifically, would bind one non-specifically, resulting, for 8 calciums, a sum over *a* from 1 to 9-8, or 1, the only possibility. For *c* = 2, there can be no addition from the 9 occupancy state, as that would presume 2 specific and 7 non-specific calciums. Because the 9 occupancy state is explicitly presumed to bind 3 specific and 6 non-specific calciums, there is no measureable frequency of 7 non-specific calciums to use to make the correction (i.e., there is no $\dot{F}_{8,i}$). Thus, for *c* = 2, the sum is from 1 to 6, the maximum number of measureable non-specific calciums.

The non-specific binding fractions, $x_{a,i}$, came from the *trans* 129/138 calcium series using the assumption that all three calcium binding sites were occupied prior to non-specific sites, thus any fraction that appeared to bind four calciums actually bound three specific and one non-specific calcium, etc. This is not an entirely accurate assumption; however, any errors *underestimate* non-specific binding, making any observed differences between *cis* and *trans smaller* than actuality.

Results

WT C9A cadherin specifically bound three calcium ions with a dissociation constant K_d = 28.5 ± 1.9 μM (errors are the boundaries of a 95% confidence interval) and extensive cooperativity with N_h = 2.85 ± 0.47, close to previously reported numbers[22] of K_d = 20 ± 0.7 μM and N_h = 2.6 ± 0.2 (**Figure 26**, with example of primary data used in plots in **Figure 27**). *Trans* 129/138 showed 2-fold weaker

affinity and less cooperativity, with K_d = 55.2 ± 5.8 μM and N_h = 1.80 ± 0.35, but also bound three calcium ions.

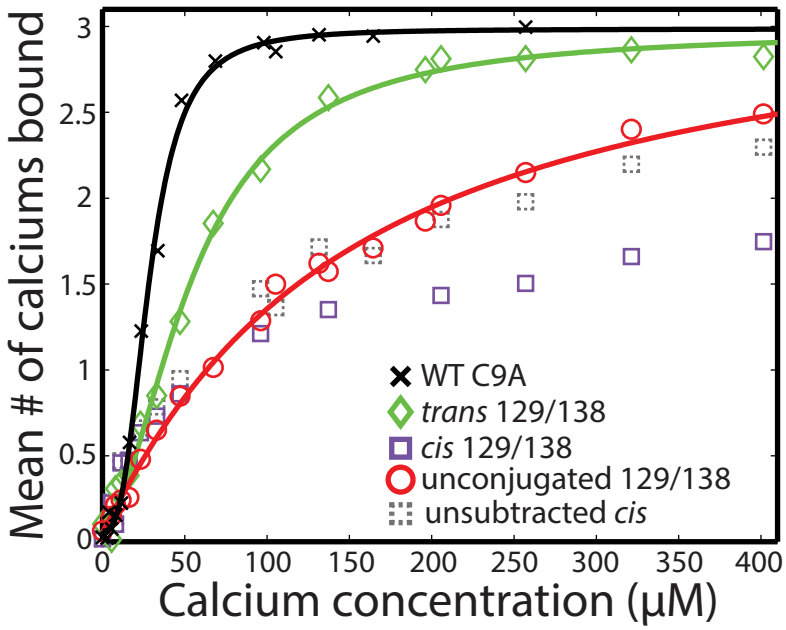

Figure 26 – Calcium binding affinity of WT C9A cadherin, as well as the different 129/138 states, as measured by mass spectrometry. Grey dotted boxes show *cis* 129/138 prior to non-specific binding subtraction.

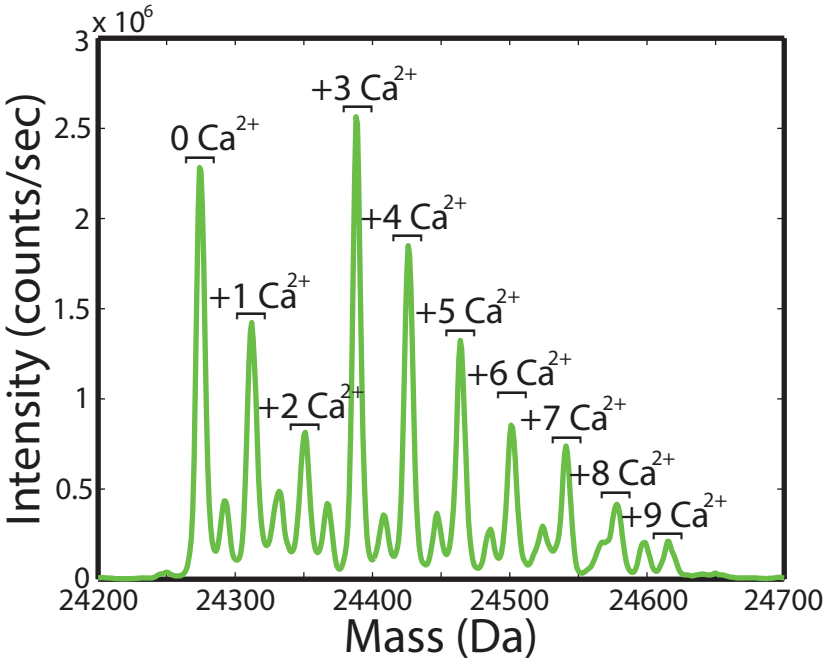

Figure 27 – Example mass spectra showing calcium binding to *trans* 129/138. [Ca^{2+}] = 96 μM. Unlabeled intermediate peaks are +1 Na.

In contrast, calcium binding to *cis* 129/138 was dominated by non-specific binding and a lack of cooperativity (**Figure 28**). *Cis* showed a significant fraction of single and double bound calcium states, which, combined with the observation of non-specific binding in the higher occupancy states, indicates that much of the observed double and triple bound occupancies are the result of non-specific binding *versus* cooperative binding at the specific binding sites. By using quadruple and higher calcium-bound states in the *trans* 129/138 sample as a reference for non-specific binding, I subtracted the estimated contribution of non-specific calcium binding from the measured average calcium occupancy for *cis* 129/138 (see Methods). The resulting line shows significantly reduced binding compared to *trans.* (A quantitative fit was not possible because saturating the specific sites required calcium concentrations outside the dynamic range of measurement.)

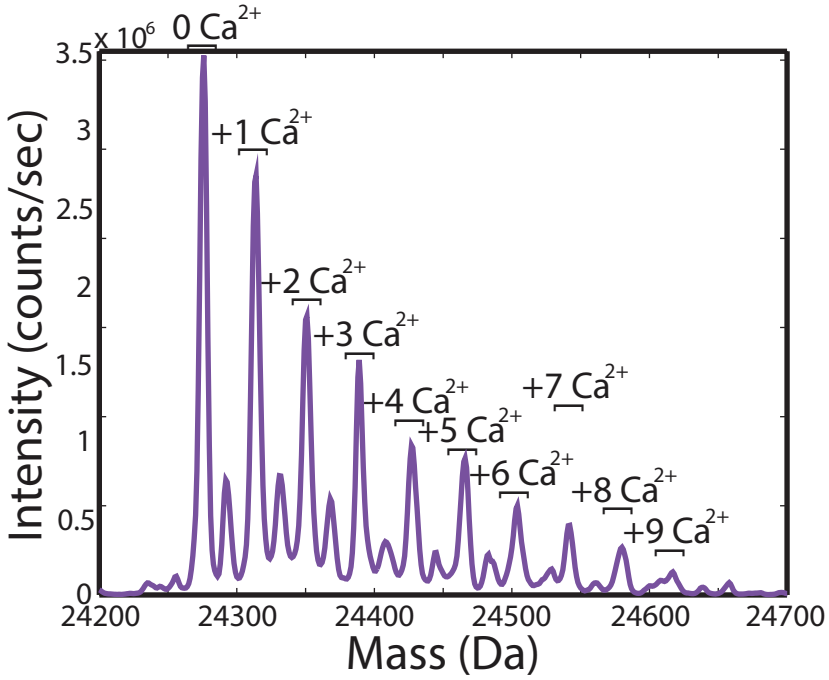

Figure 28 – Example mass spectra showing calcium binding to *cis* 129/138. [Ca^{2+}] = 168 μM. Unlabeled intermediate peaks are +1 Na.

Curiously, the unconjugated mutant also showed significantly weakened binding when compared to WT or *trans*, with K_d = 145 ± 18 μM and N_h = 1.07 ± 0.13. This would seem to directly contradict the SPR results (**Chapter 6**), where the unconjugated mutant shows stronger protein binding than *trans* at the same calcium concentration, apparently binding calcium more strongly. One possible explanation is that the addition of the exogenous cysteine residues destabilizes the calcium-bound loop conformation. The addition of the small molecule could stabilize the loops, recovering the lost affinity. This would explain the relative affinities as measured in the mass spectrometer, but still does not explain the differences in SPR *versus* MS.

Chapter 8: Changes in Half-life *vs.* Calcium Concentration

Introduction

After seeing a change in calcium binding affinity after illumination in SPR (**Chapter 6**), and MS (**Chapter 7**), but being unable to directly measure the calcium binding affinity of *cis* due to the indirectness of the assay (SPR) or lack of dynamic range (MS), I sought an assay that would allow me to accurately quantify calcium binding to *cis* 129/138. I turned to measuring the change in *cis* state half-life after the addition of calcium, relying on the coupling between calcium binding and half-life. These data gathered are for only the fraction of the population that is *cis*, allowing me to overcome the inevitable mixed populations that occur post-illumination and directly measure *cis* properties. This assay showed *cis* calcium binding to be EC_{50} = 996 ± 135 μM, a number I feel is the most accurate measurement of calcium binding affinity for this state.

Methods

For the calcium dependence of half-life, proteins were illuminated, and then $CaCl_2$ was added to final concentrations from 3.9 μM to 32 mM, using a 14-point 1:1 serial dilution series, plus a zero calcium point. Data points were collected five at a time on a Cary 50 Bio UV/Vis spectrometer (Varian) every 20 minutes for a total of 180 minutes, and the entire series was run in triplicate. Absorbance values for each calcium concentration over time were plotted and the half-life was calculated using an equation of the same form as **Equation 2**. The mean half-lives for each calcium

concentration were plotted and fit to a Hill curve of the form in **Equation 3**, using the curve fitting toolbox in MATLAB, where N_h is the hill coefficient and c is the EC_{50} for calcium-dependent half-life reduction. Absorbance curves were also fit to double exponentials to determine half-lives, without a significant increase in curve quality (data not shown).

Results

The *cis* state of BSBCA is thermodynamically unstable and it relaxes back to the stable *trans* state in the dark with a half-life of approximately 20 minutes at room temperature,[23] although conjugation to proteins can alter chromophore half-lives.[7,50] One can compute these half-lives by observing the increase in absorbance at 370 nm caused by the *cis* chromophore returning to *trans.*

One general caveat inherent to azobenzene-based strategies is that switching to the *cis* state is generally incomplete, i.e. the *cis* state always contains a minor *trans* population,[10,23] making measurements of *cis* properties difficult. However, the entirety of any change observed in half-life measurements is due only to the *cis* subpopulation, allowing measurement of pure *cis* properties unaffected by the small fraction that remains *trans*. If chromophore isomerization significantly affects calcium binding in my conjugated constructs, with stronger binding of calcium to *trans* 129/138, by thermodynamic coupling I expected to see that the free energy of the *cis* state ensemble increased as it bound calcium. This, in turn, would lower the free energy difference between the *cis* ensemble and the transition state between *trans* and *cis*, increasing the rate of thermal relaxation to *trans* and thus decreasing the half-life of *cis*. Because calcium binding is required to bring about the reduction

in half life, the change in half life can be used a proxy for calcium binding, with the midpoint of the half-life change in duration being approximately equal to the K_d for calcium binding of the isolated *cis* state. While the transition state free energy could also be changing as calcium is introduced, I assumed this was unlikely, as I assumed the major determiner of transition state free energy was the required destabilization of the azo bond in BSBCA, which would likely be unaffected by the presence of calcium.

As expected, I observed a half-life decrease from approximately 72 minutes to 28 minutes, with an EC_{50} of 996 ± 135 μM calcium (**Figures 29 & 30**). This represents a nearly 18-fold change in apparent calcium binding affinity from the 55 μM for *trans* 129/138 (mass spectrometry analysis, **Chapter 7**). I also observed a cooperative transition in half-life duration, with a measured Hill coefficient[48] (N_h) of 2.4 ± 0.74. This apparent cooperativity indicates that the re-binding of calcium to the *cis* state is cooperative, and thus that illumination likely interferes with multiple calcium binding sites.

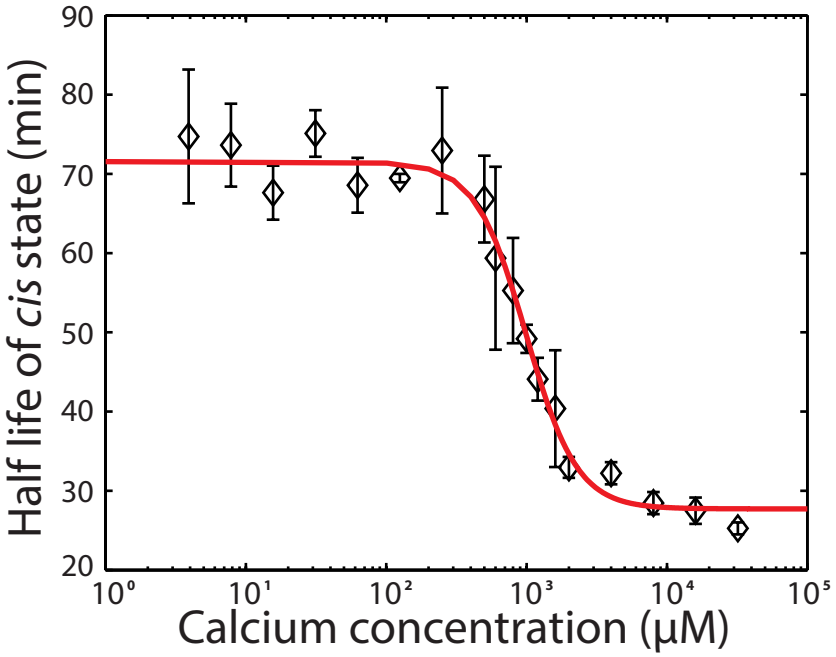

Figure 29 – The calcium concentration dependence of *cis* 129/138 half-life. EC_{50} = 996 ± 135 μM. Error bars are ±1 SD of 3 separate experiments.

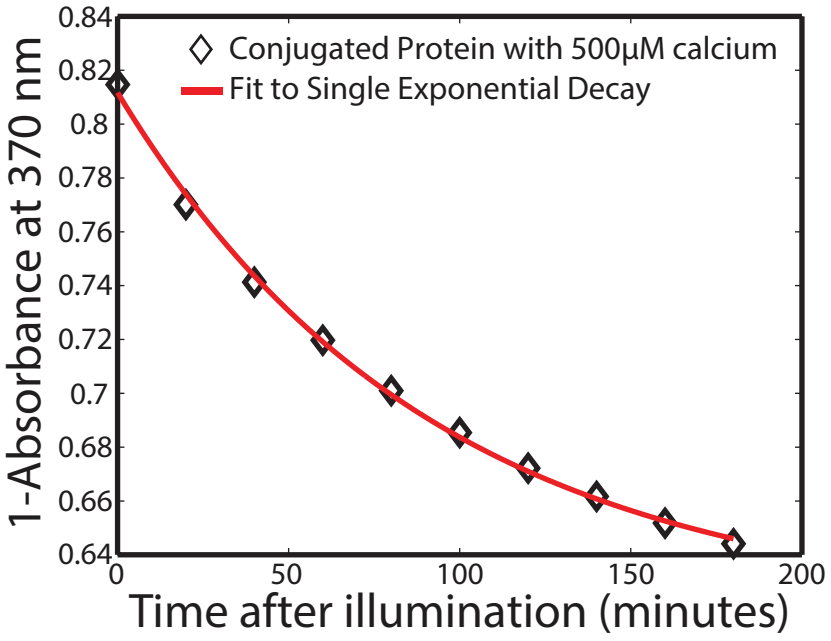

Figure 30 – A demonstration of the fit, to a single exponential, of the relaxation of conjugated protein after illumination. For this illumination, $[Ca^{2+}]$ = 500 μM, and the fit gave a half-life of 63.4 min.

Chapter 9: Conclusions & Discussion

In sum, the work presented here demonstrates the successful design of a reversibly photoswitchable cadherin variant. When illuminated with light, its calcium binding affinity changes from 56 μM to 996 μM, a nearly 18-fold change, and this change in affinity is coupled to a change in protein–protein binding. This successful design was identified after cloning a library of only 11 mutants, all of which were identified using a computational, rational design approach.

The next phase of this project is to apply this designed cadherin to experiments with cells. When applied in cell culture experiments, it could help answer many of questions about cadherin's function presented in the introduction. One way to introduce this engineered molecule into a cellular context is *via* cadherin-coated substrata. Coated substrata have been used to study cell-cell adhesion[51,52] and stem cell pluripotency[53,54]. Creation of coated surfaces allows spatial control of cadherin patterning and fine control over cadherin concentration, which could help maximize switchability of cadherin-mediated adhesion.[51] Although more difficult, the sites of mutation could also be engineered directly into the full-length cellular cadherin and used to study cadherin-cadherin signaling in a more natural context.

One natural question about this work is whether it is generalizable to other targets. Although interest in photoswitchable proteins has increased in recent years, relatively few examples exist in the literature, perhaps because finding good cysteine attachment points remains difficult. Compared to high-throughput and

other library techniques, I was able to create a successful conjugate using a rational design strategy and a small library of constructs. In principle, many of the steps used in the design process presented in **Chapter 3**, such as geometric compatibility, could be applied to any target, which is likely to significantly reduce the size of any experimental library. Additionally, the strategy presented here of targeting loop regions of proteins is likely to be successful elsewhere, though one must take care to avoid targeting loops so flexible they may accommodate the isomerization without changing function, as appears to be the case for the cadherin mutant 70/107. While disturbing helices and sheets would seem to more likely result in a dramatic change in protein function, they are also likely to require more energy from the isomerization, reducing the chances of isomerization at all.

One constraint on this current design, and all azobenzene-based designs, is the present inability of this chromophore to switch completely to *cis*. Several new chromophores have become available[50,55-57] that possess either more complete isomerization or longer half-lives that could allow for isolation and use of the pure *cis* state. Breakthroughs in this area are likely to result in large increases in the utility of these small photoisomerizable molecules. Because of the extreme spatial and temporal precision offered by light, interest and use of them is likely to continue to help solve the most recalcitrant questions remaining in biology.

Appendix A: Optimized Protocol for Cloning, Purification and Conjugation of mutants, WT C9A and WT-Avi C9A

Cloning

Natural E-cadherin contains a pro-peptide cleaved during protein maturation, leaving an N-terminal aspartate residue instead of a methionine. This aspartate forms an important intramolecular salt bridge, resulting in recombinantly produced proteins containing an N-terminal methionine having altered function.[42] To remove both the exogenous 6x-His tag and methionine, I introduced a TEV protease cleavage site. TEV protease is compatible with the reducing buffer conditions (containing 3 mM TCEP) required to maintain free cysteines in the engineered constructs prior to conjugation, whereas other common proteases are not.

Met-EC12-6xHis was cloned into pET-22b(+) (EMD Millipore # 69744-3) as described previously.[12] 6xHis-TEV-EC12 contained a 6xHis tag and TEV cleavage site (ENLFYQ) added N-terminally, as well as an inserted C-terminal stop codon between EC12 and the C-terminal 6xHis tag using the 5' oligonucleotide (GACCCCATATGCACCACCACCACCACCACGAAAACCTGTACTTCCAGGGGACTGGGTC) (blue-6xHis, orange-TEV cleavage site) and the 3' oligo (TGGTGCTCGAGTCAAGGAGCGTTGTCA). For 6xHis-TEV-EC12-avitag, the same 5' oligo was used, and the 3' oligo was replaced with one containing both an AviTag[58] and stop codon (TGGTGCTCGAGTCATTCGTGCCATTCGATTTTCTGAGCCTCGAAGATGTCGTTCAGACC GCCACCAGGAGCGTTG). (green-*avitag*) The resulting PCR products were then re-

inserted into pET-22b(+) using the same procedure as before. Cysteine 9 was mutated to alanine using site directed mutagenesis and the oligos (5') GTCATCCCTCCCATCAGCGCCCCCGAAAATGAAAAGGGC and (3') GCCCTTTTCATTTTCGGGGGCGCTGATGGGAGGGATGAC.

To insert the designed cysteines, site directed mutagenesis was used, inserting a single mutation at a time. Forward and reverse oligonucleotides were reverse complements of each other and 39 bases long. In all cases, the mutated codon was mutated to a TGC, and this codon was in the center of the oligonucleotide, with 18 bases paired to the WT clone flanking the mutation.

Protein Expression

6xHis-TEV-EC12 and Met-EC12-6xHis were expressed in *E. coli* BL21(DE3) cells as described previously,[12] with the substitution of 100 μg/mL carbenicillin replacing ampicillin. The 6xHis-TEV-EC12-avitag construct was co-transformed into BL21(DE3) cells with a plasmid containing biotin ligase (Gift from R. Fletterick). Protein was expressed under the same conditions as 6x-TEV-EC12 with the addition of 34 μg/mL chloramphenicol and 50 μM biotin (Thermo #29129) to LB culture medium. (BD #244610)

Lysis

Cells were resuspended at 20 mL/L of culture in lysis buffer (50 mM Tris-HCl pH 8.0, 2.5 mM $MgCl_2$, 0.1% Triton X-100). After resuspension, protease inhibitor tablets (Roche #11836170001) were added according to manufacturer's instruction. After addition of 1U/mL of DNAse I (NEB #M0303) and 0.1 mg/mL lysozyme (Sigma #L6876), resuspended cells were stirred for 1 hour at room temperature. Cells were

subsequently lysed using sonication, then lysates were cleared by centrifugation at 21,000x*g* for 30 minutes at 4 °C. (*Note: In my hands, EC12 is prone to aggregation if NaCl is present in the lysis buffer prior to clearing of lysates.*)

Ni-NTA Purification

After lysis, NaCl and imidazole were added to final concentrations of 300 mM and 10 mM, respectively, followed by 50% NiNTA agarose slurry (Qiagen #30230) at a volume of 5 mL/L of culture. Lysates were then nutated at 4 °C for 1-2 hours, and beads were pelleted by spinning at 3000x*g* for 10 minutes. Beads were resuspended in wash buffer (50 mM NaH_2PO_4-NaOH pH 8.0, 300 mM NaCl, 10 mM imidazole), separated *via* a gravity-flow chromatography column (BioRad #732-1010), then further washed *via* 2 column volumes (CV) of wash buffer. Protein was eluted from the beads using 1.25 CV elution buffer (50 mM NaH_2PO_4-NaOH pH 8.0, 300 mM NaCl, 250 mM imidazole). Eluted protein was injected into a HiLoad 16/600 Superdex 75 column (GE #28-9893-33) attached to an FPLC system and prequilibrated in TEV cut buffer (25 mM Tris-HCl pH 8.5, 400 mM NaCl, 1 mM EDTA, 3 mM KCl, 3 mM TCEP). Monomer fractions were pooled. Typical yields of monomer ranged from 10-20 mg protein per liter of cell culture.

Cleavage

Purified proteins were concentrated to a final concentration of 120 μM in TEV cut buffer using spin concentrators (Millipore #UFC901024) (*Note: In my hands, EC12 adsorbs to alternative protein concentrators*), after which 6xHis-TEV (Invitrogen #12575-015) was added in a mass ratio of 1:8 TEV:EC12. Protease reactions were left at 16 °C for 60 hours.

Repurification

After cleavage, protease, tags, residual contaminating proteins, and residual uncleaved protein were removed using a second round of NiNTA purification. The protease reactions were desalted using a HiPrep 26/10 desalting column (GE #17-5087-01) into repurification buffer (25 mM Tris-HCl pH 8.0, 400 mM NaCl, 30 mM imidazole, 500 μM TCEP). After desalting, Ni-NTA agarose pre-washed in repurification buffer was added at a ratio of 1mL of packed resin per 15 mg EC12 and mixtures were nutated for 1-2 hours at 4 °C. After nutating, supernatants containing cleaved protein were separated from beads using a gravity-flow chromatography column. Beads were washed using 1CV repurification buffer, and wash fractions were pooled with supernatants. Proteins were then desalted into TEV cut buffer for medium term storage. Protein purity was verified by SDS-PAGE and was estimated to be > 90%, while cleavage completeness was verified *via* mass spectrometry. (**Figure 31A**) Biotinylation efficiency was also verified *via* mass spectrometry. (**Figure 31B**) Typical yields post-repurification were 66-75% compared to precleavage, with the major loss being incomplete cleavage. Biotinylation was typically >95% of overall protein.

3,3'-bis(sulfonato)-4,4'- bis(chloroacetamido)azobenzene (BSBCA) Handling

Dry BSBCA (purchased from Linkera-osadovsk@chem.utoronto.ca) was kept at 4 °C. 16X stock solutions were created by dissolving dry BSBCA into water at a concentration of 8 mM. Stock solutions were kept at -20 °C in the dark.

Protein Conjugation

Repurified and cleaved mutant protein was concentrated to a final concentration of 160 μM in TEV cut buffer, followed by addition of BSBCA from

stock solutions to a final concentration of 500 μM. Conjugation reactions were kept in the dark at 25 °C for 72 hours. After conjugation, excess chromophore was removed *via* desalting using a HiPrep 26/10 desalting column (GE #17-5087-01) into TEV cut buffer. Conjugation efficiency was verified using mass spectrometry, looking for the characteristic increase of 453 Da[23], and efficiencies were typically >99% for the mutant K129C/D138C (**Figure 31**). (*Note: Free chromophore has a tendency to adsorb to the resin, especially in the presence of salt. It may be removed by flushing the column with several volumes of pure water.*)

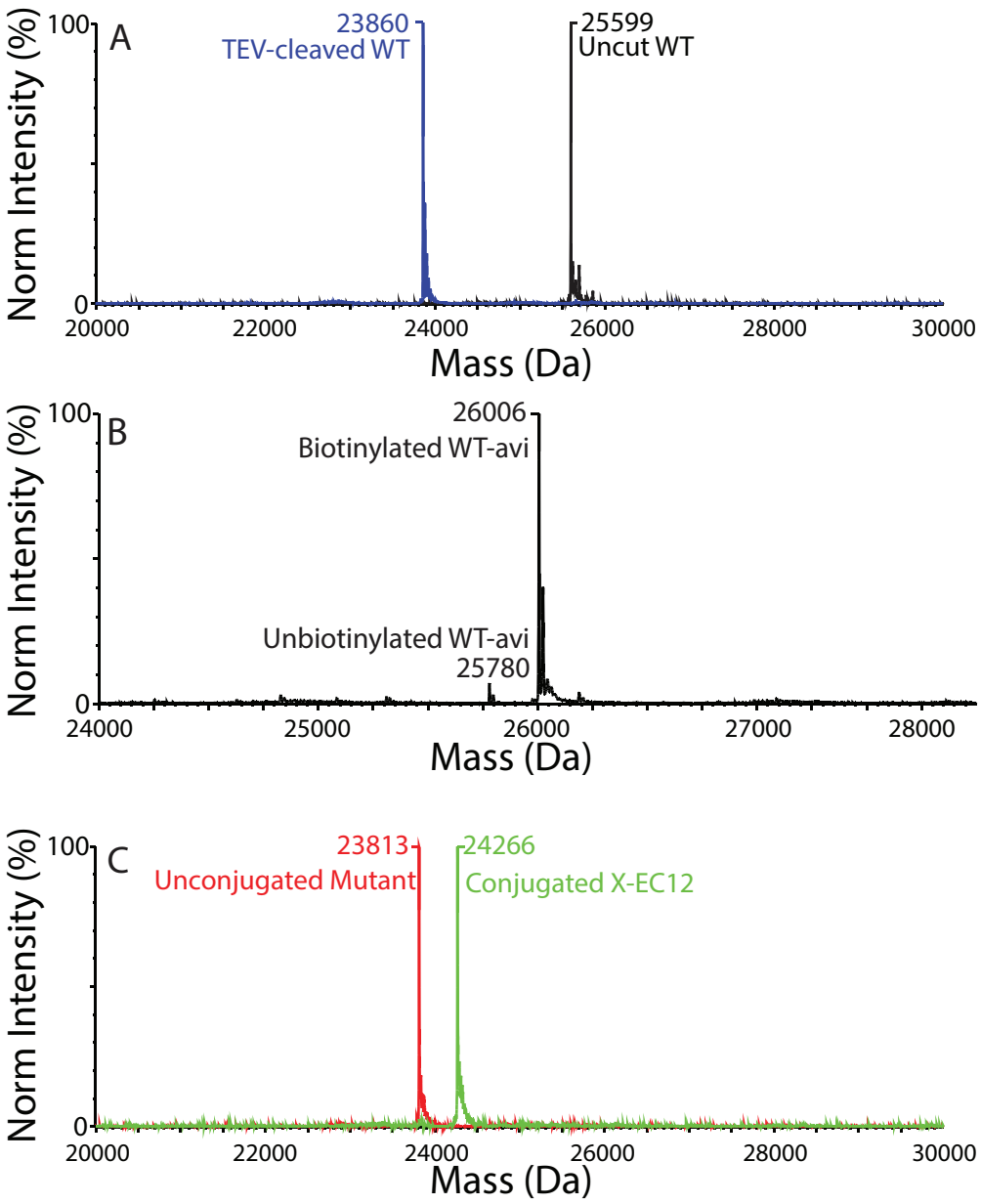

Figure 31 – Mass spectra showing steps during protein production. A and C are overlays of two spectra. (A) Demonstration of mass change after cleavage. Black, uncut WT C9A protein; blue, WT C9A protein cleaved with TEV protease. Expected monoisotopic mass change was Δ1747 Da, compared to Δ1739 Da observed. (The difference could be due to imperfect protonation state prediction, such as for the 6xHis tag.) (B) Verification of Biotinylation. The majority of the protein is biotinylated. Expected and observed mass changes were both Δ226 Da. (C) Demonstration of mass change during conjugation, for mutant 129/138: red, unconjugated mutant protein; green, conjugated mutant protein. Expected and observed mass changes were both Δ453 Da.

Final Purification and Storage

(Note: for poorly conjugated proteins, residual unconjugated protein can be removed using the protocol in **Appendix C** before final purification)

Prior to analysis, proteins were injected into a HiLoad 16/600 Superdex 75 column equilibrated in storage buffer (25 mM Tris-HCl pH 7.5, 150 mM NaCl, 500 μM TCEP) to remove any aggregates that may have developed during cleavage, repurification or conjugation. Typical monomer yields were 90% for unconjugated and 60% for conjugated protein. Proteins were stable in storage buffer for several weeks at 4 °C. For longer term storage, proteins were flash frozen in storage buffer containing 10% v/v glycerol using liquid nitrogen, then stored at -80 °C. (*Note: Conjugates had a high non-specific affinity for agarose and other common support matrices. Attempts to conjugate protein prior to cleavage and repurification led to protein aggregation. Attempts to separate TEV from conjugated EC12 using other column methods such as protease-affinity columns were similarly unsuccessful. Conjugated protein remains monomeric in the Superdex 75 and HiPrep Desalting columns used here.)*

Appendix B: Protein Concentration Determination

Unconjugated protein concentrations were determined using measurements of absorbance at 280 nm using extinction coefficients predicted by the ExPASY online protein parameter tool[29], which were ε_{280}=21430 M^{-1} cm^{-1} for EC12, and ε_{280}=26430 M^{-1} cm^{-1} for EC12-avitag. For conjugated proteins, the extinction coefficient at 280 nm was computed as the sum of the predicted protein and measured free BSBCA (ε_{280}=10100 M^{-1} cm^{-1}) extinction coefficients. BSBCA's extinction coefficient was computed by measuring BSBCA's absorbance at 370 nm and 280 nm and using

$$\varepsilon_{280} = \frac{A_{280}}{A_{370}} * \varepsilon_{370}$$

Equation 6 – Determination of BSBCA 280 nm extinction coefficient

where ε_{370}= 29000 M^{-1} cm^{-1} (see Zhang et al.[59]). BSBCA's molar extinction coefficients can vary when conjugated to protein. However, similar band intensities were observed when conjugated and unconjugated proteins of identical computed protein concentration were loaded into an SDS-PAGE gel and then stained with coomassie (data not shown).

Appendix C: Separation of Unconjugated Protein from Conjugates using Biotinylation/Streptavidin Purification

Introduction

The inspiration for the development of this method came from discussions with Dr. Matthew Banghart. I am grateful for his considered thoughts and assistance.

Prior to optimizing the conjugation conditions for 129/138 (**Chapter 5**), and while conjugating and characterizing the other 10 mutants (**Chapter 4**), I used an additional set of purification steps to remove residual unconjugated protein. This protocol involves labeling any residual free cysteines with biotin-PEG-maleimide, followed by streptavidin capture of the biotinylated protein. After optimization, it was not necessary for 129/138 due to the near-completeness of the conjugation reaction; as I never tested the other mutants using the optimized protocol, it may not have been necessary for them either. However, I include this protocol both for completeness and because it may be useful in the creation of other conjugated photoswitchable proteins where no conditions where the reaction completes can be found. One caveat of this method is that the conjugates have a tendency to aggregate in the presence of the agarose beads used to introduce the streptavidin; this effect could be avoided if a suitable replacement could be found, though this was not tested.

Method

The following steps should be inserted into the optimized purification protocol after conjugation, but prior to final purification and storage.

After conjugation, excess chromophore was removed *via* desalting using a HiPrep 26/10 desalting column (GE #17-5087-01) into TEV cut buffer, modified to pH 6.9. Proteins were concentrated to a concentration of 200 μM using spin concentrators (Millipore #UFC901024), and, afterward, EZ-Link Maleimide-PEG2-Biotin (Thermo #21901) was added to a final concentration of 2 mM. Protein solutions were placed at 4 °C in the dark overnight.

The following morning, excess biotin was removed via desalting into TEV cut buffer, pH 7.5. Protein solutions were again concentrated to 200 μM, followed by addition of pre-washed and equilibrated High Capacity Steptavidin Agarose Resin (Thermo #20359) at a ratio of 1 mL of packed resin per 3 mL of protein solution. (Note: the stated binding capacity of the resin in 10 mg of biotinylated protein per mL of resin. Adjust the ratio of resin to one appropriate for the expected mass of biotinylated protein based on conjugatability of the particular protein used). Protein-resin mixtures were nutated at 4 °C in the dark overnight. The following morning, resin and unconjugated proteins were separated from conjugates *via* a gravity-flow chromatography column (BioRad #732-1010).

The completeness of the biotinylation and streptavidin capture reactions can be monitored via mass spectrometry, *via* the appearance and disappearance of a +1050 Da peak, respectively (**Figure 32**– Note: masses for 129/138 are different from other figures because these data show an earlier version of the construct without the TEV cleavage site and with the addition of a C-terminal 6xHis tag).

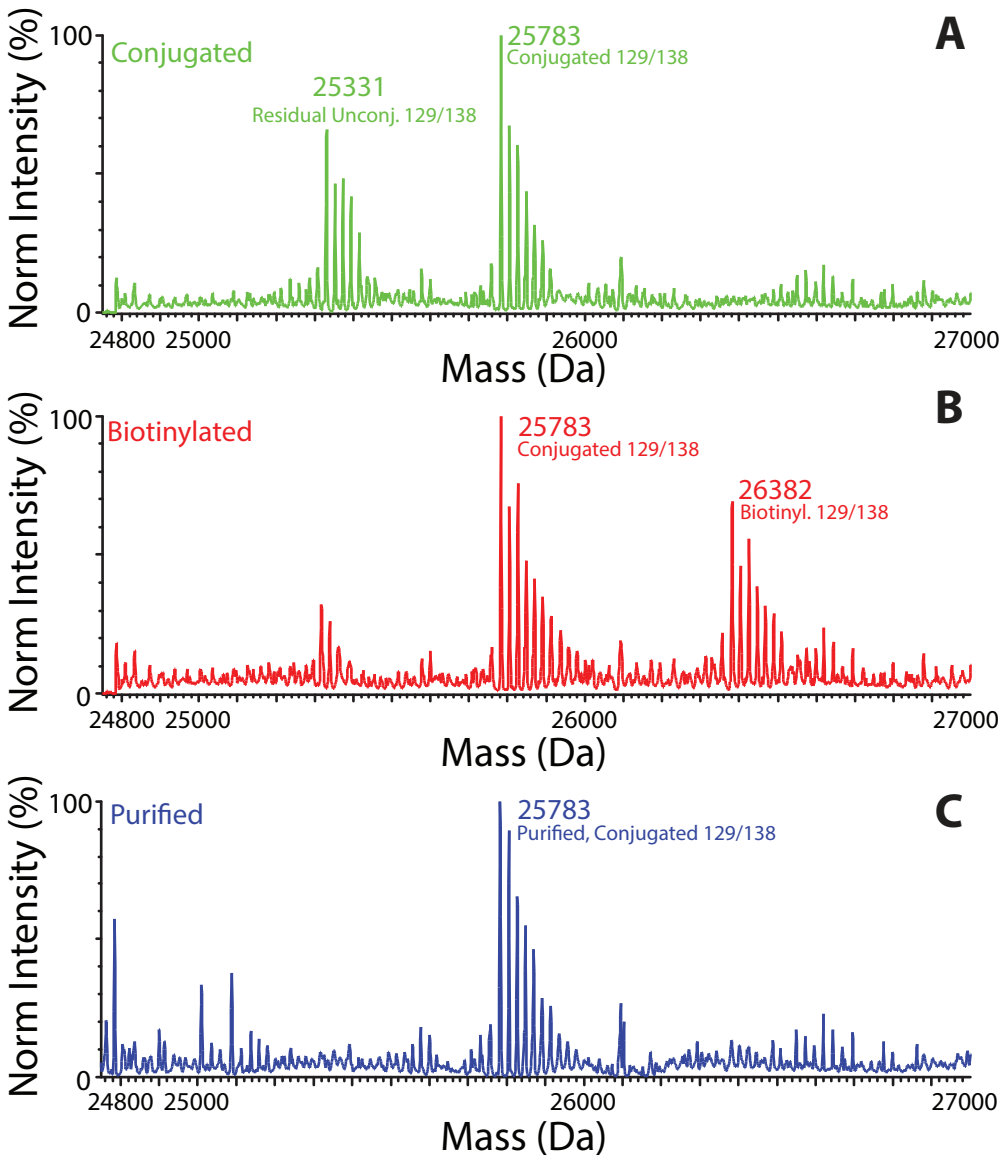

Figure 32 – Mass spectrometry analysis of biotinylation and streptavidin capture of conjugated protein. (A) Protein after conjugation, showing both conjugated and residual unconjugated protein. (B) Conjugated protein after biotinylation. The unconjugated protein has been biotinylated with 2 biotin-PEG_2-maleimides, (+1050 Da, 2x +525 Da). (C) The biotinylated sample after streptavidin purification. The biotinylated fraction has disappeared, leaving pure conjugated protein. Masses for 129/138 are different from other figures as this protein is an older version of the construct without a TEV cleavage site and containing a C-terminal 6xHis tag.

Appendix D: Size Exclusion Chromatography of *cis* and *trans* 129/138

I used size exclusion chromatography to show a decrease in 129/138 homodimerization after illumination. Protein was diluted to a final concentration of 250 μM in TBS (25mM Tris pH 7.5, 150mM NaCl) at room temperature. For *trans* 129/138, calcium was added to a final concentration of 2.5 mM and left to equilibrate for 5 minutes at room temperature, then 25 μL of protein was injected into a Superdex 75 PC 3.3/30 column (GE # 17-0771-01) equilibrated in TBS and attached to a 1200 Series HPLC (Agilent) at a flow rate of 100 μL/min. *Cis* 129/138 was illuminated with a 1W UV LED (emission maximum 365 nm) for 6 minutes prior to the addition of calcium, after which it was treated the same as *trans*. Protein was detected by monitoring absorbance at 280 nm. I observed the expected decrease in the fraction of protein forming dimers after illumination, indicating that the *cis* 129/138 homodimerization is weaker than that of *trans* 129/138. Representative traces are shown in **Figure 33**.

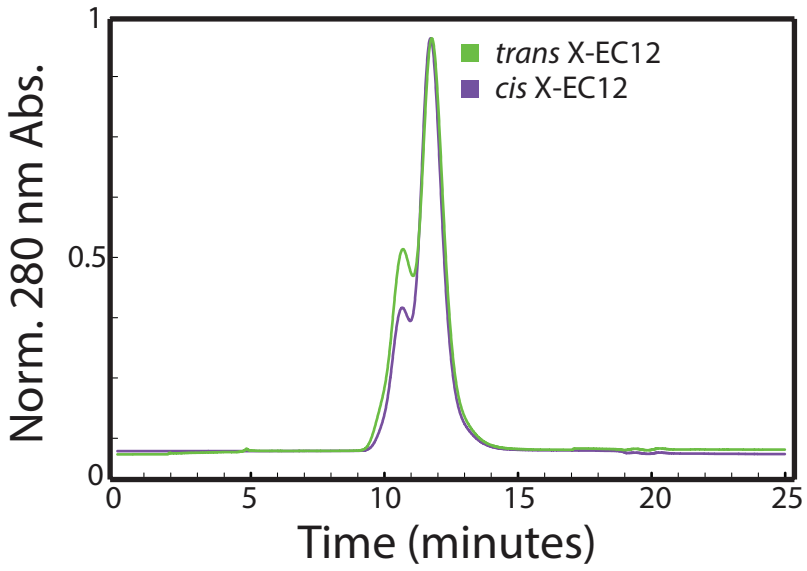

Figure 33 – Size exclusion chromatography of *cis* and *trans* 129/138.

References

1 Krauss, U., Drepper, T. & Jaeger, K. E. Enlightened enzymes: strategies to create novel photoresponsive proteins. *Chemistry* **17**, 2552-2560, doi:10.1002/chem.201002716 (2011).

2 Banghart, M., Borges, K., Isacoff, E., Trauner, D. & Kramer, R. H. Light-activated ion channels for remote control of neuronal firing. *Nature Neuroscience* **7**, 1381-1386, doi:10.1038/nn1356 (2004).

3 Caamaño, A., Vázquez, M., Martínez-Costas, J., Castedo, L. & Mascareñas, J. A Light-Modulated Sequence-Specific DNA-Binding Peptide. *Angewandte Chemie (International Edition)* **39**, 3104-3107 (2000).

4 Leung, D. W., Otomo, C., Chory, J. & Rosen, M. K. Genetically encoded photoswitching of actin assembly through the Cdc42-WASP-Arp2/3 complex pathway. *Proceedings of the National Academy of Sciences, USA* **105**, 12797-12802, doi:10.1073/pnas.0801232105 (2008).

5 Wu, Y. I. *et al.* A genetically encoded photoactivatable Rac controls the motility of living cells. *Nature* **461**, 104-108, doi:10.1038/nature08241 (2009).

6 Levskaya, A., Weiner, O. D., Lim, W. A. & Voigt, C. A. Spatiotemporal control of cell signalling using a light-switchable protein interaction. *Nature* **461**, 997-1001, doi:10.1038/nature08446 (2009).

7 Woolley, G. A. *et al.* Reversible photocontrol of DNA binding by a designed GCN4-bZIP protein. *Biochemistry* **45**, 6075-6084, doi:10.1021/bi060142r (2006).

8 Zhang, F., Muller, K. M., Woolley, G. A. & Arndt, K. M. Light-controlled gene switches in mammalian cells. *Methods in molecular biology* **813**, 195-210, doi:10.1007/978-1-61779-412-4_12 (2012).

9 Wyart, C. *et al.* Optogenetic dissection of a behavioural module in the vertebrate spinal cord. *Nature* **461**, 407-410, doi:10.1038/nature08323 (2009).

10 Beharry, A. A. & Woolley, G. A. Azobenzene photoswitches for biomolecules. *Chemical Society reviews* **40**, 4422-4437, doi:10.1039/c1cs15023e (2011).

11 Ivanov, D. B., Philippova, M. P. & Tkachuk, V. A. Structure and functions of classical cadherins. *Biokhimiya/Biochemistry* **66**, 1174-1186 (2001).

12 Koch, A., Pokutta, S., Lustig, A. & Engel, J. Calcium binding and homoassociation of E-cadherin domains. *Biochemistry* **36**, 7697-7705 (1997).

13 Prakasam, A., Chien, Y., Maruthamuthu, V. & Leckband, D. Calcium site mutations in cadherin: impact on adhesion and evidence of cooperativity. *Biochemistry* **45**, 6930-6939 (2006).

14 van Roy, F. & Berx, G. The cell-cell adhesion molecule E-cadherin. *Cellular and Molecular Life Sciences* **65**, 3756-3788, doi:10.1007/s00018-008-8281-1 (2008).

15 Pokutta, S., Herrenknecht, K., Kemler, R. & Engel, J. Conformational changes of the recombinant extracellular domain of E-cadherin upon calcium binding. *European Journal of Biochemistry* **223**, 1019-1026 (1994).

16 Capaldo, C. T. & Macara, I. G. Depletion of E-cadherin disrupts establishment but not maintenance of cell junctions in Madin-Darby canine kidney epithelial cells. *Molecular Biology of the Cell* **18**, 189-200, doi:10.1091/mbc.E06-05-0471 (2007).

17 Ohsugi, M., Larue, L., Schwarz, H. & Kemler, R. Cell-junctional and cytoskeletal organization in mouse blastocysts lacking E-cadherin. *Developmental biology* **185**, 261-271, doi:10.1006/dbio.1997.8560 (1997).
18 Nose, A., Nagafuchi, A. & Takeichi, M. Expressed recombinant cadherins mediate cell sorting in model systems. *Cell* **54**, 993-1001 (1988).
19 Flint, D. G., Kumita, J. R., Smart, O. S. & Woolley, G. A. Using an azobenzene cross-linker to either increase or decrease peptide helix content upon trans-to-cis photoisomerization. *Chemistry & Biology* **9**, 391-397 (2002).
20 Guerrero, L., Smart, O. S., Woolley, G. A. & Allemann, R. K. Photocontrol of DNA binding specificity of a miniature engrailed homeodomain. *Journal of the American Chemical Society* **127**, 15624-15629, doi:10.1021/ja0550428 (2005).
21 Guerrero, L. *et al.* Photochemical regulation of DNA-binding specificity of MyoD. *Angewandte Chemie* **44**, 7778-7782, doi:10.1002/anie.200502666 (2005).
22 Courjean, O. *et al.* Modulation of E-cadherin monomer folding by cooperative binding of calcium ions. *Biochemistry* **47**, 2339-2349, doi:10.1021/bi701340d (2008).
23 Burns, D. C., Zhang, F. & Woolley, G. A. Synthesis of 3,3'-bis(sulfonato)-4,4'-bis(chloroacetamido)azobenzene and cysteine cross-linking for photo-control of protein conformation and activity. *Nature Protocols* **2**, 251-258, doi:10.1038/nprot.2007.21 (2007).
24 Harrison, O. J. *et al.* Two-step adhesive binding by classical cadherins. *Nature Structural & Molecular Biology*, doi:10.1038/nsmb.1784 (2010).
25 Nose, A., Tsuji, K. & Takeichi, M. Localization of specificity determining sites in cadherin cell adhesion molecules. *Cell* **61**, 147-155 (1990).
26 Troyanovsky, R. B., Sokolov, E. & Troyanovsky, S. M. Adhesive and lateral E-cadherin dimers are mediated by the same interface. *Molecular and Cellular Biology* **23**, 7965-7972 (2003).
27 Grant, G. A. Synthetic peptides for production of antibodies that recognize intact proteins. *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]* **Chapter 18**, Unit 18 13, doi:10.1002/0471140864.ps1803s28 (2002).
28 *The Qiaexpressionist - A handbook for high-level expression and purification of 6xHis-tagged proteins*. June 2003 edn, (Qiagen, 2003).
29 Gasteiger, E. *et al.* ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic acids research* **31**, 3784-3788 (2003).
30 Earle, W. R., Sanford, K. K., Evans, V. J., Waltz, H. K. & Shannon, J. E., Jr. The influence of inoculum size on proliferation in tissue cultures. *Journal of the National Cancer Institute* **12**, 133-153 (1951).
31 Dulbecco, R. & Freeman, G. Plaque production by the polyoma virus. *Virology* **8**, 396-397 (1959).
32 Stanford University Functional Genomics Facility. *Mouse Exonic Evidence Based Oligonucleotide (MEEBO) Arrays*, http://www.microarray.org/sfgf/meebo.do (Accessed July 24th, 2013).
33 *Microarrays.org Protocols*, http://derisilab.ucsf.edu/microarray/protocols.html (Accessed July 24th, 2013).
34 Yamamoto, K. (University of California, San Francisco, San Francisco, CA) Personal Communication. (2008)

35 Trauner, D. (University of California, Berkeley, Berkeley, CA) Personal Communication. (2008)
36 Rosetta Commons. *Rosetta - The premier software suite for macrommolecular modeling*, http://www.rosettacommons.org (Accessed May 16th, 2013).
37 Kortemme, T. & Baker, D. A simple physical model for binding energy hot spots in protein-protein complexes. *Proceedings of the National Academy of Sciences, USA* **99**, 14116-14121, doi:10.1073/pnas.202485799 (2002).
38 Kortemme, T., Kim, D. & Baker, D. Computational Alanine Scanning of Protein-Protein Interfaces. *Science Signaling* **2004**, pl2, doi:10.1126/stke.2192004pl2 (2004).
39 Schrodinger, LLC. *The PyMOL Molecular Graphics System, Version 1.5.0.5 Enhanced for Mac OS X* (2013).
40 Leckband, D. & Prakasam, A. Mechanism and dynamics of cadherin adhesion. *Annu Rev Biomed Eng* **8**, 259-287, doi:10.1146/annurev.bioeng.8.061505.095753 (2006).
41 Woolley, G. A. (University of Toronto, Toronto, ON) Personal Communication. (2012)
42 Harrison, O. J., Corps, E. M. & Kilshaw, P. J. Cadherin adhesion depends on a salt bridge at the N-terminus. *Journal of Cell Science* **118**, 4123-4130, doi:10.1242/jcs.02539 (2005).
43 Invitrogen Corporation. *AcTEV Protease Protocol*, http://tools.invitrogen.com/content/sfs/manuals/12575.pdf (Accessed July 26th, 2013).
44 LaVallie, E. R., McCoy, J. M., Smith, D. B. & Riggs, P. Enzymatic and chemical cleavage of fusion proteins. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.]* **Chapter 16**, Unit16 14B, doi:10.1002/0471142727.mb1604bs28 (2001).
45 Troyanovsky, R. B., Laur, O. & Troyanovsky, S. M. Stable and unstable cadherin dimers: mechanisms of formation and roles in cell adhesion. *Molecular Biology of the Cell* **18**, 4343-4352, doi:10.1091/mbc.E07-01-0084 (2007).
46 Häussinger, D. *et al.* Proteolytic E-cadherin activation followed by solution NMR and X-ray crystallography. *EMBO J* **23**, 1699-1708, doi:10.1038/sj.emboj.7600192 (2004).
47 Vendome, J. *et al.* Molecular design principles underlying beta-strand swapping in the adhesive dimerization of cadherins. *Nature Structural & Molecular Biology* **18**, 693-700, doi:10.1038/nsmb.2051 (2011).
48 Weiss, J. N. The Hill equation revisited: uses and misuses. *FASEB Journal* **11**, 835-841 (1997).
49 Peschke, M., Verkerk, U. H. & Kebarle, P. Features of the ESI mechanism that affect the observation of multiply charged noncovalent protein complexes and the determination of the association constant by the titration method. *Journal of the American Society for Mass Spectrometry* **15**, 1424-1434, doi:10.1016/j.jasms.2004.05.005 (2004).
50 Samanta, S. & Woolley, G. A. Bis-azobenzene crosslinkers for photocontrol of peptide structure. *Chembiochem : a European journal of chemical biology* **12**, 1712-1723, doi:10.1002/cbic.201100204 (2011).
51 Borghi, N., Lowndes, M., Maruthamuthu, V., Gardel, M. L. & Nelson, W. J. Regulation of cell motile behavior by crosstalk between cadherin- and integrin-

mediated adhesions. *Proceedings of the National Academy of Sciences, USA* **107**, 13324-13329, doi:10.1073/pnas.1002662107 (2010).

52 Kim, N. G., Koh, E., Chen, X. & Gumbiner, B. M. E-cadherin mediates contact inhibition of proliferation through Hippo signaling-pathway components. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 11930-11935, doi:10.1073/pnas.1103345108 (2011).

53 Nagaoka, M. *et al.* E-cadherin-coated plates maintain pluripotent ES cells without colony formation. *PLoS ONE* **1**, e15, doi:10.1371/journal.pone.0000015 (2006).

54 Xu, Y. *et al.* Revealing a core signaling regulatory mechanism for pluripotent stem cell survival and self-renewal by small molecules. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 8129-8134, doi:10.1073/pnas.1002024107 (2010).

55 Blanco-Lomas, M., Samanta, S., Campos, P. J., Woolley, G. A. & Sampedro, D. Reversible photo-control of peptide conformation with a rhodopsin-like photoswitch. *Journal of the American Chemical Society*, 120406140828000, doi:10.1021/ja301868p (2012).

56 Sadovski, O., Beharry, A. A., Zhang, F. & Woolley, G. A. Spectral tuning of azobenzene photoswitches for biological applications. *Angewandte Chemie (International Edition)* **48**, 1484-1486, doi:10.1002/anie.200805013 (2009).

57 Samanta, S., Qin, C., Lough, A. J. & Woolley, G. A. Bidirectional photocontrol of peptide conformation with a bridged azobenzene derivative. *Angewandte Chemie* **51**, 6452-6455, doi:10.1002/anie.201202383 (2012).

58 Beckett, D., Kovaleva, E. & Schatz, P. J. A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Science* **8**, 921-929, doi:10.1110/ps.8.4.921 (1999).

59 Zhang, Z., Burns, D., Kumita, J., Smart, O. & Woolley, G. A water-soluble azobenzene cross-linker for photocontrol of peptide conformation. *Bioconjugate Chemistry* (2003).

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

August 20,2013

Author Signature | Date