University of Tennessee, Knoxville
TRACE: Tennessee Research and Creative Exchange

Doctoral Dissertations | Graduate School

12-2013

A Tale of Two Protein Switches: Engineering, Characterizing, and Understanding a Novel and a Natural Switch

James Vincent Price
University of Tennessee - Knoxville, jprice41@utk.edu

Follow this and additional works at: https://trace.tennessee.edu/utk_graddiss

 Part of the Biochemical and Biomolecular Engineering Commons

Recommended Citation

Price, James Vincent, "A Tale of Two Protein Switches: Engineering, Characterizing, and Understanding a Novel and a Natural Switch. " PhD diss., University of Tennessee, 2013.
https://trace.tennessee.edu/utk_graddiss/2608

This Dissertation is brought to you for free and open access by the Graduate School at TRACE: Tennessee Research and Creative Exchange. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of TRACE: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.

To the Graduate Council:

I am submitting herewith a dissertation written by James Vincent Price entitled "A Tale of Two Protein Switches: Engineering, Characterizing, and Understanding a Novel and a Natural Switch." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Chemical Engineering.

Eric T. Boder, Major Professor

We have read this dissertation and recommend its acceptance:

Paul D. Frymier, Paul M. Dalhaimer, Cynthia B. Peterson

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

A Tale of Two Protein Switches: Engineering, Characterizing and Understanding a Novel and a Natural Switch

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

James Vincent Price

December 2013

Copyright © 2013 by James Vincent Price.

All rights rerserved.

This dissertation is dedicated to my parents, Roger and Kathy Price, for all of their support and love during this journey and throughout my entire life.

Acknowledgements

I would first like to thank my advisor, Dr. Eric Boder, for giving me the opportunity to work on this project. I have gained a wealth of experience working with two very different proteins and expression systems, which I have no doubt will serve me very well the rest of my career. Eric's patience, guidance and mentorship have been invaluable and were a key component to my successful completion of this degree. I will always be very grateful for his efforts in developing me as a scientist.

I would also like to thank Dr. Paul Dalhaimer who has served as both a committee member and mentor. Paul has been wonderful providing support when I got discouraged or helping me try to figure out my next steps after finishing my PhD.

My success wouldn't have happened without the help of my labmates. First and foremost I need to thank Morgan Baltz. Morgan started working with me early in her undergraduate career and her eagerness to learn and help me explore the hemagglutinin mutations was so incredibly helpful. She performed the vast majority of the pH titration curves for the hemagglutinin mutants allowing me to focus my efforts on engineering the I domain/EF hands switch. Maryam Raeeszadeh and I started in the lab together at the same time. Our friendship has grown over the years and her support has been tremendous. She's has been a daily presence in the lab allowing me to make sure I wasn't designing flawed experiments or to empathize with about failed experiments and the general frustrations that come with research. I would also like to thank Nathan Carberry and Chris Barnes who have also helped with some of the experiments in this dissertation.

Finally, I would like to thank my girlfriend, Emily Sterchi. Emily, your love, support and friendship mean the world to me. Thank you for being patient during the late nights and long hours and for helping me relax and have fun away from the lab so I could come back recharged and ready to tackle the science that awaited.

Abstract

Throughout nature, many proteins provide a specific function in response to some input signal (e.g., phosyphorylation, pH, etc.), a process that is oftentimes described as switching a protein “on” or “off.” The advent of protein engineering has allowed for the creation and understanding of chimeric proteins for uses in a number of applications such as therapeutics, biosensors and energy production. Two proteins, hemagglutinin (HA) of fowl plague influenza virus and a chimeric protein comprising a fusion between the LFA-1 I domain and the EF3 and EF4 hands of calmodulin, have been investigated in this dissertation. Both of these proteins undergo structural rearrangements in response to an input signal in a process that could be described as molecular switching.

Hemagglutinin (HA) is a viral fusion protein that undergoes an irreversible conformational change upon acidification to catalyze the fusion of endosomal and viral membranes. Directed evolution was previously used to identify novel mutants with activation pH across a range of 4.8 – 6.0; wild-type HA activates around pH 5.2. Examination of library mutants has enabled us to identify individual amino acids responsible for phenotypic changes and thus provides additional structure-function insight highlighting the importance of HA1 headgroups in pH sensing.

Secondly, we have improved a previously created chimeric protein switch using the integrin LFA-1 I domain and the EF3 and EF4 hands of calmodulin; the original switch protein had a 2.4-fold increase for the I domain ligand, ICAM-1, in the presence of a calmodulin target peptide. Directed evolution has revealed a mutant switch with an additional twenty-fold increase in ICAM-1 binding in the presence of peptide. Additionally, the specificity of the chimeric protein for the original peptide sequence has been engineered, creating a new switch that responds to a different peptide sequence rendering the original peptide nonfunctional. Both proteins, hemagglutinin and the I

domain-calmodulin fusion, represent the importance of conformational changes in controlling protein function. The ability to manipulate these changes underscores the power of protein engineering and directed evolution and can lead to the development of complex biomolecules that can be used in a variety of applications.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

Chapter 1 : Introduction

1.1 Background

Nature is beautifully complex with a multitude of processes, proteins, and interactions to inspire new ideas for the creation of technologies to fight diseases and to better understand the mechanisms behind them. Characterizing, understanding, and engineering different proteins can lead to the development of smarter materials to be used for different applications such as targeted drug delivery, biosensing, and molecular imaging.

The unique sequence of amino acids that defines a specific protein heavily influences the three-dimensional structure, or shape, of the protein. Myoglobin, an iron- and oxygen-binding protein found in many animal cells, was the first protein to have its three-dimensional structure solved in 1958.[1] Since 1958, the number of protein structures that have been determined has grown exponentially such that the Protein Data Bank now contains close to 88,000 protein structures (www.rcsb.org). It is this three-dimensional structure that is key in determining the function of a protein. Antibodies have a distinct Y-shape, consisting of two variable regions and an Fc domain, which is critical to the immunological function of the protein. Variable domains, which contain the antigen binding site, have extreme sequence diversity allowing antibodies to bind foreign antigens with great specificity. Conversely, the Fc domain is composed of highly conserved sequences and has receptor binding cites for engagement with cells of the immune system.[2] Other proteins, such as hemoglobin can have flexible structures that provide the basis for their function.[3; 4] As one oxygen molecule binds to a heme in one of the hemoglobin subunits, strains in the protein caused by the location of the hemes are relaxed allowing the protein to change its shape to allow for better positioning of the hemes and thus allowing oxygen to bind with higher affinity.[5] It is the intrinsic ability of proteins like hemoglobin to change, or switch,

their shape that led to the understanding of allostery and serves as a fundamental principle in this dissertation.

1.1.1 Allostery

Protein structure and conformational flexibility play a critical role in controlling protein function. A number of different types of proteins exploit structural rearrangements to carry out their respective functions such as enzymes, chaperonins, membrane fusion proteins and intergrins.[6; 7; 8; 9] The conformational changes of a protein often occur in response to effector molecules (metabolites, ATP, DNA, RNA, peptides, etc.) that bind at sites distinct from the active site resulting in a reversible conformational change in the three dimensional structure of the protein, a mechanism known as allostery.[10] In the seminal paper first describing protein allostery, Monod and colleagues describe a number of enzymes that are allosterically regulated, such as ATCase and threonine-deaminase.[10] GroEL, a molecular chaperone that assists in the folding of nonnative polypeptides, performs an intricate array of allosteric movements choreographed by cooperative ATP binding. [7] Integrins, which will be discussed in more detail below, also undergo drastic conformational refolding in response to intracellular signals in order to adopt a high affinity conformation and assist in leukocyte firm adhesion.[9; 11] The integrin, lymphocyte function-associated antigen-1 (LFA-1, integrin αLβ2) is able to firmly adhere to sites of inflammation by high affinity binding through its inserted (I) domain. The affinity of αL I domain for its ligand, ICAM-1, is allosterically regulated in response to the reversible 'pulling' of its α7 helix which causes a series of rearrangements within the domain resulting in a high affinity state.

However, not all conformational rearrangements are reversible. For instance, viral fusion proteins, such as HIV-1 GP120/41 and hemagglutinin (HA), expressed on the surface of enveloped viruses mediate the merging of two biological membranes and undergo irreversible conformational changes in response to ligand binding or lowered

pH, respectively.[12] The viral fusion protein, gp120/41, is expressed on the surface of HIV-1 and effects entry of the virus into the cell. It is structurally similar to influenza hemagglutinin, which will be discussed in detail below, as both proteins are processed as fusion incompetent polypeptides (HIV-1 GP160 and influenza HA0) and are subsequently proteolytically cleaved into a fusion competent molecule. Upon exposure to an activating stimulus (CD4 binding for gp120/41 and low pH for HA) the fusion proteins undergo an irreversible conformational switch to an extended, stable coiled coil conformation exposing their fusion proteins (N-termini of gp41 and HA2, respectively) to the target membrane.[12; 13] Insertion of the fusion peptides followed by refolding of the fusion proteins brings the two membranes in close proximity to each other allowing them to merge.

Allosteric rearrangements and large-scale conformational refolding can be thought of as a protein switching mechanism. In other words, due to an external input a protein can be forced to switch conformations. This dissertation will explore this concept of protein switching by characterizing the viral fusion protein, hemagglutinin, and exploiting the concept of allostery, using the I domain of the LFA-1 integrin and the EF3 and EF4 hands of calmodulin to engineer a chimeric fusion protein switch. In the case of αL I domain, conformational switching will be shown to be induced by the binding of an exogenous peptide and hemagglutinin mutations will be identified that point to a hotspot within the protein structure that likely controls its switching mechanism.

1.1.2 Protein switches

Different design motifs have been used to engineer a number of different protein switches. Common design motifs for engineering protein switches include protein splitting, domain insertion, mutually exclusive folding and domain swapping.[14; 15] Most switches rely on using enzymes or reporter proteins, such as luciferase, GFP, β-

lactamase, or dihydrofolate reductase (DHFR) as the output domain to provide an easily readable signal upon activation by the input domain.[14; 15] An early example of an engineered switch involved domain insertion of calmodulin into a circularly permuted GFP. Upon the addition of calcium, the GFP-calmodulin switch demonstrated about an 8-fold enhancement in fluorescence.[16] β-lactamase (BLA) has been routinely used in protein switch construction as hydrolysis of beta-lactam antibiotics makes screening mutants very efficient. Domain insertions with BLA to control substrate hydrolysis include the maltose binding protein (MBP), calmodulin (CaM), and a short hexapeptide to engineer a metal ion binding site.[17; 18; 19] Mutually exclusive folding is another design motif in engineering protein switches where the folding of one protein causes the unfolding of its fusion partner. A fusion where human ubiquitin is inserted into an exposed loop of *B. amyloliquefaciens* barnase showed that by controlling barnase folding via ubiquitin, cell viability could be regulated.[20] Regardless of the design method chosen, controlling protein function by an input domain provides an additional level of regulation.

1.1.3 Possible applications of engineered switches

A large number of devastating illnesses are either virally induced (e.g. HIV/AIDS, T-cell leukemia)[21] or are associated with chronic inflammation (e.g. rheumatoid arthritis, Crohn's disease, psoriasis)[22]. Both proteins that will be studied in this thesis are either heavily involved in inflammatory responses - I domain - or from the same family of proteins responsible for seasonal flues - hemagglutinin - that also share many structural and functional properties with retroviral proteins such as the HIV fusion protein gp120/41. Therefore, these proteins serve as convenient starting points for the development of target specific molecules.

Site-specific imaging or drug delivery is a complicated problem with a number of factors such as particle size, circulation half-life, immunogenicity, and target specificity

that need to be considered.[23; 24] Recently, the αL I domain has been conjugated to the surface of nanoparticles for targeted delivery of drug and gene therapies to cells over expressing ICAM-1 both *in vitro* and *in vivo*.[25; 26; 27] These molecules use previously engineered high affinity I domain[28] to selectively target ICAM-1 making these nanoparticles highly efficient at targeting tumors where drug delivery by the permeability and retention (EPR) effect is less efficient. Designing a protein switch where I domain affinity is regulated by the presence of exogenous peptide would provide an additional layer of specificity in targeted imaging and delivery systems. Further engineering of peptide specificity would enhance the targeting capabilities.

Characterization and understanding of hemagglutinin and what leads to its conformational switching could lend insight into similar proteins' (e.g., HIV-1 gp120/41) modes of action therefore leading to more effective retroviral drugs. The fusion peptide of HA2 has also been shown to be effective in promoting fusion between a liposome and cellular membranes as an efficient means of delivering genetic material to a cell.[29; 30] A better understanding of HA allows for possible engineering of ligand binding sites into locations of the hemagglutinin that are critical in pH sensing. Grafting in a ligand binding site in a pH sensing 'hot spot' could allow for controlled release of drugs encapsulated in liposomes by the presence or absence of a ligand, providing more control over current approaches where only the fusion peptide of hemagglutinin is used to assist in liposome-membrane fusion. One concern about an approach of this nature would be immune responses due to the presence of a viral protein.

Development of a protein switch, along with the ability to alter ligand specificity, would demonstrate the profound ability to build a biomolecule that could do almost anything; possibilities include selectively delivering a drug to a specific target or developing a biosensor for security or safety purposes. An engineered protein switch could be used to benefit society in a nearly unlimited number of ways.

Chapter 2 : Characterization of hemagglutinin mutants for identification of novel pH sensing mutations

2.1 Introduction

Membrane fusion is a highly evolved process that accounts for many biological activities including cell-cell fusion (e.g. sperm and oocyte), extra- and intra-cellular fusion of pathogens with host cells (enveloped viruses), or intracellular fusion of organelles.[8] The process by which the influenza viral membrane fuses with a cellular endosomal membrane, via the viral fusion protein hemagglutinin (HA), to release its viral RNA within the host cell is a highly studied, model fusion event.[8; 12; 31; 32] Following receptor binding, internalization of the virus by endocytosis leaves the viral genome separated from the cell cytoplasm by two membranes, endosomal and viral. Fusion of these two membranes is energetically unfavorable without the additional energy released by HA upon its irreversible, conformational change in response to the lowered pH of the acidified endosome.[33; 34]

2.1.1 Hemagglutinin

The typical membrane fusion pathway involves four intermediate states that include 1) the close contact of the opposing membranes, 2) destabilization of the membranes, 3) a hemifusion state and 4) finally the complete fusion pore [8; 33]. Many mechanisms of action can lead to membrane fusion but a well characterized pathway involves fusion proteins, such as hemagglutinin (HA) (Figure 2. 1). A better understanding of membrane fusion, and in particular, the action of hemagglutinin can lead to the design of better smart molecules for such practices as targeted drug or gene delivery and to the design of novel viral inhibitors.

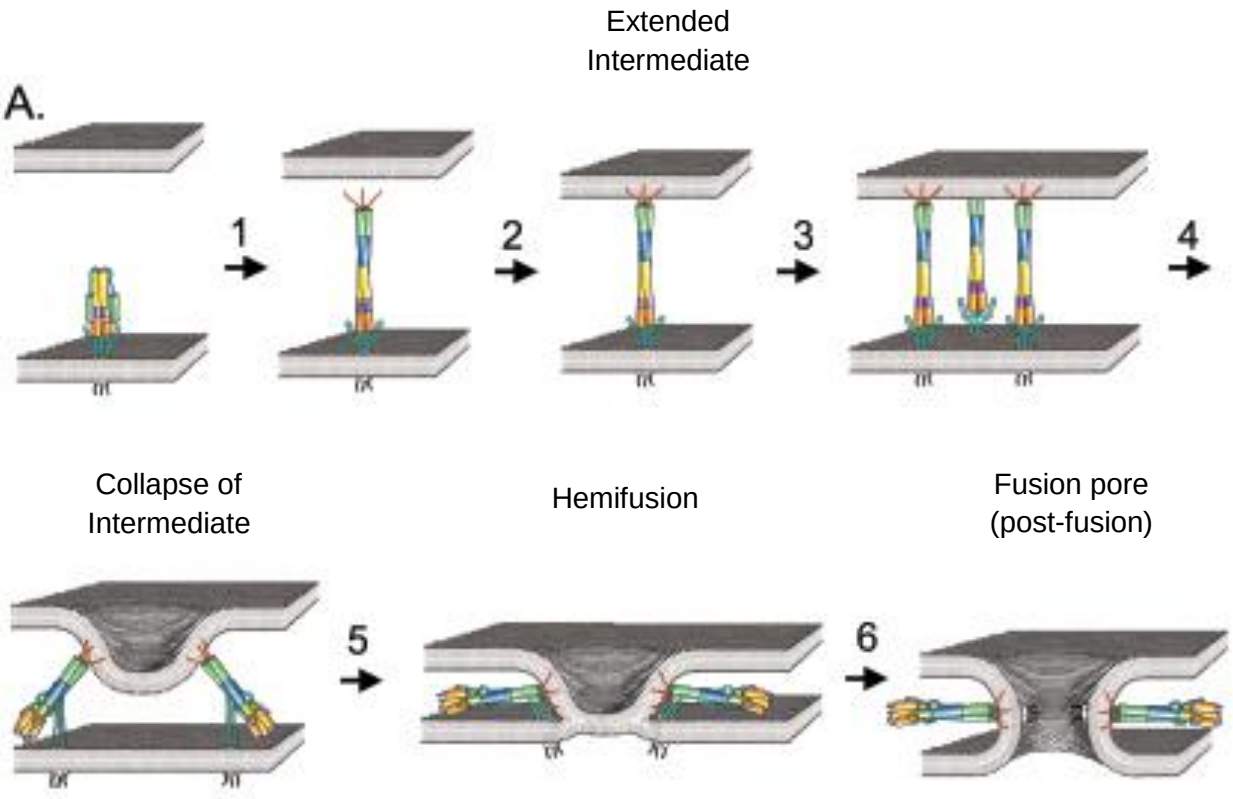


Figure 2. 1 Proposed role of HA in membrane fusion. HA1 is not shown for clarity. At acidic pH, HA refolds to expose its fusion peptide to the target membrane (1). Trimers aggregate quickly, followed by a second refolding step at the C-terminal end of HA2 where the kink region (purple) undergoes a helix-to-loop transition (3 and 4). The refolding of the C-terminal end brings the two apposing membranes into close proximity allowing for the formation of a transitive hemifusion state (5). The transmembrane domains and fusion peptides of HA are brought together allowing for pore formation (6). Figure 2.1 from Gruenke, J.A., Armstrong, R.T., Newcomb, W.W., *et. al*.: New Insights into the Spring-Loaded Conformational Change Influenza Virus Hemagglutinin. J. Virol. (2002), 76(9): 4456-4466 originally published by BioMed Central.[35]

Hemagglutinin is a membrane glycoprotein composed of three identical heterodimer subunits, HA1 and HA2 (Figure 2.2 A). The HA1 subunit is approximately 330 amino acids in length and extends 96 Å from its amino-terminal end near the viral membrane distally into a globular head domain composed of an 8-stranded β-sheet structure, which contains the binding pocket for sialic acid.[36] The carboxy-terminal region of the polypeptide extends 60Å back toward the viral membrane anti-parallel to the amino-terminal region of HA1.[36] The HA2 subunit is approximately 220 amino acids composed of a short alpha helix (29Å) connected by an extended chain to a long α-helix (76Å) followed by a transmembrane domain and short cytoplasmic tail.[36; 37] The extended loop of HA2 connecting the two helices, referred to as the B-loop, undergoes a loop-to-helix transition to extend the coiled-coil structure of the protein which relocates the fusion peptide more than 100 Å from its buried position in the neutral pH conformation towards the target membrane(Figure 2.2 A).[12; 38; 39] Each subunit is synthesized as a single polypeptide chain (HA0) and trimerizes in the endoplasmic reticulum.[40] The trimers are immediately transported to the Golgi complex where the HA0 subunits are cleaved by either a subtilisin-like (polyarginine H5 and H7 subtypes) or trypsin-like (all other subtypes are monoarginine) protease either in the *trans*-Golgi or in the plasma membrane region[39; 41]. Cleavage of HA0 into the disulfide-linked HA1/HA2 subunits results in the newly formed N-terminal residues of HA2, known as the fusion peptide, burying into a cavity of ionizable residues of the trimeric coiled-coil (Figure 2.2 B).[32; 42] The ectodomain of HA is stabilized by a number of intramolecular interactions consisting of intra-domain and -subunit disulfide bonds, polar and hydrophobic interactions between adjacent α-helices,[36] and a number of single and complex salt bridges.[43]

Hemagglutinin binds to sialyted cell-surface receptors where the enveloped virus is endocytosed. The lowered pH of the endosome causes an irreversible conformational change in the hemagglutinin structure leading to fusion of the viral and

Figure 2.2 Influenza hemagglutinin. A) Trimer of hemagglutinin monomers showing three major structures on the pathway towards low pH refolding where the HA1 subunit is shaded blue, HA2 is red and the N-terminal HA2 fusion peptide is yellow. Left panel and (B), HA is synthesized as a single polypeptide chain HA0 before proteolytic cleavage resulting in a disulfide-linked HA1/HA2. Middle panel, following cleavage of HA0 the newly formed fusion peptide imbeds itself into the trimeric coiled-coil. Right panel, under acidic conditions the protein undergoes an irreversible conformational change that results in the dissociation of HA1 headgroups (not shown) and relocates the fusion peptide more 100 Å towards the target membrane. B) Left and middle panels, rearrangement of the fusion peptide. The black sphere represents the N-terminal glycine of the newly formed HA2 subunit. Right panel, the kink angle of the fusion peptide which is essential for insertion into the target membrane as it allows the more hydrophobic N-terminal half of the peptide to associate with the hydrophobic bilayer while the C-terminal half is more exposed to a hydrophilic environment. Figure 2.2 A and B reprinted from Cross, K.J., Langley, W.A., Russell, R.J., Skehel, J.J., and D.A. Steinhauer. Compositions and Functions of the Influenza Fusion Peptide. Protein and Peptide Letters (2009), 16: 766-778, with permission from Bentham Science Publishers.[44]

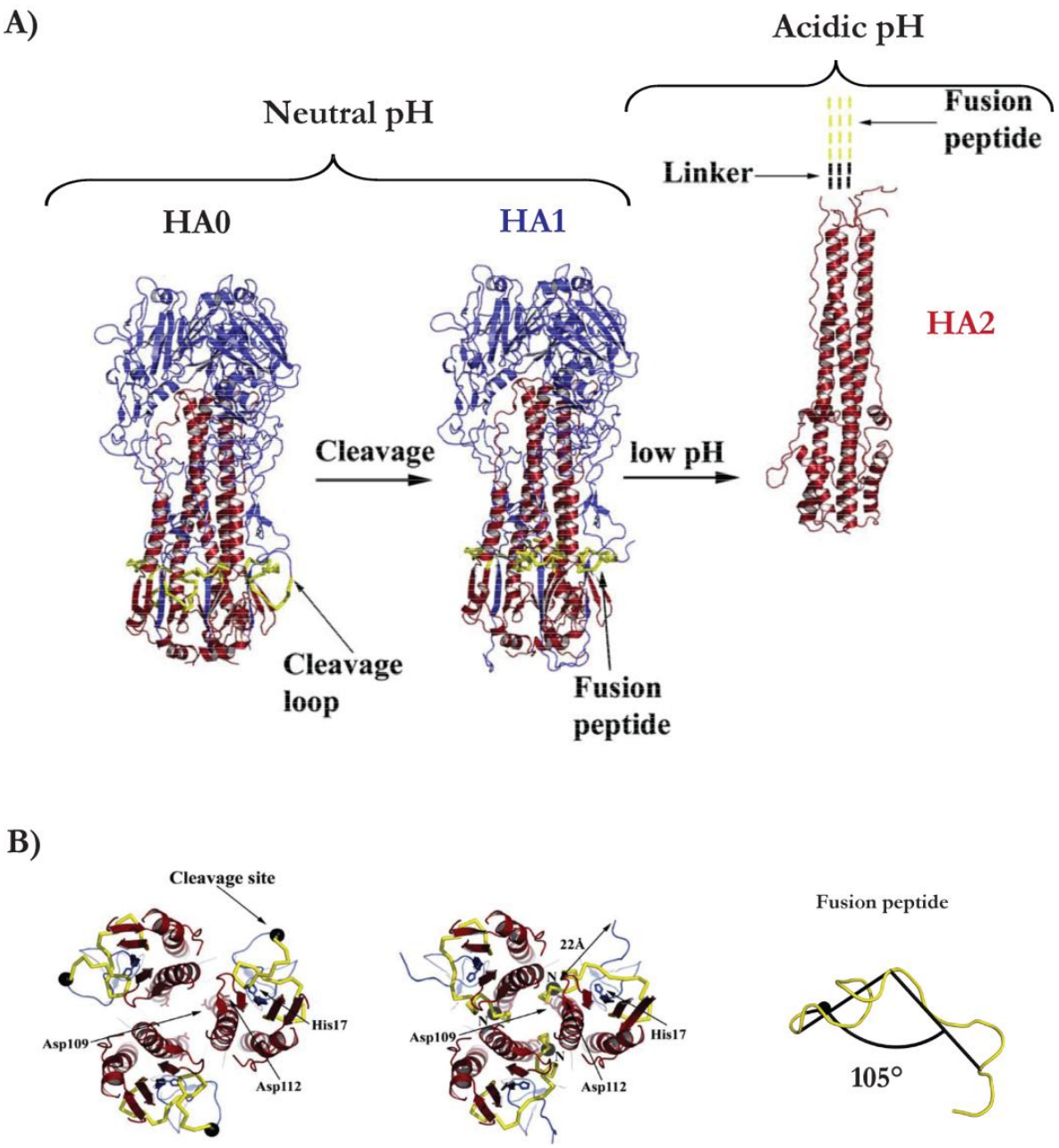


Figure 2.2 Continued

endosomal membranes. Protonation of charged residues in both the HA1 head group[45] and fusion peptide pocket[46; 47] are disrupted resulting in the HA1 head groups dissociating and allowing the fusion peptide to be exposed to the target membrane. The fusion peptide is composed of two short helices, the N-terminal helix consisting of conserved hydrophobic and glycine residues and the C-terminal half that is amphipathic with a number of charged and polar residues, that adopts a kink angle in the center of the helix of approximately 105° (Figure 2.2 B).[48; 31; 49; 50] The mechanism of membrane fusion mediated by the viral fusion peptide remains unclear (i.e., whether the fusion peptide inserts into the endosomal membrane or the viral membrane, reviewed in ref. [8]). Most current structural data suggest that the fusion peptide is extruded from its hydrophobic cavity in the coiled coil structure and extended towards the target membrane (e.g., the endosomal membrane) where it inserts. Rearrangements occur at the membrane proximal end of HA2 causing the C-terminal part of the helix to break and fold back upon itself bringing the fusion peptide and transmembrane domains into close proximity, and thus the apposed membranes[44; 46; 51]. The membrane defects imposed by the fusion peptide and the negative curvature of the membrane during the refolding of HA are predicted to lower the energetic barrier to allow fusion to proceed (Figure 2. **1**).

While the precise mechanism of hemagglutinin-mediated membrane fusion is not explicitly understood, there is agreement in the literature as to some necessary steps on the pathway toward membrane fusion. First, at least eight HA trimers rapidly aggregate at the site of fusion in a non-rate limiting step, of which only two or three are needed to refold to promote fusion.[52; 53] White and colleagues demonstrated that viral membrane fusion is a cooperative process that involves the concerted action of at least three timers, which has also been verified through single-particle kinetic experiments.[54; 55] This cooperativity is likely responsible for an autocatalytic activation phenomenon employed by HA. It is well understood that hemagglutinin aggregates at the site of

membrane fusion and that this clustering often occurs in “lipid raft” microdomains.[56] These clusters of HA coincide neatly with the idea of synchronized activation proposed by Markovic and colleagues. The authors suggest that adjacent HA proteins at high local densities will undergo a synchronized conformational change to effectively lower the energy barrier separating the metastable conformation from the final low-energy conformation, ensuring that the virus will efficiently carry out the membrane fusion process.[57] Likewise, Lee, Goulian and Boder also described a cell-scale autocatalysis mechanism where only two states, activated or inactivated HA, exist. Data show that freshly expressed protein can be activated by the presence of previously refolded HA implying a feedback autocatalysis mechanism.[58]

2.1.2 Spring-loaded model of HA

The structural transitions that lead to the final irreversible HA conformation have been well characterized with crystal structures depicting each of the important intermediate structures. As mentioned earlier, HA is processed as a single polypeptide, HA0, which is thermodynamically stable and unable to undergo the irreversible conformational change at low pH. Proteolytic cleavage of HA0 results in disulfide-linked HA1 and HA2 subunits and the repositioning of the 20 N-terminal amino acids of HA2, the fusion peptide, into a hydrophobic cavity within the HA trimer. This structure is considered to be trapped in a metastable state at neutral pH and readily refolds to a lower energy conformation at low pH or at neutral pH with the addition of a denaturant.[59] The lowest energy conformation of HA occurs at the pH of fusion where the B-loop has undergone a loop-to-helix transition resulting in a HA trimer that is stabilized by an extended coiled coil structure. The transition from the metastable neutral pH structure to the stable low pH structure is supported by the spring-loaded model for HA refolding.[38; 59]

The spring-loaded model for hemagglutinin refolding is as follows:

- Proteolytic cleavage of HA0 results in a metastable structure of HA that is primed for refolding
- HA1 headgroups act as clamps preventing the B-loop of HA2 to undergo its loop-to-helix transition
- Dissociation of HA1 headgroups likely precedes the loop-to-helix transition
- The loop-to-helix transition results in the formation of an extended coiled coil structure, a lower energy conformation

Before the low pH structure of HA was solved, the residues composing the B-loop of HA2 were shown to adopt a helical structure at low pH in solution.[38] This indicated that the B-loop of HA2 prefers a helical conformation but is blocked from refolding by the HA1 headgroups. Studies have expanded upon the importance of the B-loop transition by introducing mutations to stabilize this loop and prevent HA refolding into an extended coiled coil structure. Mutation of H2 HA2 R106H stabilizes the trimer well below the pH of fusion for wild-type H2 HA.[60] Crystal structures of this mutant at neutral and low pH indicate that side chain interactions between the loop and the long α helix of HA2 are lessened at low pH, evidence of a possible early intermediate in the spring-loaded mechanism.[60] A double proline HA mutant, F63P/F70P (H3 HA2 numbering), resulted in a fusion incompetent protein whose coiled coil structure was interrupted by the proline substitutions following irreversible low pH refolding.[35] It was argued that while the double mutant refolded at low pH, the proline mutations effectively shortened the coiled coil structure preventing the necessary tension needed for membrane fusion to occur indicating that the loop-to-helix transition is a necessary component of the spring-loaded model.[35] Intersubunit disulfides engineered to prevent the HA1 headgroups from dissociating at low pH abolished HA refolding and membrane fusion indicating the likely necessity of HA1 headgroup dissociation in HA refolding.[61; 62] Finally, targeted molecular dynamics using neutral and low pH crystal structures (1HGF

and 1HTM, respectively) indicated that the spring-loaded conformational mechanism is the most energetically favorable pathway between the structures.[63]

A number of studies have examined mutations throughout the structure of HA to identify pH sensing mutations and their effect on HA refolding.[64; 65] These studies genereally relied upon sequence analysis and/or structural information to identify residues or parts of the structure to target for identification of pH sensing residues. Previous work in our lab developed a flow cytometry assay that decouples HA refolding from membrane fusion.[58] This assay was then used in a directed evolution experiment to generate HA mutants whose structures are either stabilized or destabilized with respected to wild-type H7 HA. We have analyzed the different mutations that came from this previous directed evolution work in order to try and identify individual mutations that are responsible for pH sensing and to gain further insight into the structure-function relationship of HA. Analysis of individual mutations mapped the majority of the changes to the HA1 domain, and more specifically to the interface of two HA1 subunits. While there is agreement in the literature that some loosening of the HA1 domains precedes the large structural rearrangements leading to the fusion competent structure of HA, our data suggest that HA1 plays a more significant role in controlling HA refolding.

2.2 Materials and Methods

2.2.1 DNA, library and point-mutant construction

Kozak sequence GCCACC was added to the wild-type FPV HA gene[51] (from A/FPV/Rostock/34) by PCR in order to increase the expression in mammalian cells. Amino acid residues HA2 11-20 (i.e. fusion peptide residues 11–20) were replaced with the c-myc epitope tag sequence via PCR, and referred to as HA3.1, as previously described.[58] HA3.1 DNA was randomly mutated by non-biased error-prone PCR using Mutazyme II polymerase (Stratagene, Santa Clara, CA). PCR products were purified by electroelution and digested with Xho I and Not I. They were then ligated into pLNCX2

retroviral expression vector (Clontech, Mountain View, CA). *E.coli* transformation was done by electroporation and the cells were cultured in 250 ml LB Amp medium in a 37 °C shaker for 36 hours. The library size was determined by plating 20 ul of the culture after 36 hours of incubation and counting the number of colonies on the plate after a 20 hour incubation at 37 °C. The number of colonies was then multiplied by 20,000. Point mutants were created by site-directed mutagenesis using the QuikChange II XL kit (Stratagene) following the manufacturer's protocol. Primers were created with a minimum of 15 nucleotides of homology on either side of the desired amino acid change. Point mutants were confirmed by sequencing prior to retroviral transduction of the point mutant to create a stably-expressing HA with the mutation.

2.2.2 Cell culture

All mammalian cell cultures were maintained at 37°C and 5% CO_2. NIH-3T3 cells (a gift from Dr. John Crocker, University of Pennsylvania), EcoPack2-293 and AmphoPack-293 (Clontech, Mountain View, CA) were grown in Dulbecco's Modified Eagles Medium (DMEM) (Hyclone, Logan, UT) supplemented with 10% (v/v) fetal bovine serum (Hyclone), 100 U/ml penicillin and 100 ug/ml streptomycin (Life Technologies, Grand Island, NY). EcoPack2-293 and AmphoPack-293 are HEK293 derived cell lines stably expressing viral structural genes gag, pol and env for safe production of infectious, replication-incompetent virions. Media was replaced roughly every 48 hours and cultures were passaged prior to reaching confluency (~4 days). During passaging, cells were washed with Dulbecco's phosphate buffered saline (DPBS) containing 8 g/L NaCl, 200 mg/L KCl, 1150 mg/L Na_2HPO_4, 200 mg/L KH_2PO_4. The cell monolayers were then detached using either 0.25% trypsin-EDTA (routine passage) or 0.05% trypsin-EDTA (for analysis).

2.2.3 Retroviral expression and viral titer determination

1×10^6 EcoPack 2-293 (BD Biosciences, Mountain View, CA) packaging cells were plated in each well in a 6-well plate 16-20 hours before transfection, and transfected with 10 µg of retroviral expression vector library. Virus containing medium was collected at either 24 hours or 48 hours post-transfection and filtered through a 0.45 µm cellulose acetate filter. NIH-3T3 cells were plated 24 hours prior to infection in 6-well plates at a density of 1×10^5 cells per well. In order to determine the viral titer, six 10-fold serial dilutions of the virus containing medium was prepared, and NIH-3T3 cells were infected by adding 1 mL of the diluted medium to the wells with the final polybrene concentration of 4 µg/ml. The medium was replaced with selective medium containing 100 µg/ml of G418 (Life Technologies) 24 hours post-infection. The viral titer was estimated by counting the number of colonies in the highest dilution that contained colonies after one week of selection and by multiplying the number of colonies by the dilution factor. For flow cytometry analysis of HA3.1 library expression in NIH-3T3 cells, cells were stained with HC2 mouse mAb and anti-mouse-phycoerythrin (PE) 24 or 48 hours post-infection..

2.2.4 Stable selection of HA library expressing cells

Cells were infected and selected by the method used for viral titer determination. Non-infected cells were used as a negative control and the cells were cultured in the selective medium until there were no viable cells observed in the control population. Expression levels before selection and after selection were analyzed by flow cytometry after HC2 mouse mAb and anti-mouse-PE staining.

2.2.5 Analysis of pH-induced activation:

Adherent, non-confluent HA3.1 or wt HA and single point mutant expressing HEK293 cells were detached using 0.05% trypsin-EDTA and cell suspension aliquots

were transferred to 1.5mL microcentrifuge tubes and placed on ice. The HA expressing cells were washed twice with PBS+0.1%BSA (8 g/L NaCl, 200 mg/L KCl, 1440 mg/L Na_2HPO_4, 240 mg/L KH_2PO_4, and 1 g/L BSA) at 500xg using a benchtop microcentrifuge. The cells were then pH-pulsed for 5 minutes at room temperature at the appropriate pH. After the pH pulse HA3.1 were double-stained with mouse monoclonal anti-FPV HA1 (HC2) and chicken anti-c-myc mAb. Wild-type HA or single mutant HA expressing cells were labeled with either anti-FPV HA1 (HC2) or the trimer and conformation specific mAb, HC58.[66] Primary antibody incubations were done at 1:100 dilution on ice for 30-60 minutes. Following the primary incubation, the cells were washed again with PBS+0.1% BSA and then labeled with a fluorescent secondary antibody at a 1:100 dilution and incubated on ice 30-60min. Following the secondary incubation, the cells were washed again in PBS and then analyzed by flow cytometry. Only viable cells, as determined by propidium iodide (BD Biosciences) exclusion (10 µL/10^6 cells for 15 minutes at room temperature) and/or light scattering, were included in the analysis by flow cytometry, and at least 10,000 gated events were analyzed for all flow cytometric experiments.

2.2.6 Determination of Transition pH

Viable cell populations were determined by light scattering and gated for analysis. The median fluorescence intensity (MFI) of the gated population was determined using FlowJo flow cytometry analysis software (Ashland, OR). Plots of MFI versus pH were created using the statistical software package Prism (Graphpad, La Jolla, CA) and a dose-response curve was fitted using the following equation:

$$Y = Bottom + \frac{(Top - Bottom)}{\left(1 + \left(EC50 / pH\right)^{HillCoefficient}\right)}$$

Where Y is the median fluorescence at a given pH, X is the pH, 'Bottom' is the lower plateau of the dose response curve, 'Top' is the upper plateau of the curve and EC50 is the pH at which HA is assumed to have transitioned to an irreversible conformation. An EC50 value was extracted from the dose-response curve fit and considered to be the "transition pH" at which HA refolds from a closed to an open conformation. Reported transition pH values are averages from independent experiments.

2.3 Results

2.3.1 Previously identified HA mutants with altered activation pH

Linking phenotypic changes to genetic changes creates the ability to probe the structure-function landscape and identify residues, or areas of the protein structure, that are involved in the conformational refolding of fowl-plague virus hemagglutinin (FPV HA). Previously, research in our lab developed a flow cytometry assay where HA refolding is decoupled from membrane fusion and can conveniently be quantified by immunofluorescent staining and flow cytometry (Figure 2.4).[58] The development of this flow cytometry assay for HA refolding allowed for the construction of three separate libraries by error-prone PCR and the ability to use fluorescence-activated cell sorting (FACS) to isolate HAs whose neutral pH conformations were either stabilized or destabilized with respect to wild-type hemagglutinin (Figure 2.5). Mutants from each of the three libraries were isolated and their genes sequenced to identify mutations that may or may not be responsible for pH sensing (Table 2.1). This chapter will analyze these previously identified mutations to try and identify individual mutations responsible for pH sensing and how these mutations likely affect the structure-function relationship of hemagglutinin.

In the previously developed assay, a set of monoclonal antibodies is used to assess HA refolding where two antibodies, HC58 and anti-c-myc (9e10) (when HA2

Figure 2.3 HA refolding decoupled from membrane fusion. A) HA expressing HEK293 cell monolayers are detached from culture plates using 0.05% Trypsin-EDTA. Cells are then transferred to 1.5 mL microcentrifuge tubes for washing and immunofluorescent labeling steps. B) Schematic of surfaced expressed HA refolding and immunofluorescent staining flowing 5 minute pH pulse. At neutral pH (left), HA is in its metastable neutral conformation with HC2 and HC58 epitopes available for antibody binding (upper right) while the c-myc epitope is not. Following the pH pulse and irreversible conformational change, HC2 and c-myc epitopes are available for immunofluorescent staining while HC58 is not (bottom left). C) Crystral structure of H7 HA showing locations of HC2, HC58 and c-myc eptiopes. D) Following immunofluorescent staining, cells are sorted (library) or analyzed (characterization) for refolding with a flow cytometer. For all HA depictions (B and C), only monomers are shown for clarity.

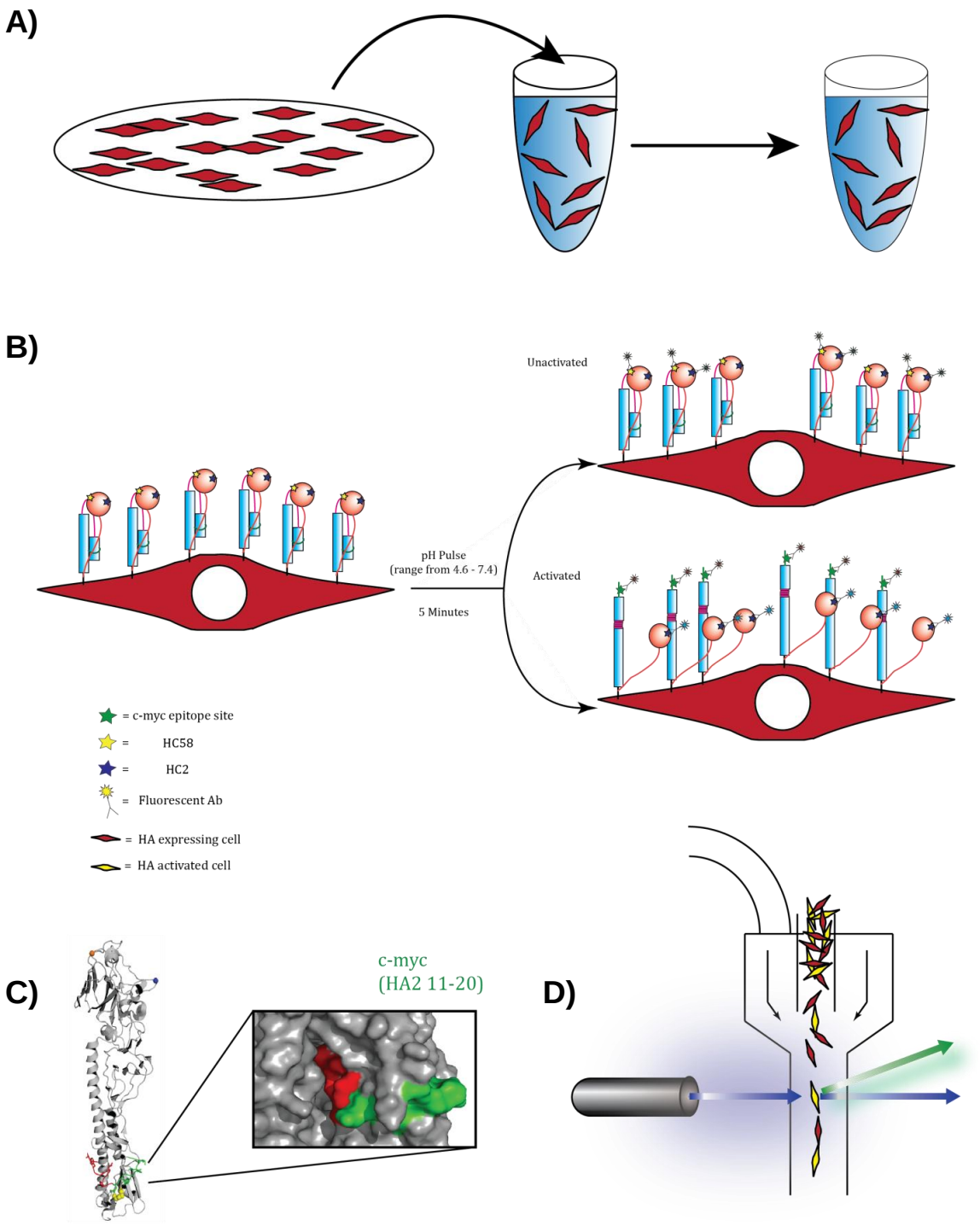


Figure 2.3 Continued

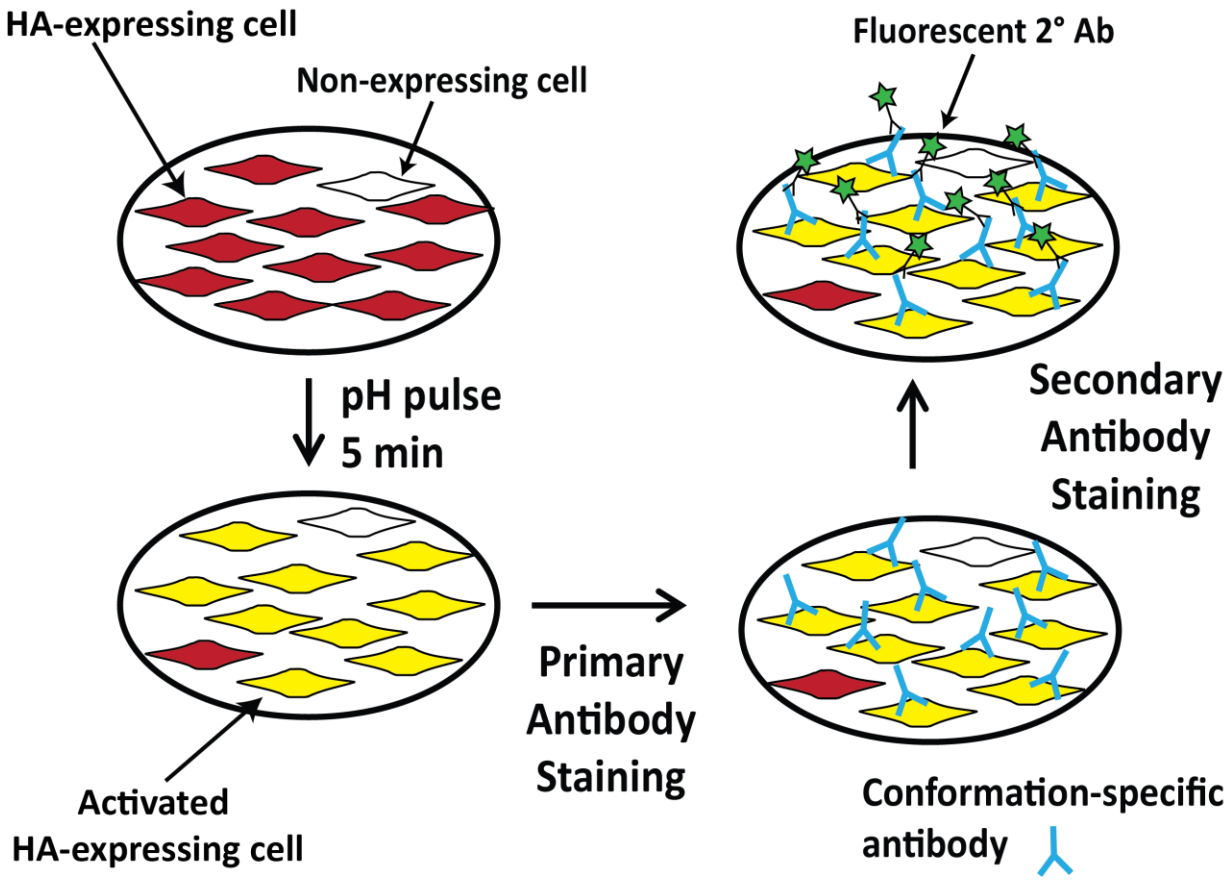


Figure 2.4 pH pulse experiment for HA refolding. HA-expressing cells (red, upper left) are subjected to a 5-minute pulse of pH buffer (pH 4.6 to 6.4 in increments of 0.2 pH units). Following the 5 minute pulse, cells are quenched with pH 7.4 buffer to stop any additional refolding. Cells are then labeled with a conformation specific mouse antibody (HC58, specific for the neutral pH conformation) for 45-60 minutes at 4°C. The labeled cells are then washed once and labeled with an anti-mouse second antibody for 45-60 minutes at 4°C. Cells are washed a final time and then analyzed by flow.

Figure 2.5 Previous directed evolution of HA. A) Design of three separate libraries to alter the conformational stability of HA. The top two library depictions illustrate that libraries were sorted following a higher pH pulse step to activate only destabilized HA. The bottom library design was more complicated. The library was initially pulsed at pH 5.2 to activate wild-type and destabilized HA. Thermolysin was then used to cleave the fusion peptide, eliminating the c-myc epitope signal. The library was then re-pulsed with pH 4.8 buffer to activate stabilized HA. B) Top panel represents the pH 5.6 library and the bottom panel represents the pH 6.0 library before and after 3 rounds of sorting. C) pH 4.8 library population following pH 5.2 pulse (right), thermolysin cleavage (middle), and pH 4.8 pulse (left). Data in parts B and C reprinted from Lee, Jeong H. "Engineering and Characterization of Influenza Hemagglutinin Properties in Mammalian System", PhD Dissertation, 2007.[67]

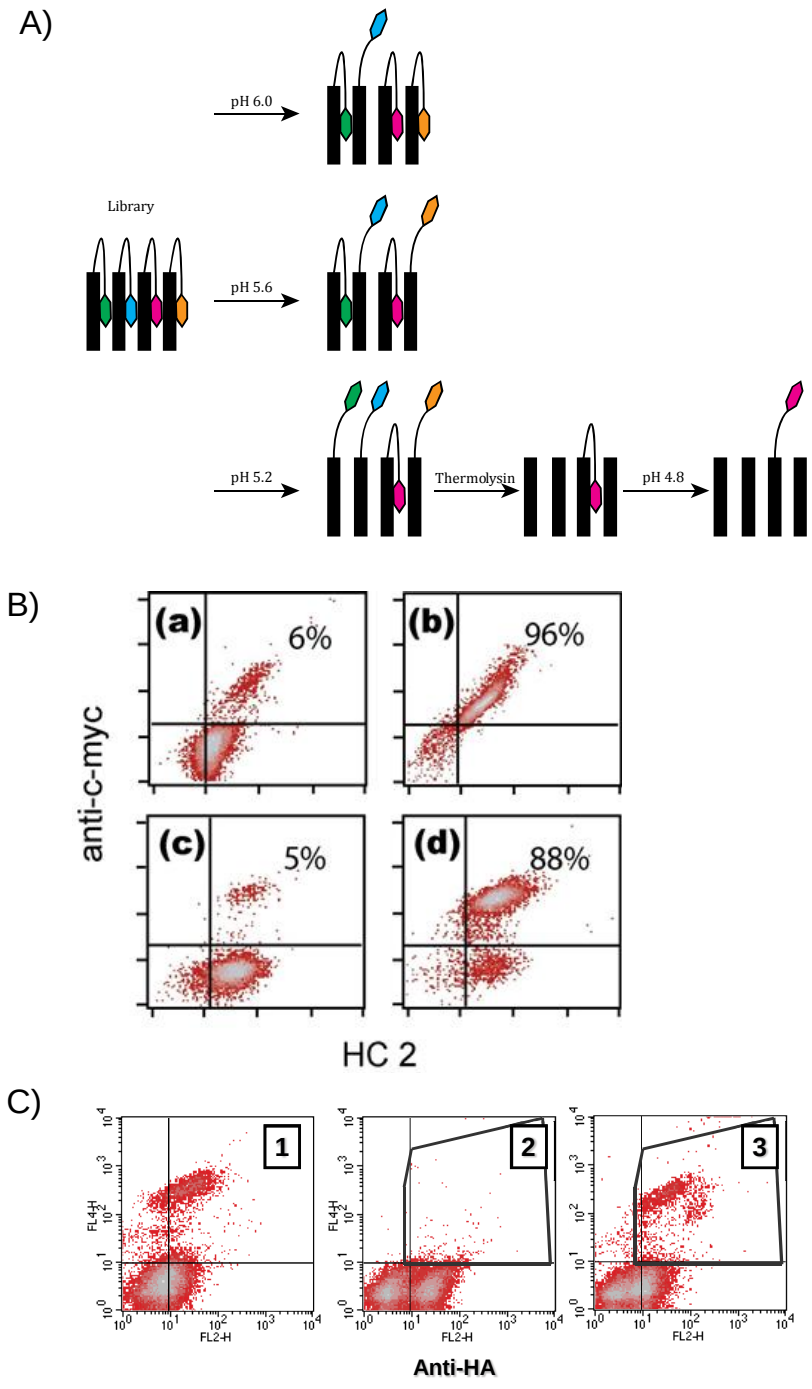


Figure 2.5 Continued

11-20 is replaced with the c-myc epitope) (Figure 2.3C), only detect the neutral and low pH conformations, respectively, while HC2 binds irrespective of HA conformation.[66] HA-expressing cells are first trypsinized from the culture dish using 0.05% Trypsin-EDTA and transferred to 1.5 mL microcentrifuge tubes for the washing and labeling steps. At this point, cells are kept as close to 4°C as possible. HA-expressing cells are washed twice with PBS+0.1% BSA. Following the second wash step, a 5 minute pH pulse is applied to induce HA refolding. Citrate buffers with different pHs ranging from 4.6 to 6.4 in increments of 0.2 units were used to subject the HA-expressing cells to a 5 minute pH pulse to induce HA refolding. Following the 5 minute pH pulse, cells were quenched with pH 7.4 PBS. After the pH pulse and quenching with pH 7.4 buffer, the cells were labeled with a conformation specific antibody. In the previously performed library selections, anti-c-myc (9e10) was used to identify exposure of the fusion peptide upon HA refolding. In this work, the conformation specific antibody, HC58, specific for the neutral HA conformation was used. Cells were incubated with primary antibodies for 45-60 minutes, then washed once and labeled with a fluorescent secondary antibody. Both 9e10 and HC58 are mouse monoclonal antibodies, so a anti-mouse secondary antibody was used to detect binding of the primary antibody to HA. Again, cells were incubated with the secondary antibodies 45-60 minutes on ice and then washed once to remove any unbound antibody. The immunofluorescently labeled cells were then run on a flow cytometer to analyze for HA refolding.

A number of studies using various biochemical assays have demonstrated that wild-type hemagglutinins are typically fully refolded by pH 5.2 and transition from a neutral conformation to the extended, low pH conformation around pH 5.4-5.5. Previous work was performed to try and isolate HAs that were destabilized compared to wt HA. Two HA libraries, HA3.2 and HA3.3, were created and exposed to a five minute pH pulse of either pH 5.6 or 6.0, respectively (Figure 2.5 A and B). A brief pulse at elevated pH allowed for any library members with destabilized HA to refold prior to

library members that may be expressing wt-like or stabilized HA. Three rounds of fluorescence-activated cell sorting (FACS) were used to enrich a population of HA that was destabilized with respect to wild-type HA. Following the third round of sorting, the two destabilizing libraries had been enriched to the point where 96% of HA3.2 and 88% of HA3.3 had been activated by their target pH (Figure 2.5 B). Mutants from these two libraries were isolated and mutations were identified by sequencing.

While there are numerous studies where investigators have either created, or found, destabilizing hemagglutinin mutations there are very few reports of mutations that stabilize HA in the presence of low pH.[46] Even though Thoennes et. al. report on a few mutations that appear to be stabilizing ($H17_1Y$ and $K51_2A$), the authors were not explicitly looking for stabilizing mutations but rather exploring the effect of ionizable residues around the fusion peptide to trigger HA refolding.[46] To identify mutations that had a stabilizing effect, the HA3.4 library was first pulsed at pH 5.2 to activate all destabilized and wild-type-like mutants. Refolded proteins with an exposed fusion peptide with the c-myc epitope were then susceptible to proteolytic cleavage by thermolysin,[58; 68] making their conformational change undetectable. Following the thermolysin digest with another pH pulse at pH 4.8 allows any stabilized mutants to be activated and expose their fusion peptide along with the c-myc epitope for fluorescent staining (Figure 2.5 A and C). Flow cytometry shows that after the first pulse at pH 5.2, a significant population is positive for c-myc; however, following cleavage with thermolysin, the population seen in the left panel of is no longer present in the middle panel. Finally, reapplying another pulse with pH 4.8 buffer activates another significant population presumably expressing stabilized HA. Library clones were isolated after two rounds of sorting and mutations contributing to a stabilized HA conformation were identified. This chapter will focus on the identifying mutations that are responsible for the altered phenotypes selected for in the previous directed evolution experiments that have just been described.

2.3.2 Identification of Destabilizing Mutations

To further understand which mutation(s) were responsible for destabilization of the hemagglutinin trimer, and to better understand what mechanism(s) drives HA refolding, individual mutations were analyzed. Five representative clones from the HA3.2 library and three clones from the HA3.3 library were previously isolated and sequenced following the three rounds of sorting to identify a set of mutations that may have contributed to the destabilization of the protein. From the sequenced clones of the mature HA3.2 library, a total of ten unique mutations were identified out of fourteen total mutations (Table 2.1). The three sequenced clones from the HA 3.3 library produced 3 unique mutations out of 7 total mutations. Interestingly, while both the HA3.2 and HA3.3 library were designed to identify destabilized HA, the two distinct libraries only shared one common mutation D166G.

The panel of mutations was thoroughly analyzed by creating hemagglutinin trimers with individual mutations using site-directed mutagenesis. To capture the behavior of wild-type HA, the c-myc epitope in the fusion peptide was not used to assess HA refolding, but rather the conformationally specific antibody, HC58. HC58 binds to HA in its neutral-pH conformation at interfaces between HA1 headgroups of two monomers; therefore as the headgroups dissociate upon HA refolding HC58 binding is abolished.[66] Flow cytometry was used to measure the median fluorescent intensity of the HC58 binding to HEK293 expressing cells following pH pulses in a range of pH around the expected pH of activation. The median fluorescent intensity (MFI) was plotted against pH and a dose-response curve was fitted to the plots (Figure 2.6). An EC50 value was extracted from the plots and used as the transition/activation pH for that mutant protein.

From the flow cytometry results two mutations each from the 5.6 and 6.0 pH libraries appear to destabilize HA. I96M and W243S, originating from the HA 3.2 library

Table 2.1 List of hemagglutinin mutants previously isolated by directed evolution for altered conformational stability

Clones	Activation pH	Mutations
HA3.4.1	4.8	D93S E462Q
HA3.4.2	4.8	Y155F A168V D240Y D247N
HA3.2.1	5.6	W243S
HA3.2.2	5.6	D166G W243S L407A
HA3.2.3	5.6	G142D D166G Y218H
HA3.2.5	5.6	I96M V188G W243S L407A S429P
HA3.3.1	6.0	D166G G365C
HA3.3.2	6.0	D166G G237R G365C
HA3.3.3	6.0	G237R G365C

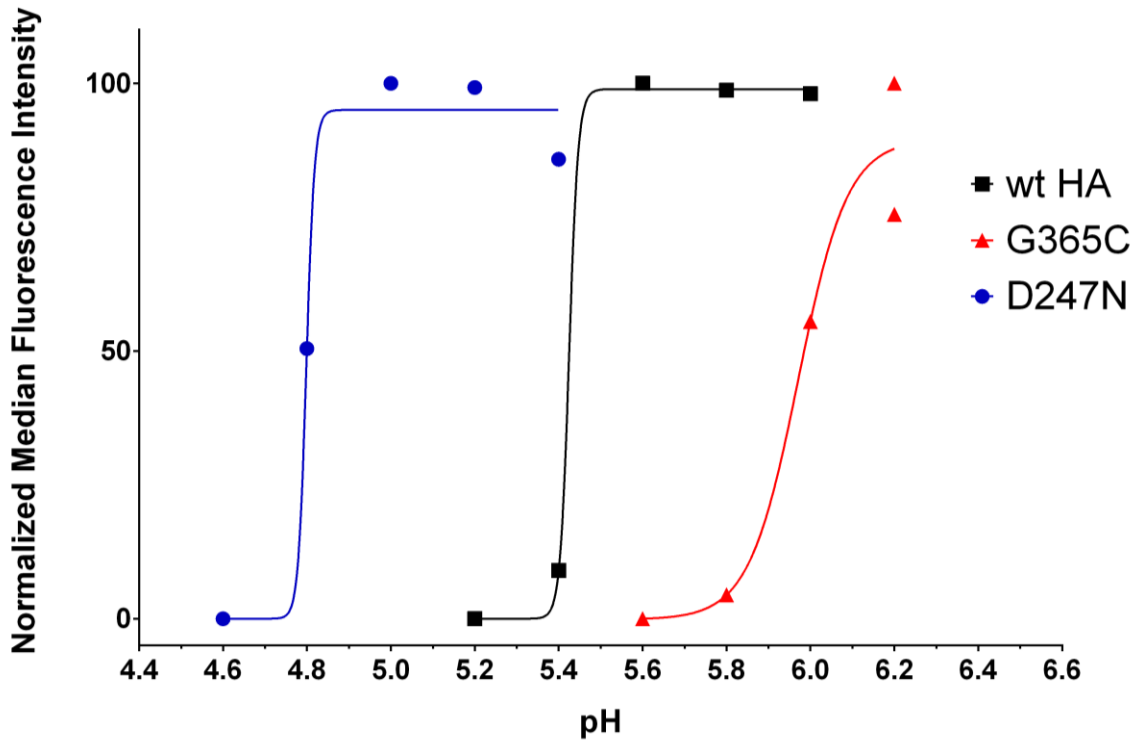


Figure 2.6 HA normalized pH transition curves. HEK293 cells expressing HA point mutations were exposed to buffered solutions at a range of pH spanning the expected transition pH. Viable expressing cells, gated on light scattering, were analyzed for refolding by flow cytometry using conformation specific antibody, HC58, and median fluorescence was plotted against pH. Shown here are representative pH transition curves for mutant D247N (blue), wt HA (black) and G365C. EC50 values, referred to in this dissertation as the 'transition pH', were extracted from these curves using Prism software package.

(pH 5.6 selection pressure), appear to be slightly destabilizing (Table 2.2). Meanwhile, two mutations from the HA 3.3 library, G237R and G365C, are significantly destabilizing (Table 2.2). Two of the three clones isolated from the HA3.3 library contained both G237R and G365C mutations, while the third clone from the HA3.3 contained D166G and G365C mutations with the D166G mutation not contributing to the high pH activating phenotype. Meaning G365C is likely solely responsible for the destabilized protein. The double mutation of G237R and G365C likely made HA highly destabilized.

2.3.3 Identification of Stabilizing Mutations

After analyzing individual mutations from the HA3.4 library clones the stabilization of HA in response to lower pH appears to be due to a single mutation, D247N. Interestingly, a second stabilizing mutation was also found in the HA3.2 library which was screened after a pH pulse of 5.6. D247N has an activation pH midpoint of 4.8 about 0.6 less than wild-type hemagglutinin (Table 2.2). The S429P mutant was further stabilized by about 0.4 pH units compared to wild-type HA making this mutation moderately stabilizing. These mutations will be further discussed below.

2.4 Discussion

Here we report the use of directed evolution to identify novel mutations in fowl plague hemagglutinin that contribute to either a stabilized or destabilized neutral pH structure. Previous work to better understand HA refolding has relied mainly on two approaches, selection of virus variants and rationally designed point mutants[35; 46; 47; 60; 64; 65; 69; 70; 71; 72]. While these approaches have proven successful in their ability to isolate HA variants that are either acid-stable or undergo the irreversible conformational change at elevated pH, they are intrinsically limited by the constraints of nature. HA mutants isolated in amantadine resistance experiments, where amatadine hydrochloride

Table 2.2 Transition pH for individual mutants

Mutation	Selection pH	Transition pH (EC50)
wild-type	-	5.545 ± 0.08
D93S	4.8	5.51 ± 0.04
Y155F	4.8	5.54 ± 0.05
A168V	4.8	5.56 ± 0.16
D240Y	4.8	5.35 ± 0.05
D247N	4.8	4.83 ± 0.18
E462Q	4.8	5.65 ± 0.08
I96M	5.6	5.79 ± 0.001
G142D	5.6	5.51 ± 0.05
D166G	5.6/6.0	5.50 ± 0.09
V188G	5.6	5.43 ± 0.22
Y218H	5.6	5.49 ± 0.10
W243S	5.6	5.67 ± 0.15
L407A	5.6	5.61 ± 0.02
S429P	5.6	5.06 ± 0.07
G237R	6.0	6.30 ± 0.16
G365C	6.0	6.09 ± 0.17

Note: Highlighted rows represent mutations that have a transition pH that is significantly different than wild-type hemagglutinin. Green represents stabilizing mutations, orange represents destabilizing mutations, and red significantly destabilizing mutations.

affects the intracellular pH,[65; 66] have been chosen based on the ability of cells treated with amantadine to produce viral particles capable of lysing cells.[64; 65] These mutants are therefore somewhat limited in the sequence space they can sample because of the necessity to retain membrane fusion activity and viral infectivity. Studies employing site-directed mutagenesis often changed residues based on: locations previously identified as affecting pH-sensing[65], sequence alignment[46; 60; 71], structurally interesting locations such as the fusion peptide or fusion peptide pocket[35; 72], or a combination of these.[47] Either way, previous studies have been limited because of selections based on membrane fusion or sequence alignments where highly conserved amino acids are typically not changed.

Directed evolution uses repeated rounds of selection to quickly enrich rare populations and has been used to successfully engineer a variety of proteins with improved fitness from single-chain antibodies[73] to enzymes that can be used as biocatalysts.[74] Using error-prone PCR to introduce mutations into the gene for FPV HA removed the natural constraint of fusion competency allowing for residues that are normally highly conserved across all HA subtypes to be mutated. Out of the sixteen unique mutations found in previous directed evolution screens, two of the mutations (G365C and W243S) were found at residues that are conserved across all HA subtype consensus sequences.[75; 76] To our knowledge, of the 16 mutations found screening the library of mutated H7 HAs, 15 have never been identified as mutations affecting pH sensing and the sixteenth mutation, G365C, was found to be stabilizing in a previous study.[71]

As previously mentioned, four mutations (I96M, G237R, W243S, and G365C) were found that destabilized the neutral conformation of HA. Two of the mutations are located at residues that are conserved among all HAs (W243 and G365) and three of the four mutations are located in the HA1 headgroup (Figure 2.7). The majority of

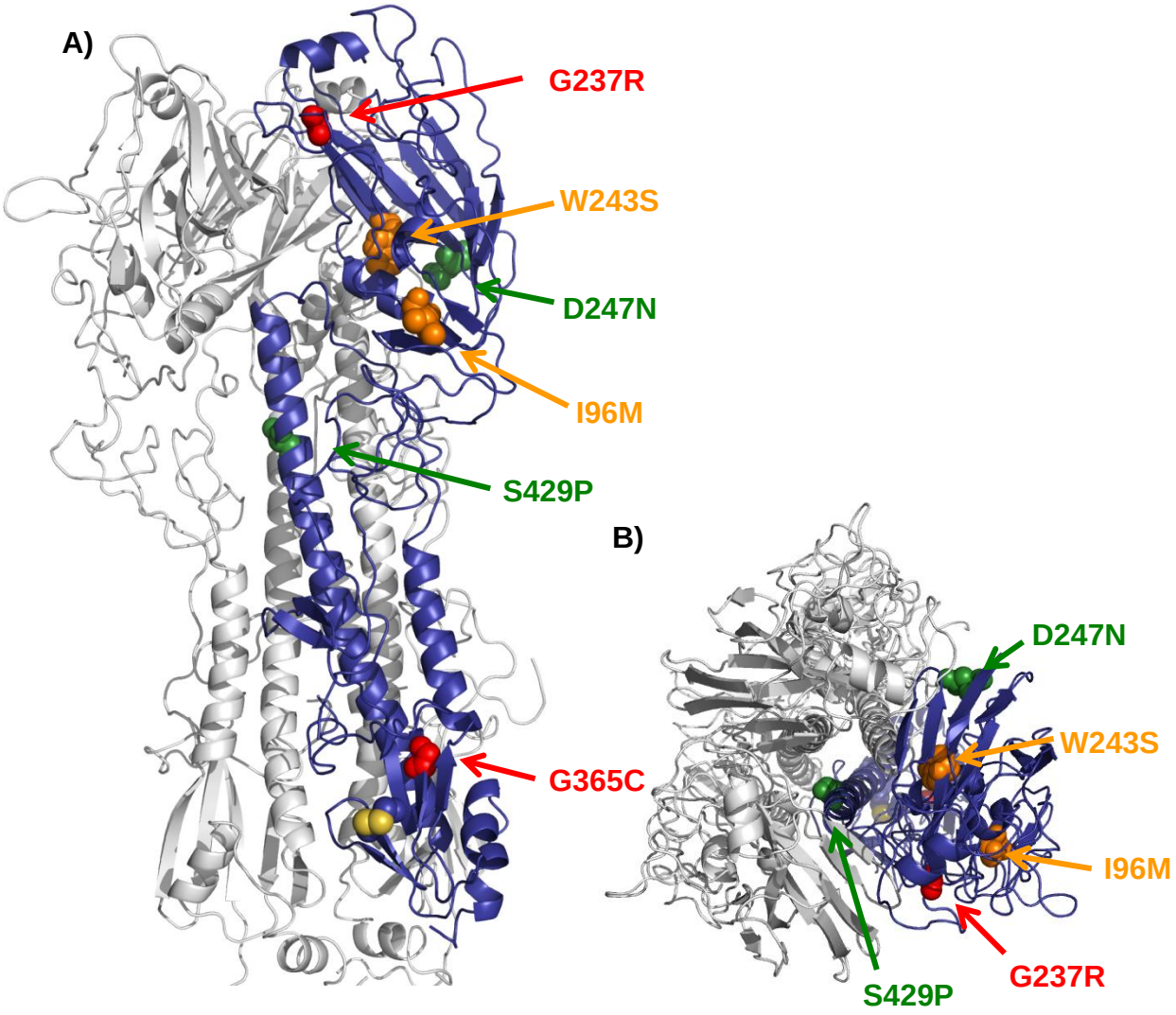


Figure 2.7 Structural locations of pH sensing mutations. A) Side image of hemagglutinin trimer with mutations that affect pH-sensing shown as spheres and color-coded according to Table 2.2. Coordinates used in the analysis are from PDB file 1HGF[77] and the images were generated using PyMol software (http://www.pymol.org). B) Top-down image of the HA trimer with same pH-sensing mutations displayed and clustering toward trimer interfaces.

previously identified pH-sensing mutations have been located in the HA2 subunit with very few mutations found to sense pH in the HA1 headgroup.[35; 46; 47; 64; 65; 70] The I96M mutation allows HA to undergo its acid-induced refolding ~0.3 units higher than wild-type FPV HA most likely due to disrupting the side chain packing. In the neutral pH conformation, the smaller side of isoleucine packs tightly into a small pocket created by three loops of HA1: T64 to D66, D93 to E97, and M274 to V282 (Figure 2.8). The additional length of the methionine side chain may have trouble fitting into this small pocket and that likely destabilizes the surrounding contacts. One contact that is likely altered as a result of Ile-Met substitution is the hydrogen bonding between Q277 and E411. E411 has been previously identified as a residue that makes critical contacts between the B-loop of HA2 and surrounding residues in the prefusion conformation and that B-loop rearrangement is a critical step in the loop-to-helix transition that occurs in HA2.[60] Taken together, it seems most plausible that steric hinderances caused by the methionine side chain eventually lead to disruption of hydrogen bonds between E411 and Q277, which allows for the neutral pH structure to become destabilized more readily.

While the structural implications of I96M are somewhat ambiguous, the effects of a Gly-to-Arg mutation at HA1 237 seem more straightforward. G237R raised the pH at which HA undergoes its structural transition by 0.8 units, similar to that for G365C. Glu-237 is located near the HA1-HA1 interface and is next to Arg-238. Introducing a second arginine residue close to the HA1-HA1 interface likely contributes to long-range electrostatic interactions that destabilize the interface. Two positively charged arginines next to each other and located near positively charged residues on the apposing HA1 headgroup (K217 and H219) could provide a strong electrostatic repulsive force to

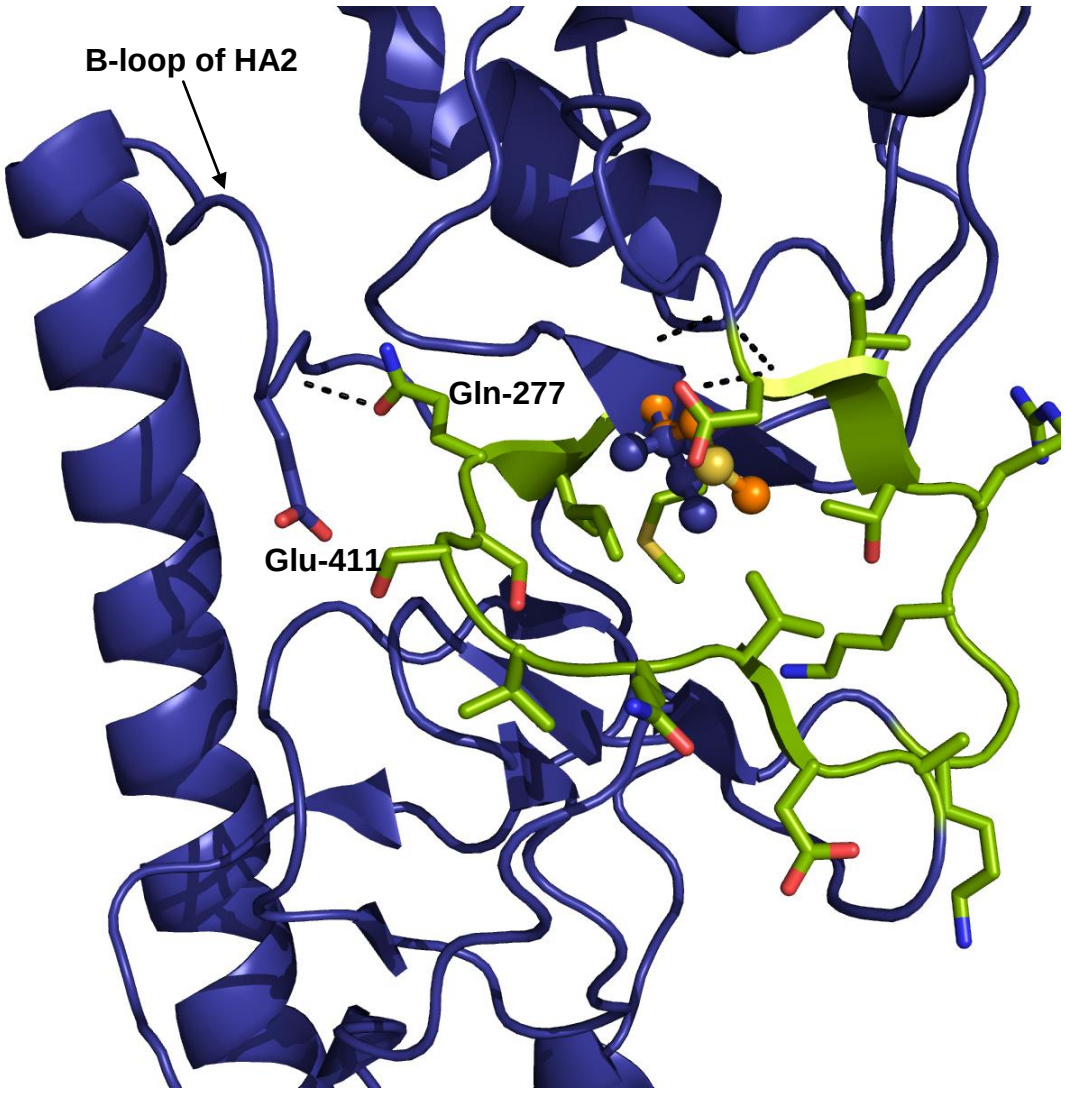


Figure 2.8 I96M mutation structural implications. Ile-96 is shown as ball-and-stick in dark blue and the Ile-96-Met mutation is shown as ball-and-stick also, and colored orange with the sulfur of the side chain yellow. The I96M mutation was modeled using Swiss pdbViewer and the rotamer with the best score was chosen. Possible interactions with the loop (colored lime green) containing Gln-277 may disrupt its interaction with Glu-411 of the HA2 B-loop ultimately destabilizing the neutral pH structure.

further destabilize the neutral pH conformation (Figure 2.9). The pKa of histidine is 6.0 and Mittal et. al. report that the pKa for each hemagglutinin monomer is between 5.6 and 5.7 indicating that histidine residues could play an important role in pH sensing.[78] Thus, while it is impossible to precisely know the pKa of H219, it likely becomes protonated providing additional positive charge at the same interface of the G237R mutation. The additional charge provided by the mutation, as well as the positive charge associated with the protonation of His-219 as the pH drops, disrupts the HA1 interface sufficiently enough for the headgroups to dissociate and HA2 to refold into the extended coiled coil, low pH conformation.

Arguably, a more interesting mutation to be identified as being responsible for the destabilization of the neutral pH conformation is W243S. W243S elevated the transition pH of HA by ~0.2 but W243 is a conserved residue among all HA subtypes and is located at an interface between HA1 and HA2 of an adjacent monomer. W243 fits into a hydrophobic pocket created mainly by the HA1 headgroup and Val415 of the B-loop of the neighboring HA2. As previously mentioned, the B-loop of HA2, which under goes the loop-to-helix transition upon activation, has been proposed to transition to a relaxed state under acidic conditions prior to undergoing its irreversible conformational change.[79] When W243 is occupying the hydrophobic pocket in the metastable structure, Val415 is unable to occupy the hydrophobic pocket of W243 thus maintaining the prefusion conformation (**Figure 2.10**). However when serine is introduced in place of W243, the hydrophobic pocket is unoccupied and valine is able to position its side chain in the hydrophobic pocket allowing the B-loop to adopt its “relaxed” form prior to the loop-to-helix transition. The proposed ability of valine to occupy the hydrophobic pocket would allow HA2 to more readily undergo its acid-induced conformational change thus elevating the transition pH of the protein.

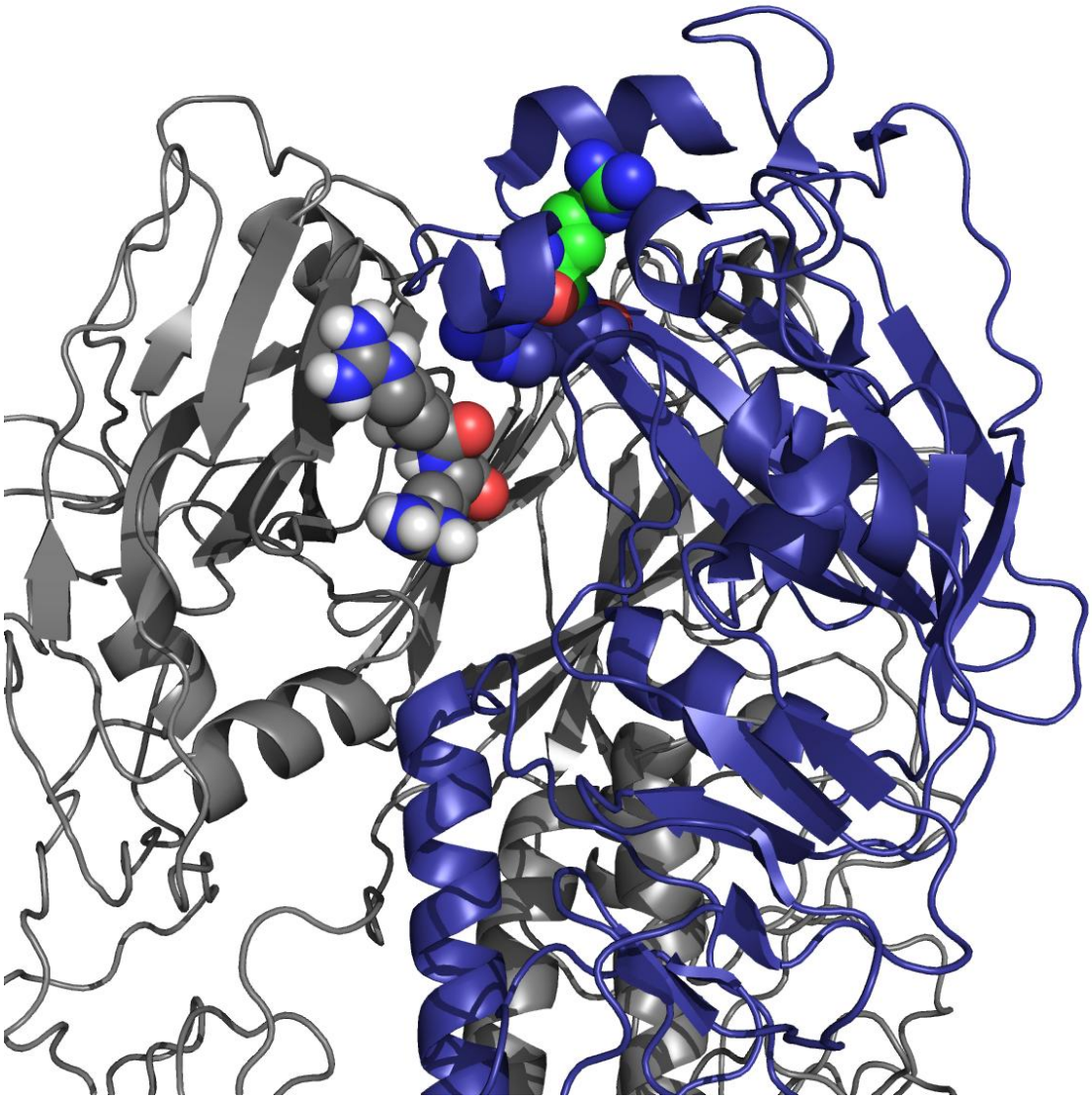
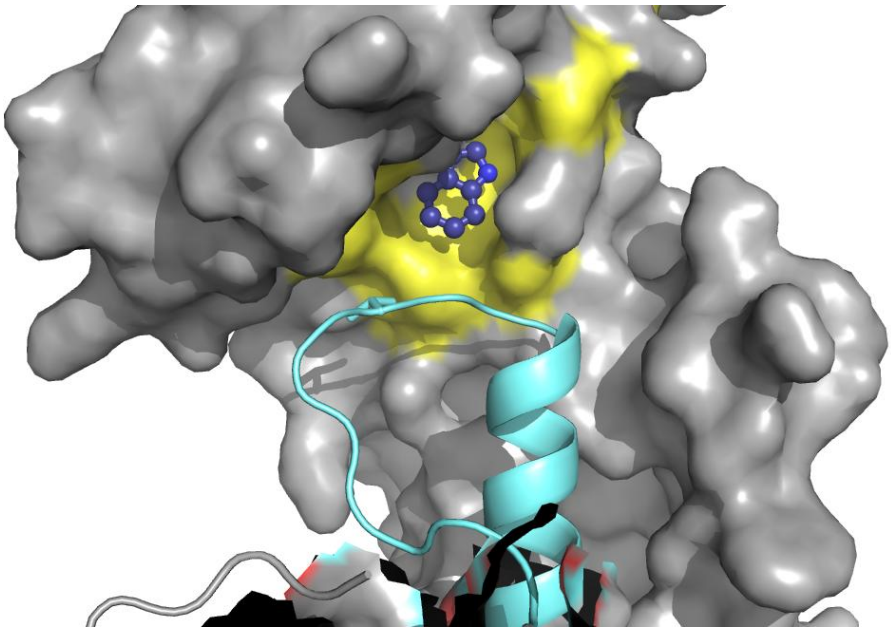


Figure 2.9 G237R electrostatic interactions. The G237R mutation was modeled using Swiss pdbViewer and shown as green spheres. R238 (purple) of the same monomer, as well as the locations of K217 and H219 (gray) of FPV HA, are also shown as spheres. The proximity of these positively charged residues likely contributes to electrostatic repulsion of the HA1 headgroups destabilizing the protein.

Figure 2.10 Effect of W243S mutation on HA structure and pH sensing. A) Structural rendering of W243 occupying a hydrophobic pocket in the neutral HA0 structure (PDB 1HA0).[80] Val-415 does not readily occupy the hydrophobic pocket at neutral pH when W243 is present. B) The mutation W243S clears way for Val-415 to move into the hyrodrophobic pocket thus allowing the B-loop of HA2 to swing outward into a "relaxed" conformation, primed to undergo the loop-to-helix transition that is essential in the spring-loaded refolding model. The two HA2 structures shown are at pH 8.1 (orange, PDB 3QQB) where the B-loop is still closely packed with the α helix, and at pH 5.3 (cyan, PDB 3QQO). The pH 5.3 structure was stabilized in the closed conformation by H2-HA2R106H and shows the B-loop adopting a "relaxed" position assumed to be an intermediate structure before undergoing the loop-to-helix transformation.[60]

A)



B)

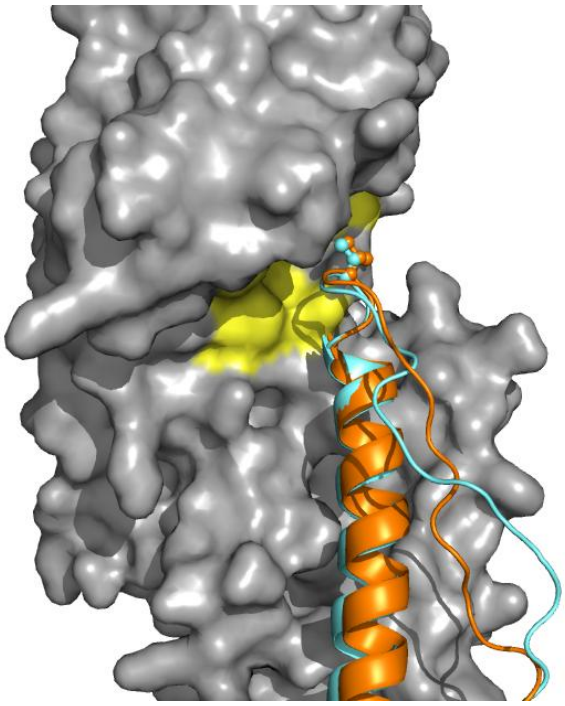


Figure 2.10 Continued

Results in this study show that the one of most destabilizing mutations found G365C, and together with D247N, had the largest deviation in transition pH from type HA (~0.6-0.7). Gly-365 is located at the end of the fusion peptide in the proximal region of HA2 near where HA1 and HA2 are covalently attached via a bond between C14 and C479. Substitution of Gly-365-Cys places an additional cysteine less than 12Å away from the HA1-HA2 disulfide (

Figure 2.11). We propose that G365C interferes with the native disulfide in at least one HA monomer, causing a gross displacement of that HA1 head group leading to a loss, or disruption, of all stabilizing contacts in the trimer head group. Previous work examining disulfide bond formation in HA found that mutant HAs with a C14S or C479S mutation failed to form disulfide bonds with any of the native cysteine residues of HA1, but confirmed that a small portion of mutants were transported to the cell surface in a non-native conformation.[81] The formation of a new disulfide bond between C14 and G365C would likely only occur in a few monomers given the propensity of HA to form native disulfide bonds. Formation of new disulfides in only one or two monomers would destabilize the entire trimer while still allowing the refolding of the other monomer(s) and exposure of the c-myc epitope that was used in the directed evolution work to identify destabilized HA, from which the G365C mutation was isolated. Interestingly, our flow cytometric data is in direct conflict with a previous study that made the same mutation in another H7 variant, A/Netherlands/219/03 (H7N7).[71] The study by Ilyusina et. al., found that substituting Cys for $Gly23_2$ resulted in a stabilized HA that refolded 1.0 point below the pH of the wild-type A/Netherlands/219/03 HA. While the different HA variants may explain some of the differences between the previously reported results and ours, it is likely not enough to explain a ~1.7 pH difference in transition pHs between two closely related HAs.

Directed evolution was also used to identify mutations that stabilized hemagglutinin and lowered the pH at which the protein undergoes its irreversible

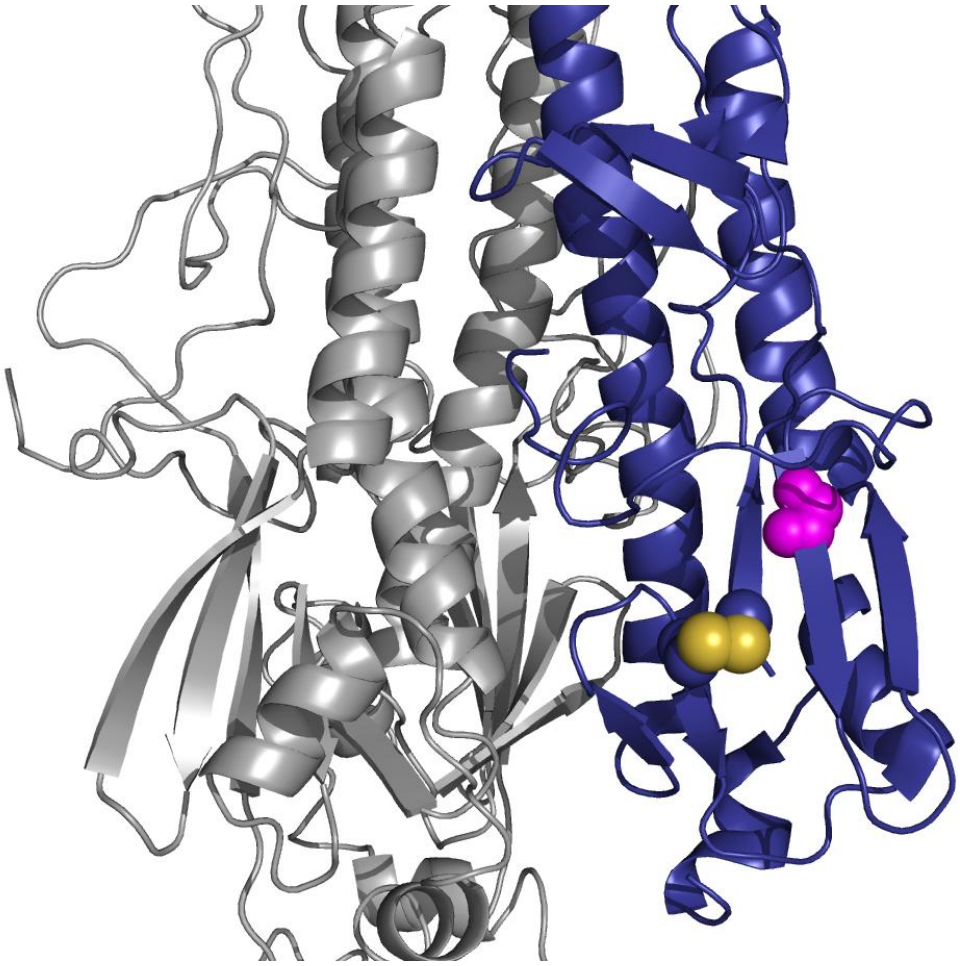


Figure 2.11 G365C mutation and disulfide bonding. G365C is shown as magenta spheres near the native disulfide that covalently links the HA1 and HA2 subunits. The destabilizing affect arising from the G365C mutation may be due to disulfide bond formation between the new cysteine of HA2 and Cys-14 which is one-half of the native disulfide. A new disulfide would likely shift the entire HA1 headgroup therefore destabilizing the entire protein.

conformational change. Two mutations have been identified as stabilizing the neutral conformation, D247N and S429P. However, only D247N was found via screening of the 4.8 pH-pulsed library. S429P was identified in clone HA3.2.5 along with two destabilizing mutations I96M and W243S, supporting the idea that multiple mutations with in HA can be cumulative.[65] S429P was found to stabilize the trimeric structure by 0.4 units from the wild-type transition pH. S429P is located near the membrane distal segment of the long α-helix of HA2. A serine to proline mutation likely stabilizes the neutral pH conformation through interactions between the large α helix of HA2 and the neighboring B-loop. In H2 HA, Phe-405 of the B-loop packs tightly against Phe-430 of a neighboring HA2 stabilizing the neutral pH structure.[60] In H7 HA, and specifically for Rostock/FPV HA, there is no aromatic side chain at position 430 or adjacent residues in a neighboring HA2 to make an aromatic stack with the conserved Phe-405 to stabilize the structure. The aliphatic ring structure of a proline substitution provides additional hydrophobic contacts with Phe-405 compared to serine providing a stabilizing force to keep the HA in an acid-stable conformation(Figure 2.12).[82] D247N stabilized the transition pH of HA by 0.6 and is located in the HA1 headgroup at the HA1-HA1 interface. The side chain of Asp-247 is positioned directly across from Glu-414 of an adjacent HA2 subunit, creating a charge repulsion that destabilizes the neutral conformation to some extent and is a critical component of the metastable conformation (Figure 2.13). Substitution of a neutral asparagine residue is not only a conservative mutation, but it acts to eliminate the charge repulsion between D247 and E414, stabilizing the neutral conformation of HA.

Interestingly, all of the destabilizing mutations were found in multiple clones except for I96M and each of the stabilizing mutations only appeared in one clone. Of the six mutations identified in this study, all but G365C and S429P were located in the HA1 headgroup. The locations of the mutations found through this directed evolution study reinforces the model of the HA1 headgroups acting as a “clamp” preventing

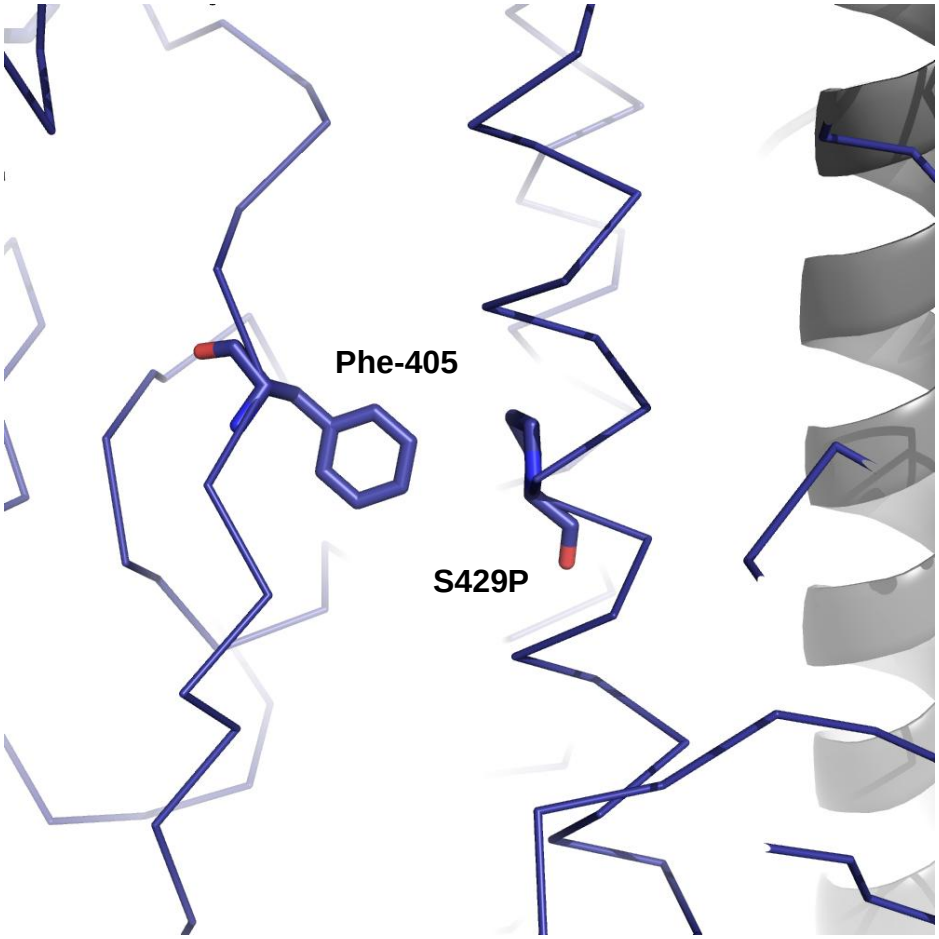


Figure 2.12 Stabilization of B-loop by S429P mutation. The addition of a aliphatic ring by the mutation of Ser-429 to proline allows the aromatic residue of Phe-405 of the B-loop to be stabilized in a ring packing manner. The S429P mutation was modeled in using Swiss pdbViewer and the lower scoring trans oriented side chain was chosen.

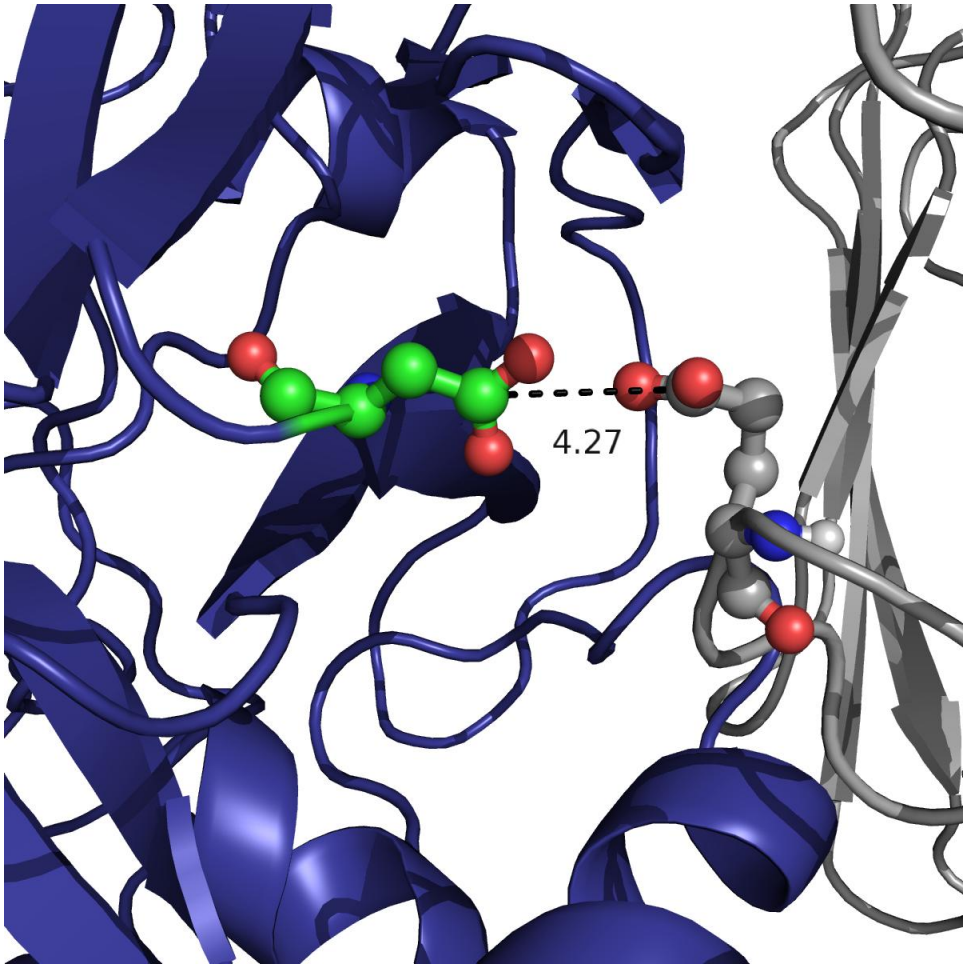


Figure 2.13 Asp-247-Asn mutation relieves electrostatic repulsion. In the native HA structure the negatively charged side chain of Asp-247 is directly opposed by the negatively charged side chain of Glu-414 and separated by roughly 4Å. The proximity of both negative charges would destabilize the HA1-HA1 interface priming it for refolding. By substitution of Asp with Asn, that charge repulsion is relieved stabilizing the interface.

refolding of hemagglutinin. Other studies have proposed that the B-loop of HA2 and regions of HA1 that interact with this loop are critical for controlling the pH of fusion.[35; 60] Our results tend to agree with this finding based on the number of mutations that likely interact with the B-loop of HA2 in some capacity. We propose, more specifically, that it is the HA1 headgroup that acts as the primary pH-sensing mechanism and HA1-HA1 interactions must be sufficiently interrupted by the increasing concentration of protons within the endosome, so that HA1 headgroups dissociate allowing the HA2 subunit to undergo its conformation rearrangement. The total displacement of HA1 globular domains is in agreement with a previous computational study that showed complete dissociation of the head groups is the most energetically favorable pathway to the low pH, irreversible conformation.[63] Through directed evolution, mutations were made to a few highly conserved residues that had previously not been explored for pH-sensing and the location of both stabilizing and destabilizing mutations within the HA1 headgroup provide clues to HA1 acting as the primary pH sensor that controls hemagglutinin switching.

Chapter 3 : Engineering and characterization of an improved chimeric adhesive switch via directed evolution using yeast surface display

3.1 Introduction

Nature has evolved to include many complex processes that require sophisticated proteins capable of controlling cellular signals and functions. This is often accomplished through allosteric regulation where output signals are modulated in an 'on/off' fashion by the presence or absence of effector molecules (e.g., small molecules or ligands). Allosteric regulation was first proposed fifty years ago as a way to describe proteins in metabolic networks that regulated the production of metabolites and revolved around two core concepts which haven't changed; first, the effector binding site is spatially distinct from the active site, and, second, the effect is entirely due to conformational changes within the protein structure.[10]

These key concepts have been the basis for a number of different approaches to engineering functional molecules for use in sensing, diagnostic or therapeutic functions. Controlling transgene networks for the treatment of diseases is reliant upon precise control of genes at the transcription and translation levels with small molecules, such as siRNA, or gene switches.[83; 84] One such example of a gene switch is nuclear hormone receptor, hERα, which had its ligand selectivity engineered from its natural ligand 17β-estradiol to a synthetic ligand 4,4'-dihydroxybenzil (DHB) using site saturation mutagenesis and directed evolution. [85] Protein switches have also been created for diagnostic and sensing applications by fusing heterologous genes together to couple two disparate protein functions in a modular fashion. Examples of modular switches created by varying degrees of domain insertion include fusions of TEM1 β-lacatamase (BLA) and maltose binding protein (MBP), photoprotein (AEQ) and sulfate-binding protein (SBP), and between the photoactive LOV2 domain of *Avena sativa* and *E. coli trp* repressor.[86; 87; 88] Until recently, proteins such as GFP, luciferase, dihydrofolate

reductase (DHFR) and β-galactosidase have been heavily relied upon in engineering protein switches as they provide simple phenotypic screening and selection strategies.[14; 15]

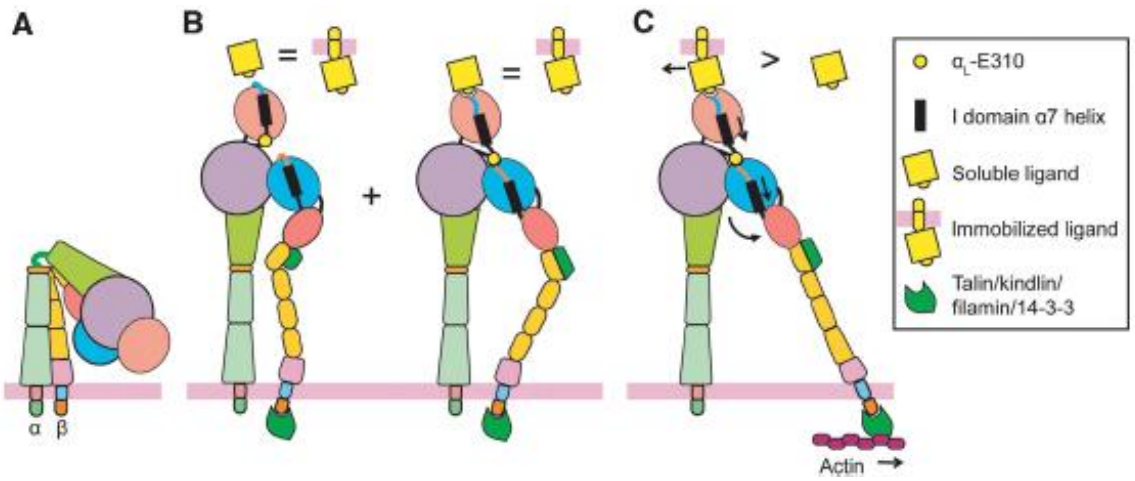
Recent work in our lab has demonstrated the ability to modulate cellular adhesion between the αL inserted (I) domain and an intercellular adhesion molecule-1 (ICAM-1)-coated surface under shear flow by the presence or absence of an effector peptide.[89] The allosteric rearrangement within the I domain was exploited to make a functional switch by fusing the EF3 and EF4 hands of calmodulin to the N- and C-termini of the I domain, respectively (Figure 3.1 D).

3.1.1 I domain

In an innate immune response, inflammation is a front-line defense against infection, injury or trauma. Inflammatory responses typically only last a few days and include a wide array of cell types, receptors, selectins, and chemokines. While inflammation tends to be a transitory response, prolonged inflammation is responsible for a number of pathological and clinically relevant conditions such as rheumatoid arthritis, Crohn's disease, and transplant rejection where prolonged inflammation leads to tissue and organ damage. Among the many players of the inflammation cascade arguably the most important are integrins.

The inflammatory cascade begins when blood vessels dilate and flow is slowed.[90] Inflammed endothelial cells express E- and P-selectin on their surface, which in conjunction with reduced flow allows leukocytes expressing P-selectin glycoprotein ligand 1 (PSGL1) to transiently tether to the vasculature. The transient interactions between endothelial cells and leukocytes are mediated by the high on- and off-rates of the selectin-ligand interactions; this selectin-mediated rolling allows for intracellular signaling to occur within the leukocyte via chemokine binding to G-protein coupled

Figure 3.1 Role of αL I domain in LFA-1 and a fusion protein switch. A, B and C) Activation of LFA-1 from a bent, closed, low-affinity conformation (A), to an extended, closed, low-affinity conformation (B, left). An inside-out signal is transduced up the β-leg resulting in alteration of the β I-like domain (blue domain) MIDAS allowing Glu-310 to bind and pull down the α7 helix of the I domain (pink domain, B right). In nature, the bond between ICAM-1 and the αL I domain is strengthened by force (C). D) The pulling down of the α7 helix is mimicked in our switch construct by the binding of the EF3 and EF4 hands of calmodulin to a peptide of smooth muscle myosin light chain kinase (smMLCK, purple helix). The original chimeric fusion between I domain and the EF3 and EF4 hands of calmodulin, prior to any mutagenesis or directed evolution, is referred to in this dissertation as “wild-type switch.” Upon binding, the α7 helix (orange) is displaced downward allowing for the adoption of a high affinity conformation by the I domain. Mg^{2+} is depicted as a red sphere at the top of the MIDAS.



D)

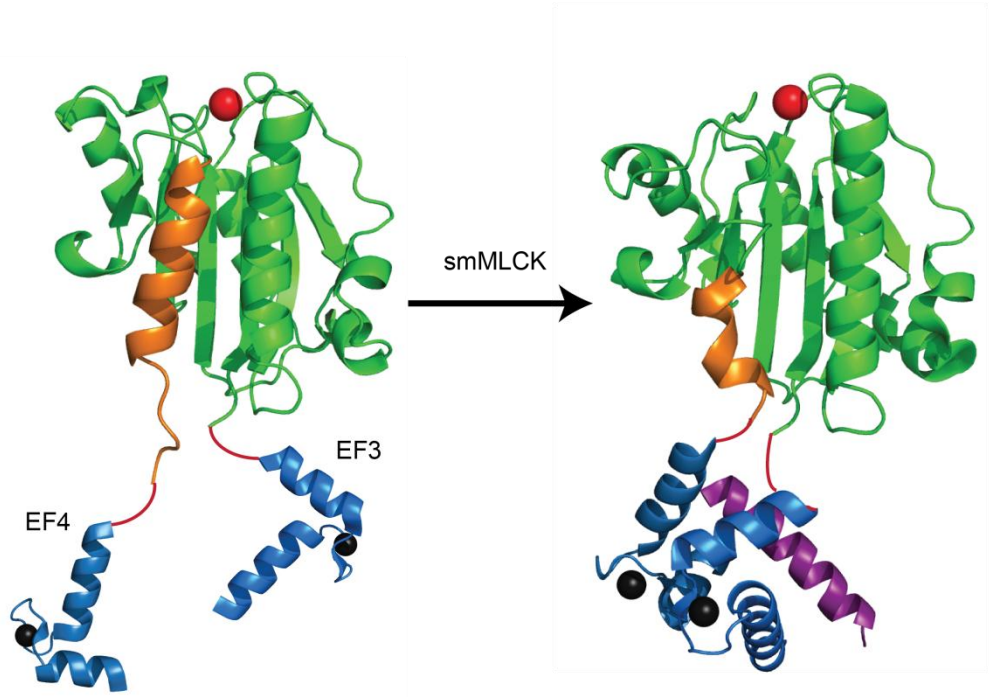


Figure 3.1 Continued

receptors (GPCRs).[91] Activation of GPCRs by chemokines induces a complex signaling network, which at this time is not completely understood, [91] but that activates integrins by actin-binding proteins such as talin-1, kindlin, or filamin.[11; 91] Binding of integrin cytoplasmic domains by actin-binding proteins separates the α and β cytoplasmic domain tails inducing a 'switchblade-like' conformational rearrangement into an extended, primed conformtation in a process termed 'inside-out signaling' (Figure 3.1C). [11; 90; 92] Integrin avidity, the total adhesiveness of the ligand-protein interaction, is composed of integrin affinity for its ligand and the valency (number of contacts) between the protein and ligand. Once activated, integrins facilitate leukocyte firm adhesion to the endothelium followed by crawling and then paracellular and transcellular transmigration into the site of inflammation to carry out its immunological activity. [91]

Lymphocyte function-associated antigen-1 (LFA-1) is probably the most widely studied and and well characterized member of the integrin family. It is a heterodimeric transmembrane protein composed of noncovalently associated α and β subunits. While each subunit is large and complex, the ligand binding site is contained solely within the 180 amino acid inserted (I) domain of the αL subunit of LFA-1 (αLβ2, CD11a/CD18) [93]. The crystal structure of the αL I domain reveals a central hydrophobic pocket composed of six β-strands adopting a Rossman fold surrounded by 7 amphipathic α helices.[9; 93; 94] A Mg^{2+} ion is located at the "top" of the crystal structure as part of a metal ion-dependent adhesion site (MIDAS), and is ligated by five residues, D137, S139, S141, T206 and D239 (Figure 3.2).[93] LFA-1 binds intercellular adhesion molecules (ICAMs), of which it has the highest affinity for ICAM-1.[11; 93] ICAM-1 binds to the I domain of LFA-1 through coordination of an acidic residue (Glu34) to the MIDAS of I domain, which is surrounded by a ring of hydrophobic residues in both the ICAM-1 and I domain. A salt bridge between Glu-241 and Lys-39 of I domain and ICAM-1, respectively, is also crucial for ligand binding.[93] The exposure of the MIDAS is in conjunction with the axial

Figure 3.2 αL I domain. A) Low-affinity, “closed”, conformation of αL I domain. The magnesium ion of the MIDAS (metal ion-dependent adhesion site), in the top portion of the structure, is colored in yellow. Residues coordinating the Mg^{2+} of the MIDAS are shown as ball and sticks, where oxygen atoms are red and nitrogen atoms are blue. Glu-310 acts as a pseudo-ligand for the β2 I-like domain and ligates the Mg^{2+} of the I-like MIDAS in a similar fashion to Glu-34 of ICAM-1 to the MIDAS of the αL I domain. B) High affinity, “open”, conformation I domain structure. Again, the Mg^{2+} of the MIDAS is shown as a yellow sphere. The α7 helix has been displaced downward by 10Å compared to the low affinity conformation. Additionally, the coordinating residues have also rotated as a result of the allosteric rearrangements due to the motion of the α7 helix and are now in a position to facilitate ICAM-1 binding. C) ICAM-1 bound to the I domain showing ligation between Glu-34 of ICAM-1 and the Mg^{2+} of the I domain MIDAS.

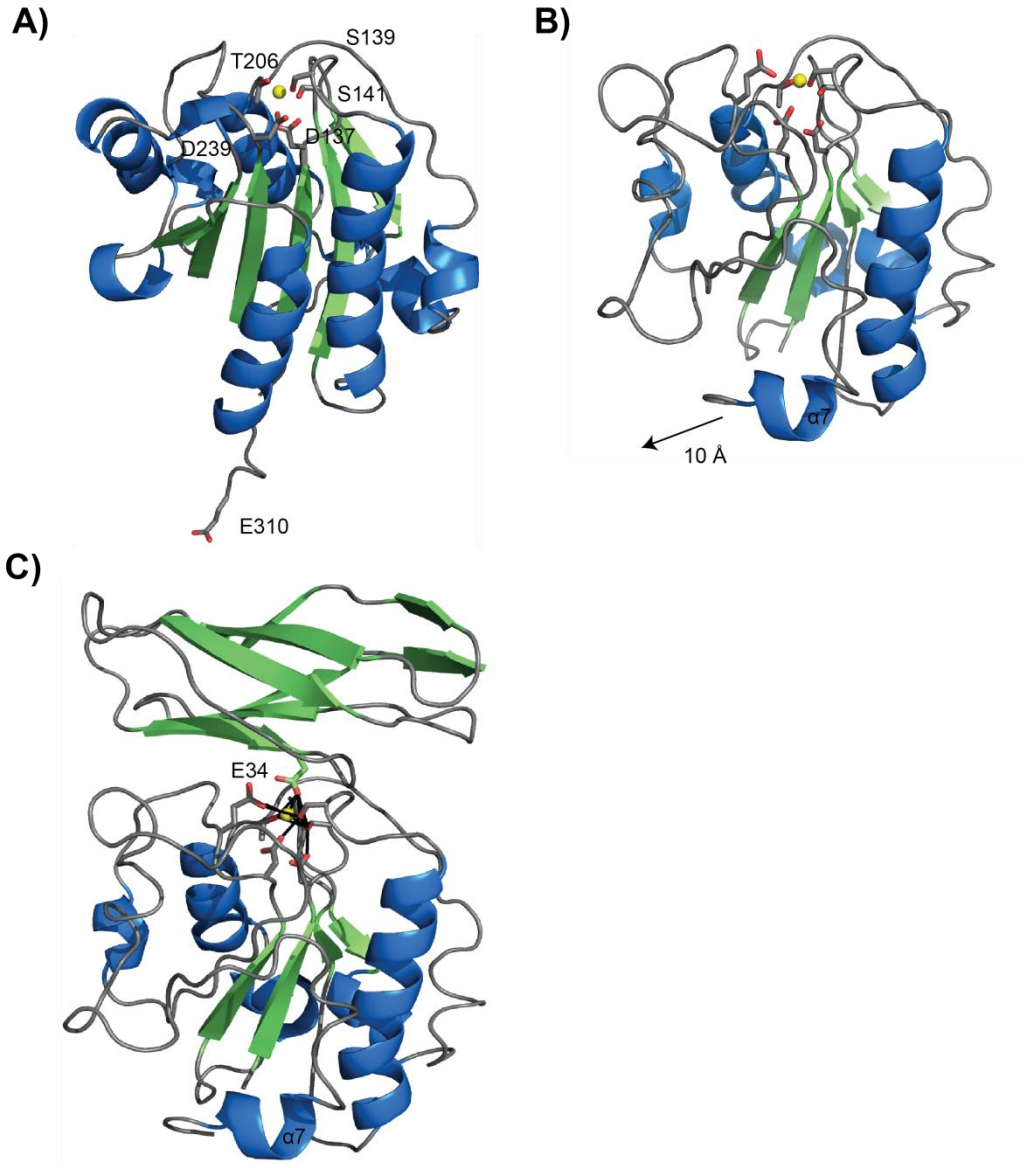


Figure 3.2 Continued

downward movement of the α7 helix of I domain by about 10 Å. The downward movement of the α7 helix of the I domain acts as a pulley causing a rearrangement of the metal ion-dependent adhesion site (MIDAS) drastically altering the affinity for the ligand, ICAM-1.[9; 11; 92; 95] This downward movement toward the β I-like domain also allows for interaction between Glu-310 of the linker between the α I domain and β propeller domain and the β I-like MIDAS. This conformational interaction between the different subunits allows LFA-1 to transduce signals bidirectionally through the different domains.[9; 92; 93; 94; 96] This allosteric interaction between the I domain and the I-like domain in the β subunit makes the I domain an ideal candidate for engineering a novel molecular switch.

The allostery of the I domain can be coupled with other protein domains to create a chimeric protein switch. Additionally, the flexibility seen within I domain with the proximity of its N- and C-termini makes I domain very suitable for circular permutation. Circular permutation, coupled with allosteric proteins, gives engineers a powerful combination to create novel molecular switches with increased affinities or functions [14; 86; 97]. Previous work has demonstrated that the individual EF-hands of each calmodulin subdomain can solubly recombine in the presence of the target peptide, smooth muscle myosin light chain kinase (smMLCK).[98] Fusion of the EF-hands to the termini of the I domain likely produces a disruption in the allosteric region of the I domain due to the repositioning of the α7 helix upon binding of the EF3 and EF4 hands to smMLCK peptide leading to a three-fold increase in adhesion (**Error! Reference source not found.** D).[89] Here we show that through focused mutagenesis on the allosteric region of the I domain and directed evolution a near twenty-fold improvement in switch activity was achieved. This improvement in switch activity has resulted in mutant switches that were effectively “on/off” in their behavior.

3.2 Materials and Methods

3.2.1 Yeast strains and plasmid

S. cerevisiae EBY100 (MAT**a** AGA1::GAL1-AGA1::URA3 ura3-52 trp1 leu2Δ1 his3Δ200 pep4::HIS3 prb1_1.6R can1 GAL) cells were transformed with pCT302-derived vectors allowing for selection under Trp- and Ura- conditions and surface expression of proteins under the galactose inducible GAL1,10 promoter.[99] pCT302 was used as the expression vector for all experiments except for circular permutation and fusion constructs were cloned in using NheI and BamHI restriction sites. Routine sticky-end restriction enzyme digestion and ligation reactions were performed as follows. Briefly, 10 µL of purified plasmid DNA was digested using 1 µL (20 Units) of NheI-HF (NEB) and 1 µL (20 Units) of BamHI-HF(NEB), 4 µL of NEB CutSmart buffer (10X), 0.4 µL BSA (100X) and 23.6 µL of dH_2O in a total reaction volume of 40 µL. The preparative double digest was placed in a 37°C water bath for 2-3 hours. The reaction was stopped with 10X stop buffer (50% (w/v) glycerol, 0.1 M EDTA (pH 7.5), 1% (w/v) SDS, and 0.1% (w/v) bromophenol blue). Digested DNA was analyzed by gel electrophoresis and the correctly sized bands were excised and purified using the Promega Wizard gel purification kit following the manufacturer's instructions. Ligation reactions contained 3 µL of purified linearized vector, 6 µL of purified DNA to be inserted, 2 µL of T4 Ligase buffer, 8 µL of dH_2O and 1 µL of T4 Ligase. Ligation reactions were carried out at 16°C overnight and the following day, 2 µL of ligation reaction was used to transform E. coli. Use of NheI and BamHI restriction sites allows two epitope tags (HA and c-myc) on either side of the gene of interested to be expressed in the ORF. For the circular permutation library, the AfeI restriction site was inserted into pCT302 via PCR on the C-terminal side of NheI and the N-terminal side of BamHI, respectively, allowing for blunt-end cloning and termed vector pCT-JVP (primer sequences can be found in Appendix 6). Briefly, the PCR reaction to insert the AfeI restricition sites used 5 units of *Taq* DNA polymerase (NEB) with 5 µL of 10X standard

Taq buffer, 200 µM dNTPs, roughly 50 ng of plasmid DNA, and 0.15 µM of each primer. PCR cycles used were as follows: initial denaturation at 95°C for 2 minutes, then 30 cyles of 95°C for 30 seconds, annealing at 62°C for 30 seconds, extension at 68°C for 50 seconds, followed by a 5 minute final extension at 68°C. The PCR product was inserted into the TA-cloning vector, pCR2.1, using the TOPO TA cloning kit following the manufacturer's instructions (Invitrogen). The gene fragment containing the newly inserted AfeI restriction sites was then transferred from the cloning vector to a linearized pCT302 vector using restriction enzyme digestion and ligation. The AfeI restriction sites were confirmed by gel electrophoresis and sequencing. Yeast were grown in the semi-defined media SD-CAA (20g/L dextrose, 6.7 g/L yeast nitrogen base without amino acids, 5.4 g/L Na_2HPO_4, 9.7 g/L $NaH_2PO_4 \cdot H_2O$, and 5 g/L CAA) at 30°C overnight. The following day, expression was induced by subculturing at an OD_{600} of 1 and grown overnight in SGR-CAA (same as SD-CAA expect 2% (w/v) dextrose is replaced with 2% galactose and 2% D-Raffinose).

3.2.2 Magnetic-activated Cell Sorting (MACS)

Yeast were grown and induced as previously described. Following overnight induction in SGR-CAA at 30°C, $3x10^6$ cells (OD_{600} of 0.3) were transferred to microcentrifuge tubes for washing and immunostaining. 10^8 yeast were washed twice with 1 mL of PBS(137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4) supplemented with 1 mM $MgCl_2$ and 0.1% BSA. A MACS LS column (Miltenyi Biotech, Auburn, CA) was prepared by washing the column with 3 mL of PBS supplemented with 2 mM EDTA and 0.05% BSA. 60 µL of ICAM-Fc (R&D Systems, Minneapolis, MN) at 200 µg/mL was added to 200 µL of Protein G magnetic beads to saturate the beads with ICAM; the ICAM-Protein G beads were incubated at 4°C for 45 minutes. The ICAM-Protein G beads were applied to the LS column and washed three times with 5 mL of PBS with 1 mM $MgCl_2$ and 0.1% BSA to remove an unbound ICAM-Fc. The complexed ICAM-Protein G beads were eluted with 500 µL of buffer and 450 µL of the

ICAM-Protein G beads were then applied to 10^8 yeast. The library and complexed ICAM-Protein G beads were incubated for 1 hour at 4°C. Meanwhile, a new LS column was prepared as previously described. The yeast/bead mixture was added to the column, washed with buffer as before, and eluted with 5 mL of PBS plus 2 mM EDTA and 0.05% BSA. The eluate was then used to inoculate 245 mL of SD-CAA.

3.2.3 Flow cytometry

Yeast were grown and induced as previously described in section 3.2.1. Following overnight induction in SGR-CAA at 30°C, $3x10^6$ cells (OD_{600} of 0.3) were transferred to microcentrifuge tubes for washing and immunofluorescent labeling. Yeast were washed twice with 500 µL Labeling Buffer (137 mM NaCl, 2.7 mM KCl, 10mM Na_2HPO_4, 2mM KH_2PO_4, 10 mM $MgCl_2$, and 1 mM $CaCl_2$). Following the washes, cells were incubated with mouse anti-c-myc antibody 9E10 (ascites fluid, Covance, Princeton, New Jersey) at a 1:100 dilution in Labeling Buffer for 45 to 60 minutes at 4°C along with either smMLCK peptide or a scrambled version of smMLCKp (henceforth referred to as 'scrampep') at a concentration of 10 µM. Yeast were washed once with 500 µL of Labeling Buffer then resuspended in 30 µL for labeling with fluorescent secondary antibodies and precomplexed ICAM-Fc/anti-Fc AF647. 1.5 µL of ICAM-Fc per sample (200 µg/mL) (R&D Systems, Minneapolis,MN) was precomplexed with Alexa Fluor 647-conjugated goat anti-human anti-Fc secondary antibody (Jackson ImmunoResearch, West Grove, PA) at a ratio of 1:1.6 at 4°C for at least 60 minutes. Yeast were costained with precomplexed ICAM-Fc/AF647 anti-Fc, at a final ICAM-Fc concentration of 10 µg/mL, goat anti-mouse Alexa Fluor 488 (Invitrogen) at 1:100 and smMLCKp or scrampep at room temperature, protected from light for 1-2 hrs. Following the secondary incubation, yeast were spun down at 10,000xg for 1 minute and resuspended in 1 mL of fresh Labeling Buffer. Samples were analyzed on an Accuri C6 flow cytometer (BD Biosciences, Ann Arbor, MI). 10,000 events were collected per sample and gated based on size. Background-corrected mean AF647 fluorescence

(i.e., ICAM-1 binding) was normalized by background-corrected mean AF488 fluorescence (i.e., full length protein surface expression). Normalized mean fluorescence of ICAM-binding of samples incubated with smMLCKp was divided by normalized mean fluorescence of ICAM-binding of samples incubated with scrampep to calculate a fold-change in ICAM binding. The following equations outline the fold-change calculations:

$$NMFI = \frac{MFI647_{sample} - MFI647_{bkgd}}{MFI488_{sample} - MFI488_{bkgd}}$$

$$Fold \quad Change = \frac{NMFI_{smMLCKp}}{NMFI_{scrampep}}$$

where NMFI is the normalized mean fluorescence intensity, MFI 647_{sample} is the mean fluorescence intensity due to ICAM-binding by the sample protein , MFI647_{bkgd} is the mean fluorescence intensity of ICAM binding by a negative control population expressing a disparate protein on the yeast surface, MFI 488_{sample} is the mean fluorescence intensity of protein expression by the sample protein and MFI 488_{bkgd} is the mean fluorescence intensity of protein expression by uninduced yeast.

3.2.4 Library construction of circularly permutated I domain

A short, six amino acid linker containing two BamHI restriction sites used to circularize the gene (GGGSG) was added by PCR and referred to as Id6L. 400 µg of plasmid DNA containing Id6L was digested with 2,000 units of BamHI (New England Biolabs Inc., Ipswich, MA) for six hours at 37°C. The digest was stopped with 10X stop buffer (recipe). Linearized Id6L was gel purified and then ligated into circular genes

using T4 Ligase (NEB) overnight at 16°C. Overnight ligations were desalted using Promega PCR clean-up kit (Promega, Madison, WI). Circularized Id6L (10 µg) was then digested with 20 mUnits of DNaseI (NEB) for 10 minutes at room temperature in a total reaction volume of 300µL. The DNase digest was stopped with 6 µL of 0.5M EDTA to a final concentration of 10 mM and heat inactivated at 65°C for 10 minutes. Following heat inactivation, digested DNA was desalted as previously described. About 3.5µg of digested DNA was repaired using 12 U of T4 ligase, 3U T4 polymerase and 200µM dNTPs with 10X T4 ligase buffer in a total reaction volume of 200µL. The DNA repair reaction was carried out at 12°C for 30 minutes and heat deactivated at 75°C for 20 minutes using a thermocycler. Each step is the circular permutation process was verified by gel electrophoresis. pCT312 (~12 µg plasmid DNA) was digested with 15 units of AfeI for at least 2 hours and heat inactivated at 65°C for 20 minutes. Following heat inactivation of the AfeI enzyme, digested vector was dephosphorylated using Antarctic phosphatase at 37°C for 40 minutes and then heat inactivated at 65°C for 10 minutes. Digest, dephosphorylated pCT312 was analyzed by gel electrophoresis and the linearized, dephosphorylated vector was gel purified for ligation with the circular permuted Id6L library. Circularly permuted Id6L library was ligated with linearized, dephosphorylated pCT312 at a ratio of 3:1 using T4 ligase and 5% PEG3350 with 10X T4 ligase buffer in a total reaction volume of 120µL overnight at 16°C. Overnight library ligations were transformed into MegaX DH10B electrocompetent cells (Life Technologies, Carlsbad, CA) using 3 µL of ligated DNA and 20 µL cells per transformation. Three transformations were pooled and recovered for 1 hour at 37°C prior to plating serial dilutions and library size was estimated to be $1x10^6$ the following day. Library DNA was purified and frozen for future use. The circularly permuted library was transformed into yeast using the lithium acetate method.[100]

3.2.5 Construction of switch library

Random mutagenesis of the α7 helix and EF4 hand was performed using nucleotide analogs. Ten cycles of error-prone PCR using 2.5 units of *Taq* polymerase and 500μM each of dPTP and 8-oxo-dGTP (TriLink Biotechnologies, San Diego, CA) introduced mutations into the switch gene between two ClaI restriction sites located within the gene. Homologous recombination using 50-bp of homology on each side of the ClaI restriction sites allowed for efficient recombination of the mutagenized gene. Seven electroporations each containing 1 μg of vector DNA and 1.2 μg of library insert per 50 μL of EBY100 were pooled together and transformed yeast were allowed to recover in YPD (2% w/v dextrose, 2% w/v peptone, and 1% w/v yeast extract) for 1 hr at 30°C. The transformed yeast library was then spun down at 1,000xg for 3 minutes and resuspended in 10mL of SD-CAA and serial dilutions were plated. Library size was estimated to be $1.4x10^7$.

3.2.6 Fluorescence-activated cell sorting

The library was sorted a total of four times using alternating rounds of positive and negative selections to enrich for yeast that displayed improved switching behavior in response to the presence of the smMLCK peptide. The first round of library sorting was a positive selection to enrich for a population that bound ICAM-1 with high affinity in the presence of 10 μM smMLCK peptide. The second and third rounds were negative sorts to deplete the library of yeast that bound ICAM-1 in the presence of the scrambled peptide at 10 μM. The fourth round of sorting was a positive sort to re-enrich for ICAM-1 high affinity binders in the presence of 10 μM smMLCK peptide. Library expressing yeast were grown and induced as previously described. Following overnight induction at 30°C for at least 20 hours, $3.5x10^7$ cells (OD_{600} = 3.5) were aliquoted and immunostained as previously described. Briefly, cells were washed with Labeling Buffer then labeled with anti-c-myc mouse monoclonal antibody, 9E10 (Developmental Studies Hybridoma Bank, University of Iowa) at a dilution of 1:50 for 1 hour on ice. During the

primary incubation, library expressing yeast were incubated with either 10 μM smMLCK peptide or 100 μM scrambled smMLCK peptide. Following the primary incubation, cells were washed with Labeling Buffer then labeled with precomplexed ICAM-1-Fc/anti-Fc AlexaFluor 647, smMLCKp at 10 μM or scrampep at 100 μM, and goat anti-mouse AlexaFluor 488 at a 1:100 dilution.

3.2.7 Protein purification

pET28b(+) (Novagen, Darmstadt, Germany) was used as the soluble protein expression vector. Restriction sites, NheI and XhoI, were used to insert the gene of interest into the ORF of pet28b(+) allowing for proteins to be expressed with a N-terminal His6 tag and a C-terminal c-myc tag to allow for diagnostic determination of full length protein. *E. coli* BL21(DE3) were transformed by electroporation using 1 μL of plasmid DNA per 50 μL of BL21(DE3) cells with plasmids containing genes encoding for either high-affinity I domain mutant F265S or fusion, switch proteins. Overnight BL21 cultures were diluted 1:50 into fresh LB-Kan (50 μg/mL) media and induced with 1M isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1mM at an OD_{600} between 0.4 and 0.6. Cells were allowed to continue growing following induction for another 3-6 hours. Cells were harvested by centrifugation at 5,000xg for 10 minutes and frozen cell pellets were stored at -20°C until the protein was extracted. Cell pellets were thawed on ice and then resuspended in 10 mL/g cells B-PER (Thermo Scientific, Rockford, IL) and 5 μL/g cell pellet of DNaseI (NEB). Cell suspensions were homogenized with a needle and syringe for about 5 minutes and allowed to incubate at room temperature for 15 minutes with mild shaking. Soluble protein (fusion protein switches) was collected in the soluble fraction following centrifugation at 15,000xg for 15 minutes. High affinity I domain is found in inclusion bodies and had to be further purified from inclusion bodies. The pellet containing inclusion bodies was resuspended in 5 mL B-PER and 4 μL Lysozyme (50 mg/mL) and homogenized, incubated at room temperature and centrifuged as with the separation of the soluble fraction previously

described. This process was repeated twice more to isolate inclusion bodies and then resuspended in Buffer A (50 mM phosphate, 8 M urea, pH 7.5) and boiled for 1 hr. Protein refolding was accomplished using an on-bead method.[101] Briefly, solubilized H.A. Id was added to prepared TALON cobalt resin (Clontech, Mountain View, CA) and shaken overnight at 4°C. The slurry was then washed twice with 40 mL of refolding buffer (50 mM phosphate, 1 mM $MgSO_4$, 5% (v/v) glycerol, pH 7.5) and resuspended in 30 mL of refolding buffer and allowed to shake gently overnight at 4°C. The beads/protein were collected in a column, washed twice with 5 CV of Equilibrium/Wash Buffer (50 mM phosphate, 300 mM NaCl, pH 7.0) and eluted with 2X 1.5CV of Elution Buffer (50 mM phosphate, 300 mM NaCl, 300 mM imidazole, pH 7.0). Fractions were confirmed by western blotting or Coomassie staining. His-purified proteins were further purified using the ÄKTApurifier with a Superdex 200 10/300 GL column (GE Healthcare Bio-Sciences, Pittsburgh, PA) according to the manufacturer's instructions to remove impurities that eluted from the TALON resin. Briefly, purified proteins were dialyzed into SPR Buffer (20 mM Tris, 150 mM NaCl, 10 mM $MgCl_2$, 2 mM $CaCl_2$, 0.01% (v/v) Tween-20, pH 8.0) using Slide-A-Lyzer cassettes (Thermo) following the manufacturer's instructions. Briefly, 3 mL of eluted protein was added to the cassette, dialyzed in 2X500 mL of SPR buffer at room temperature for 2 h each. Buffer was then replaced with 500 mL of fresh SPR buffer and dialyzed overnight at 4°C with gentle stirring. Dialyzed protein was then removed from the cassette and stored at 4°C until purification on the FPLC. Prior to loading protein on the FPLC, it was concentrated using Amicon Ultra centrifugal filter units (Millipore, Billerica, MA). 500 µL of concentrated protein was loaded onto a Superdex 200 10/300GL and eluted with 1.5 column volumes of buffer. Fractions were collected, analyzed by SDS-PAGE and saved for surface plasmon resonance experiments.

3.2.8 Surface Plasmon Resonance

Protein switch activity was analyzed using surface plasmon resonance. ICAM-Fc was immobilized at a level of 1220 RU (response units) on a CM5 sensor chip using the amine coupling kit (BIAcore, Piscataway, NJ). Analytes were flowed over the immobilized ICAM-Fc at a flow rate of 10 µL/min, 25°C in SPR buffer plus either scrampep or smMLCKp at 10 µM. Analyte was injected for 240 sec at the following concentrations: 200 µM, 100 µM, 20 µM, 10 µM, 2 µM, 1 µM, 200 nM, 100 nM and 2 nM. Analyte was dissociated using 20 mM Tris (pH 8), 300 mM NaCl and 20 mM EDTA. An adjacent flow cell was left empty and blocked and used as a control to subtract bulk effects. Kinetic analyses were performed using Scrubber2 after subtracting bulk change and nonspecific binding.

3.3 Results

3.3.1 Circular Permutation

Circular permutation is an efficient method for identifying sites amenable to domain insertion. This approach has been successfully applied to create a number of different proteins whose function is allosterically regulated following domain insertion.[16; 19; 97] Following the creation of the circularly permuted I domain library and transformation into *E. coli*, seventeen colonies were randomly selected and sequenced to verify that the library contained full-length, circularly permuted I domain. Out of the 17 sequences, 2 (11.8%) didn't contain the I domain gene (i.e., they were re-ligated vector), 9 (52.9%) contained full-length, circularly permuted I domain sequences, and 6 (35.3%) contained extremely truncated (<40% of total sequence) versions of the gene. Since the I domain gene with the GGGSG linker is 570 bp the library needs to contain at least 1615 clones to theoretically sample each DNase I cut site within the gene assuming that 35.3% of genes remain full-length following the circular permutation. The

circularly permuted I domain library was transformed into yeast and contained 10^6 transformants, thereby ensuring that the entire library was adequately sampled.

Clones were initially screened by two rounds of magnetic-activated cell sorting (MACs) to deplete the library of nonfunctional clones and enrich for binders (Figure 3.3). The first round of MACS resulted in roughly two-fold enrichment for ICAM binders. Following the second round of MACS the library was enriched 10-fold over the starting library for ICAM binders (Figure 3.3). This enriched library was then recovered by growing in SD-CAA media and sorted by FACS. Following one round of fluorescence-activated sorting, a population was enriched 17-fold for permutants that retained the ability to bind ICAM. Twenty clones were then analyzed for their specific ability to bind ICAM and to determine their new sequence. Interestingly, in all of the permutants that retain ICAM binding, only changes to their C-terminal allosteric region including the β6-α7 loop and the α7 helix were tolerated. Previous studies have shown that the allosteric region, residues Y257 to E310, of LFA-1 I domain is critical for relaying the conformational changes of the whole integrin to the active site in order to bind ICAM-1 with high affinity.[28; 93] The circular permutation results indicating that this region of the protein is the most tolerant to changes in its structure, coupled with previous studies identifying the importance of this region for ICAM-1 binding, led to the design of primers to focus random mutagenesis in this region.

3.3.2 Directed evolution for improved switches

A library of approximately 1.5×10^7 EF3-I domain-EF4 switch mutants was created using error prone PCR with 500 μM nucleotide analogs. Four rounds of fluorescent activated cell sorting (FACS) were conducted to try and isolate a population of switch proteins whose affinity for ICAM-1 was allosterically regulated by the binding of an agonist peptide, smMLCK. The naïve library was phenotypically similar to the wild-type switch where 15.9% and 14.9% of the naïve library bound ICAM-1 in the

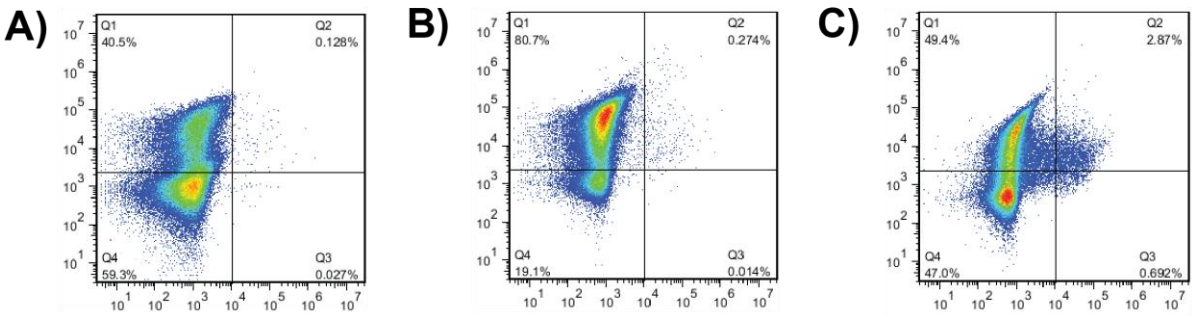


Figure 3.3 MACS enrichment of functional circular permuted I domain library. A) Naïve library of circular permuted I domain. B) Following a single round of MACS, the library was enriched by roughly two-fold over the naïve library. C) A second round of MACS resulted in 10-fold enrichment of functional permuted I domains compared to the naïve library.

absence or presence of scrambled peptide, respectively, while 12.9% of the original I domain-calmodulin fusion, referred to henceforth as wild-type switch, bound ICAM-1 in the absence of peptide (Figure 3.4 A and B). In the presence of smMLCK peptide, 22.9% of the naïve library bound ICAM-1, similar to the wild-type switch with 23.8% bound ICAM-1 with smMLCK present. The first round of sorting conducted with 10 μM smMLCK peptide was used to isolate populations that bound ICAM-1 with increased affinity. Following the first sort, 42.8% of the library was able to bind ICAM-1 in the presence of the antagonist peptide (i.e., scrambled peptide) while 47.5% of the library population bound ICAM in the presence of the agonist peptide (i.e., smMLCKp) (Figure 3.4C). The next two rounds of sorting were considered to be negative sorts where library expressing yeast that bound ICAM in the presence of 10 μM scrambled peptide were excluded and only library expressing yeast that didn't bind ICAM in the presence of the scrambled peptide were collected. Following these two negative sorts, the library was depleted 30-fold of proteins that bound ICAM in the presence of scrambled peptide resulting in a phenotypic change in the library population resembling a true off-state (Figure 3.4E). A final positive sort for ICAM binders in the presence of smMLCK was conducted to enrich the library population that, as a whole, looked to bind ICAM only in the presence of smMLCK. The final sort resulted in a library population that had a near 3-fold increase in ICAM binding in the presence of smMLCK compared to ICAM binding in the presence of scrambled peptide (Figure 3.4F).

Following the final round of FACS, yeast were plated and a total of 50 mutants were analyzed for switch activity. The majority of clones showed improvements in switch activity with some mutants displaying as much as a 20-fold increase in ICAM-1 binding in the presence of peptide (Figure 3.5). A few clones demonstrated decreased ICAM binding in the presence of smMLCKpeptide. Eight of the top clones and four of the clones that had decreased switch activity were sequenced to identify mutations that

Figure 3.4 Directed evolution of EF3-I domain-EF4 switch via FACS. Flow cytometry plots showing both protein surface expression and ICAM-1 binding in the presence of either smMLCKp (agonist) or the scrambled version of smMLCKP (scrampep, antagonist). Surface expression is displayed on the vertical axis in mean fluorescence intensity (MFI) and cells above the horizontal gate can be considered to be expressing full length switch constructs. ICAM-1 binding is displayed on the horizontal axis in MFI and cells to the left of the vertical gate can be considered to have proteins binding ICAM-1. A) Wild-type switch B) Naïve library (i.e., library expressing yeast prior to sorting) C) Switch library following the first sort with smMLKCp D) Library following a negative sort in the presence of scrampep E) Library following a second negative sort in the presence of scrampep F) Library following the last round of sorting; a positive sort in the presence of smMLCKp.

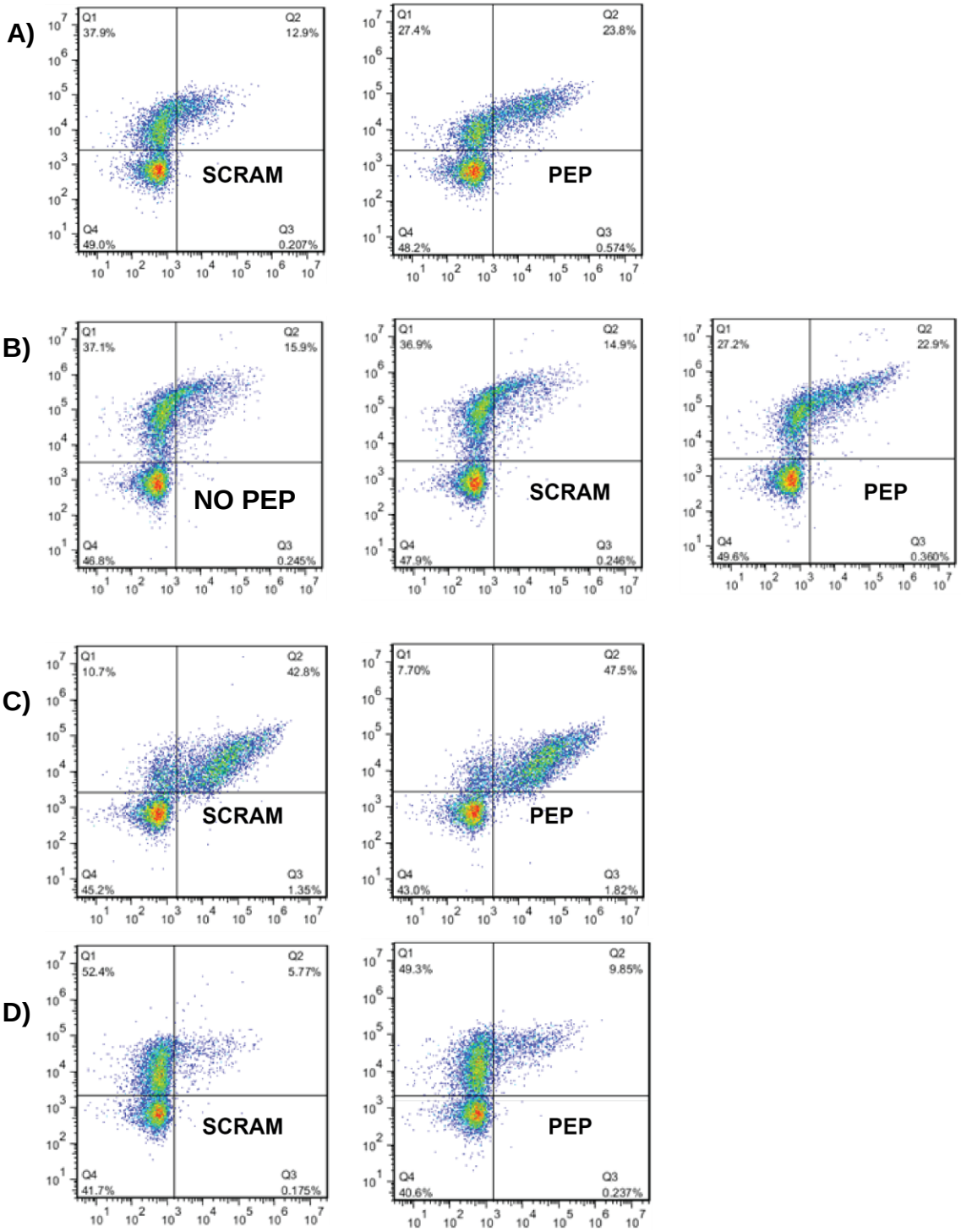


Figure 3.4. Continued

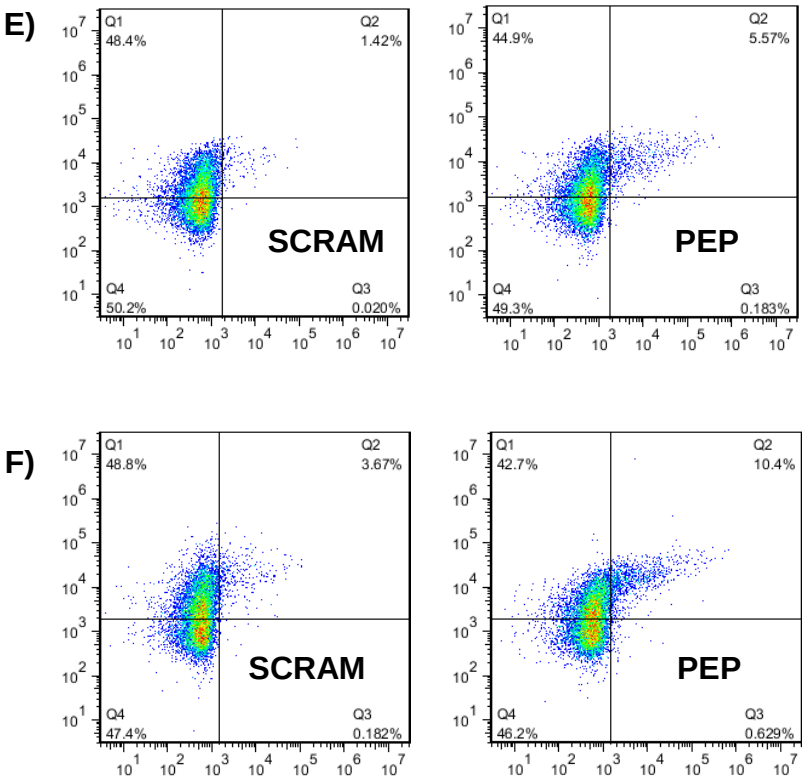


Figure 3.4 Continued

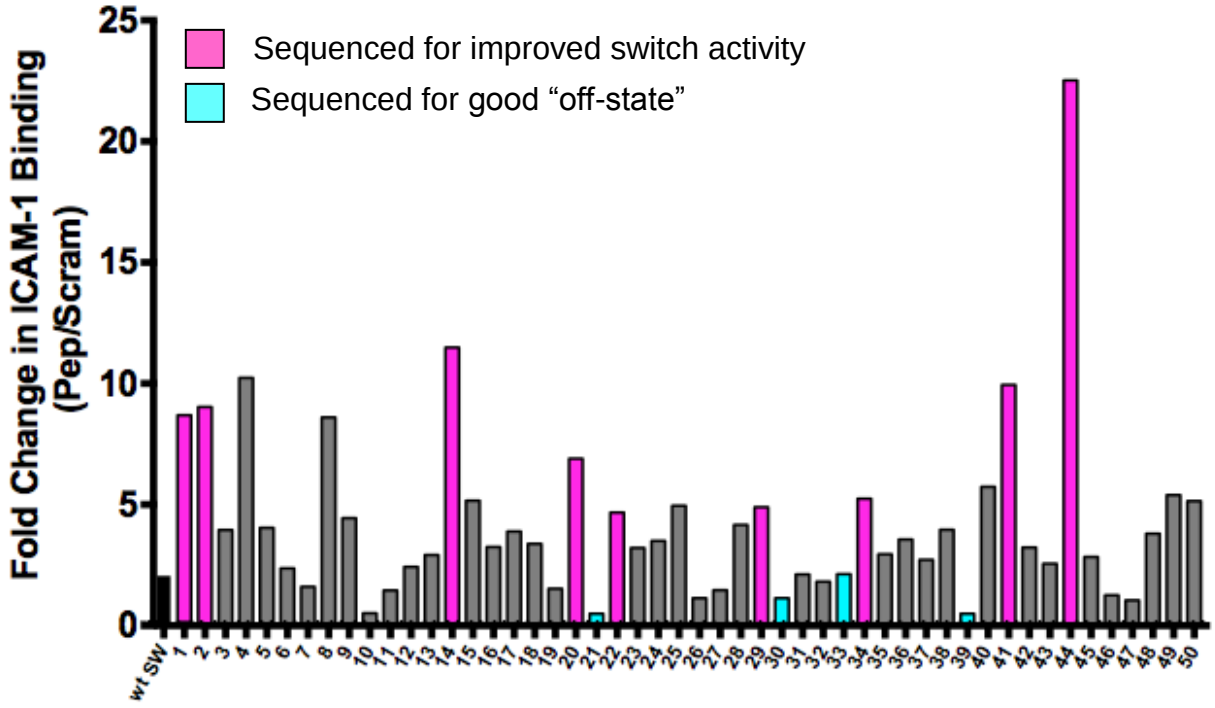


Figure 3.5 Analysis of individual clones following four rounds of directed evolution for improved switched. Fifty clones were analyzed for their switching ability relative to the wild-type switch. Fold change is calculated as discussed in the materials and methods section. The top 9 clones were sequenced to identify mutations that improved switch activity in the presence of smMLCKp and three of the lowest switch activity clones were also sequenced to identify mutations that improved the “off-state” of the switch in the presence of scrampled peptide.

both decreased ICAM binding in the presence of the scrambled peptide and increased ICAM binding in the presence of smMLCKp.

3.3.3 Mutations affecting switch activity

Seventeen unique mutations were found in the twelve mutants that were sequenced (Table 3.1). Three mutations (L289R, E293A, and Y307D) were found in multiple mutants, and all were located in either the β6-α7 loop or the α7 helix of the allosteric region of I domain. Three mutations (I318S, M337V, and K341E) were located in the EF4 domain of calmodulin. The double mutant, Leu-289-Arg and Tyr-307-Asp, was found in three of the seventeen mutants sequenced. Tyr-307 was also mutated to a serine. The frequency in which position 307 was mutated highlighted the need to explore the role of this position in affecting switch activity. A fourth double mutant, clone 21, which demonstrated decreased switch activity compared to the wild-type switch contained the mutations L289R and I231L. Given the conservative nature of the isoleucine to leucine mutation, it was hypothesized that L289R is responsible for locking the switch in a low affinity state. A series of sixteen single, double, triple and quadruple mutants were created from the mutations found through directed evolution to try and identify a synergistic set of mutations to further increase switch activity (Table 3.2).

The only single mutant, M8, contained the Leu-289-Arg mutation thought to lock the switch in a low affinity state. As predicted, M8 saw a decrease in switch activity relative to wild type switch and lacked the ability to bind ICAM-1 in the presence of smMLCK peptide (Table 3.2 and Figure 3.6). This mutation was included in the other 15 mutants to improve the off-state of the switch in the presence of scrambled peptide. Mutants M5, M6, and M7 all contained the EF4 mutation I318S and resulted in decreased switch activity. Double mutants, M2 and M3, differing in which EF4 residue was mutated (M337 and K341 for M2 and M3, respectively) showed markedly different

Table 3.1 Mutations from directed evolution of chimeric switches

Clones	Mutations	Fold Change in ICAM Binding
Wt	Wild type	2
1	K263R, D290A, Y307D, K341E	8.5
2	Y307S	9
14	M337V	12
20	L289R, Y307D	7
21	L289R, I231L	0.5
22	I258L, E293A	5
29	L289R, Y307D	5
33	V286G	2
34	T273P	5.5
39	E241V, F277C, K280N, I318S	0.5
41	L289R, Y307D	10
44	K252Q, E293A	23

Table 3.2 Designed point mutations for improved switch activity

Clone	Mutations	Fold Change in ICAM Binding
wt	Wild type	4.4 ± 0.8
M1	L289R, I318S	0.3 ± 0.2
M2	L289R, M337V	11.2 ± 2.8
M3	L289R, K341E	19.9 ± 16.8
M4	L289R, M337V, K341E	8.1 ± 2.5
M5	L289R, I318S, M337V	4.3 ± 2.3
M6	L289R, I318S, K341E	1.9 ± 1.0
M7	L289R, I318S, M337V, K341E	3.6 ± 2.4
M8	L289R	6.8 ± 4.0
M9	L289R, Y307D	11.2 ± 3.2
M10	L289R, Y307S	10.3 ± 5.7
M11	L289R, Y307D, M337V	2.0 ± 1.0
M12	L289R, Y307S, M337V	94.3 ± 28.7
M13	L289R, Y307D, K341E	31.1 ± 12.9
M14	L289R, Y307S, K341E	33.4 ± 11.6
M15	L289R, Y307D, M337V, K341E	68.0 ± 30.7
M16	L289R, Y307S, M337V, K341E	124.1 ± 113.3

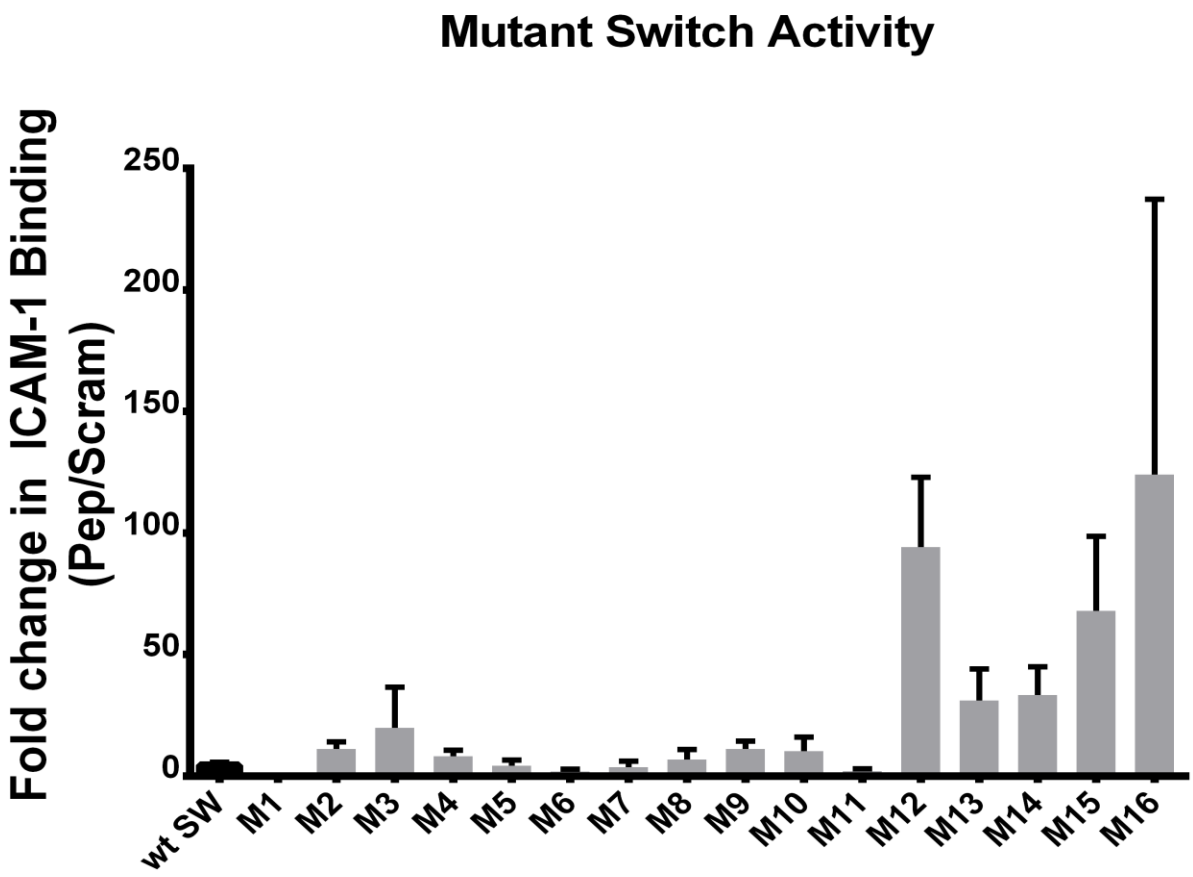


Figure 3.6 Mutant switch activity measured by flow cytometry.

switch activities indicating that M337V is likely more beneficial than K341E is to improved switching. The effect of these two mutations is also seen when comparing M12 and M14. Comparing M11 to M12, M13 to M14 and M15 to M16, it appears that Y307S is more beneficial than Y307D to improving switch activity. Two of the three mutations of the EF4 domain, M337V and K341E, likely increase the specificity of the EF hands for smMLCKp in relation to scrambled peptide. The third mutation, I318S, is thought to disrupt the hydrophobic pocket created by the two EF hands in which Trp-799 of smMLCK peptide is buried. This disruption of the hydrophobic pocket by substitution with a polar residue diminishes peptide binding and therefore switch activity.

3.3.4 Effect of peptide concentration on switch activity

The effect of peptide on switch activity was tested by incubating the wild-type switch and one of the top performing switches, M12, with differing concentrations of smMLCK and scrambled peptide. Peptide concentrations ranging between 200 nM and 200 μM were tested. At concentrations of smMLCK peptide above 100 μM the amount of ICAM-1 bound by both switches increases significantly and both wild-type and M12 switches bind ICAM-1 with similar affinities in the presence of 200 μM smMCK peptide (Figure 3.7). However, a noticeable difference in ICAM-1 affinity is seen between the two switches in the presence of the scrambled peptide. Scrambled peptide at 200 μM causes about at 40% increase in ICAM-1 binding with the wild-type switch compared to switch M12 (comparing ICAM-1 binding fluorescence intensity at 200 μM with scrambled peptide). Similarly, the fluorescence intensity associated with ICAM-1 binding in the presence of the scrambled peptide is consistently lower across all peptide concentrations with the M12 switch compared to the wild-type switch. The lowered affinity also increases the dynamic range of the M12 switch allowing for improved switch activity.

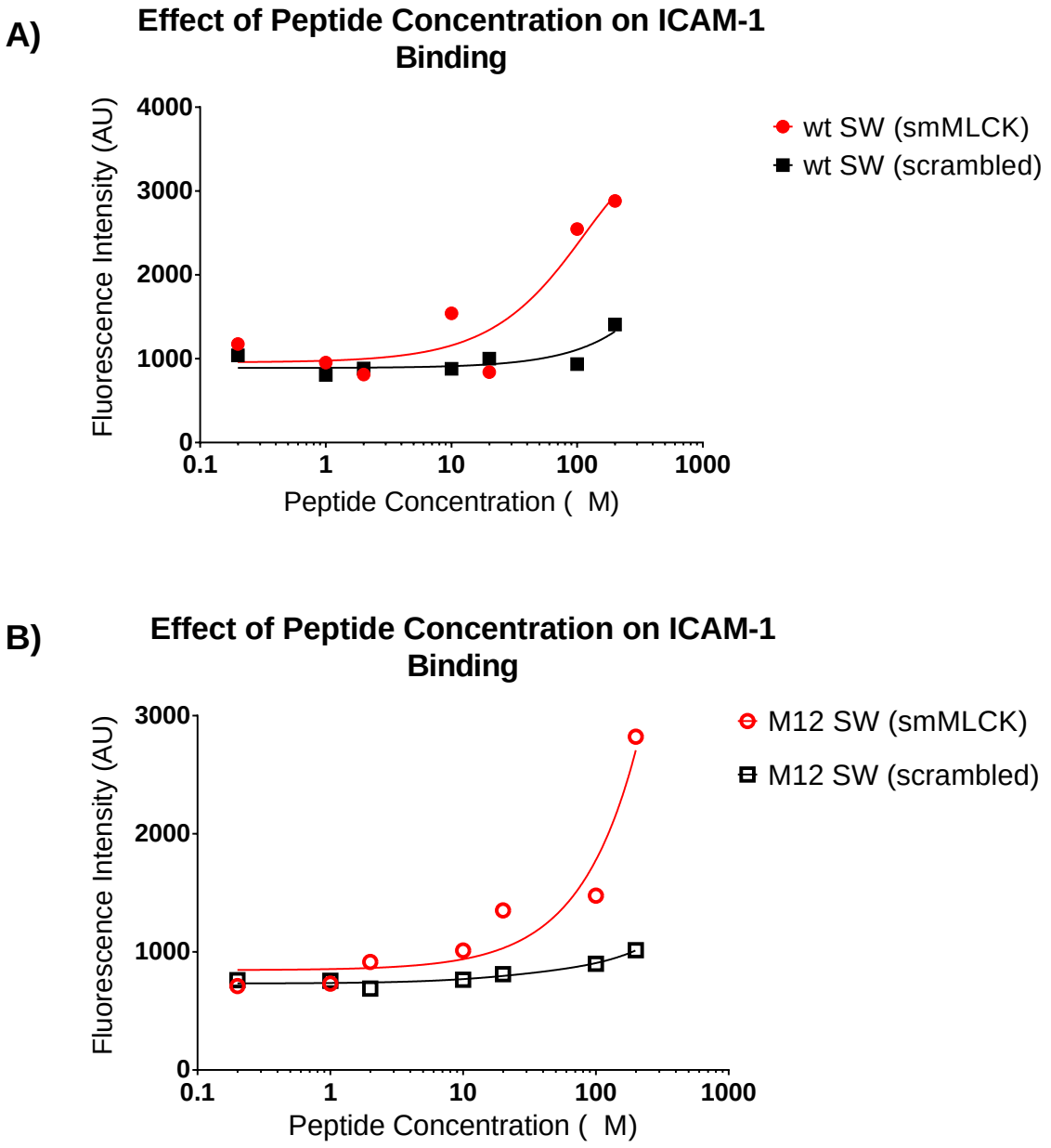


Figure 3.7 Effect of peptide concentration on switch activity. Wild-type (A) and M12 (B) switches were incubated with peptide at final concentrations between 200 nM and 200 μM and ICAM-1 binding was measured by flow cytometry. Mean fluorescence intensity from ICAM-1 binding was plotted against peptide concentration and curves were fit using Prism.

3.3.5 Soluble protein analysis

To accurately determine binding affinities for the engineered switches, as well as rule out any artifacts due to the surface display construct, soluble switches were produced for binding kinetics analysis using surface plasmon resonance. ICAM-Fc was immobilized on the sensor surface of a CM5 chip and switches in the presence of 10 μM smMLCK or scrambled peptide were flowed over the senor surface and ICAM-1 binding was measured in real time. Affinity measurements for a first generation switch, Switch 9, was found to be 132 μM in the presence of smMLCK peptide and 550 μM in the presence of scrambled peptide (Table 3.3 and Figure A10.1). This 4.2-fold change in ICAM binding in the presence of the effector peptide, smMLCK, is consistent with what has been observed by flow cytometry with this switch. Additionally, by inspection of ICAM binding by flow cytometry reveals that this switch likely has low to 'intermediate' ICAM-1 affinity based on comparison between high affinity and low affinity I domain controls. Additionally, comparison of the k_{on} and k_{off} for this switch are consistent with previously measured association and dissociation rate constants for different I domain mutants regarded as having intermediate ICAM affinities.[93] I domain mutants from this study had reported k_{on} between 133 x 10^3 and 3.4 x10^3 M^{-1} s^{-1}, k_{off} between 0.43 and 1.2 s^{-1}, and K_D between 3.0 and 450 μM for intermediate to low-intermediate affinity I domain mutants.[93]

3.4 Discussion

Previous work has developed an allosteric switch by fusing the EF3 and EF4 hands of calmodulin to the N- and C-termini of I domain, respectively, resulting in a chimeric fusion protein with a modest two-fold increase in ICAM-1 binding in presence of smMLCK peptide .[89] This work sought to improve the switch activity of the initial switch construct, herein referred to as the wild-type switch, by improving the off-state of the fusion protein through directed evolution (i.e., minimize ICAM-1 binding in the absence of peptide/presence of scrambled peptide) while retaining ICAM-1 binding at

Table 3.3 Binding kinetics of soluble switches

Switch	Peptide	k_{on} ($M^{-1} \cdot s^{-1} \times 10^{-3}$)	k_{off} (s^{-1})	K_D (μM)	Fold Change
SW 9	scrambled	0.96 ± 0.5	0.48 ± 0.05	550 ± 220	
SW 9	smMLCK	2.7 ± 0.8	0.34 ± 0.05	132 ± 15	4.2
SW 1	scrambled	1.0	0.57	570	
SW 1	smMLCK	1.2	0.22	190	3.0

near wild-type levels in the presence of smMLCKpeptide. Previous studies have shown the power of combining circular permutation and domain insertion to create protein fusions that effectively couple β-lactamase activity to the presence of maltose.[97] Our first approach to improving the switch activity of the EF3-I domain-EF4 construct was to create and screen a library of circularly permuted I domains to identify a non-obvious site within the I domain structure that would be amenable for domain insertion. Interestingly, only new termini in the allosteric region of the I domain were tolerated. This finding was significant in two ways. First, it reiterated the importance of this part of the I domain in controlling ICAM-1 binding through allosteric rearrangements in the I domain, reaffirming previous findings.[28; 93] Second, it provided a site to focus mutagenesis to try and introduce mutations that could improve the switch activity of an EF3-I domain-EF4 fusion. As the β6-α7 loop is displaced by the downward movement of the α7 helix, three residues, Leu-295, Phe-292 and Leu-289, located along the β6-α7 loop occupy a hydrophobic pocket in succession as the I domain goes from a closed, to an intermediate, to an open conformation, respectively. Additionally, the movement of the α7 helix and the β6-α7 loop disrupts the hydrophobic core of the protein along with a network of hydrogen bonds within the central β-sheet structure of the I domain, this disruption causes rearrangements of loops containing MIDAS coordinating residues bringing the MIDAS residues into a high affinity conformation and completing transmission of the allosteric signal from the allosteric region to the ICAM-1 binding site of the MIDAS.[28; 93; 102] Prior to performing focused random mutagenesis on the allosteric region of the protein, a fusion was constructed where the α7 helix of the I domain was replaced by a short GGG linker connecting the EF4 hand directly to the allosteric region of the I domain to hopefully cause a greater disruption of the allosteric region upon peptide binding. However, this construct resulted in negligible improvement in switch activity necessitating the need for focused random mutagenesis (data not shown).

The mutation, Leu-289-Arg, which appeared in four of the clones, three of which had the identical double mutation of Leu-289-Arg and Tyr-307-Asp, is critical for improved switching. Upon examination of the low affinity crystal structure of the αL I domain Arg-289 likely forms a salt bridge with residue Glu-301 (**Error! Reference source not found.**). A salt bridge prevents the transition from a low affinity conformation to a high affinity conformation, thus locking the α7 helix in place and preventing the I domain from binding ICAM-1. Clone M8, which contained the single mutation L289R did indeed show negligible ICAM-1 binding confirming that Leu-289-Arg does keep the I domain in a low affinity conformation.

Two amino acids that are beneficial to improved switch activity, aspartic acid and serine, were found in place of the wild-type tyrosine at position 307. Both aspartic acid and serine are much smaller than tyrosine and position 307 is the last I domain residue prior to a short linker and the EF4 hand on the C-terminal end of the fusion. The smaller side chains of aspartate and serine may provide the EF4 hand more flexibility allowing the EF4 to adopt a large range of conformations to improve smMLCK peptide binding. Furthermore, it appears that shorter is better. Comparing mutants where the only difference is whether Tyr-307 has been changed to an aspartate or serine, mutants with the shorter serine side chain mutation have marked improvement in switch activity.

The two mutations associated with the EF4 hand that improve the switch activity of the fusion are M337V and K341E. Each of these mutations probably contributes to a number of small beneficial interactions and cumulatively these interactions increase the affinity of the EF4 hand for the effector peptide. The smaller side chain of valine likely allows for tighter packing of Trp-799 of the peptide into the hydrophobic pocket of the EF4 hand strengthening the hydrophobic interaction between the peptide and EF4. Similar electrostatic interactions between the mutated Lys-341-Glu and the positively charged N-terminus of the peptide contribute to improved binding of the EF-hands to the

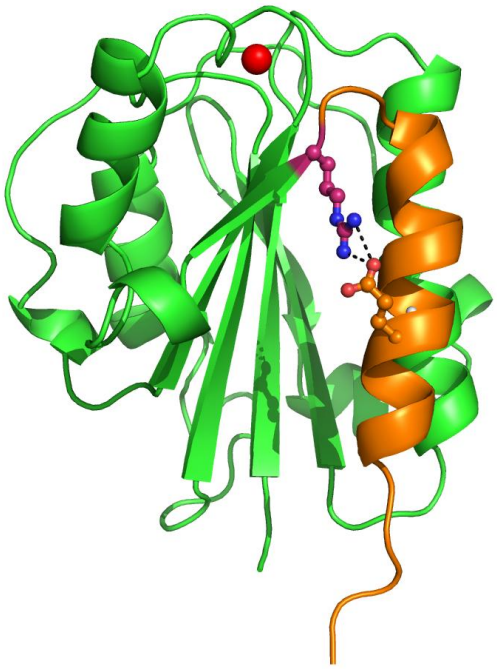


Figure 3.8 L289R mutation creates an ion pair with Glu-301. The Leu-289-Arg mutation was modeled into the low affinity I domain crystal structure (PDB 1LFA) using Swiss-PdbViewer (**http://www.expasy.org/spdbv/**)[103] and the rotamer with the lowest score was selected. The L289R mutation was then analyzed in PyMol and determined to make an ion pair with Glu-301 of the α7 helix. Amino acid side chains are shown as ball and stick with Arg-289 in purple and Glu-301 in orange.

peptide. The N-terminus of smMLCK peptide is littered with positively charged residues providing ample opportunity for attractive charge-charge interactions to occur between the positively charged N-terminus of the peptide and the negatively charged glutamic acid. The magnitudes these mutations have on improved switch activity are likely very similar as neither M337V or K341E clearly outperforms the other mutation. Thus, the effect of either of these mutations in combination with Y307S and L289R improve the switch activity of the protein.

Taken together, mutant M12, which contains L289R, Y307S and M337V mutations, was consistently the best performing switch. The effects of the different mutations seem to indicate that L289R provides a stabilizing force maintaining the I domain in a closed conformation. Presumably, binding of the EF3 and EF4 hands to smMLCK peptide sufficiently disrupts the salt bridge between L289R and E301 allowing the I domain to relax into a higher affinity conformation. The overall effect is that, in the presence of scrambled peptide, the EF3 and EF4 hands do not associate strongly enough to disrupt the Arg-289–Glu-301 salt bridge providing a true, low affinity state for the switch. In the presence of smMLCK peptide, the EF3 and EF4 hands are able to associate strongly enough to disrupt the salt bridge and allow the switch to adopt a higher affinity conformation.

We have demonstrated the ability to use directed evolution to improve the switch-like functionality of a fusion between the LFA-1 I domain and the EF3 and EF4 hands of calmodulin. Previous work had developed this fusion as the first example of a protein switch where the controlled output function was adhesion. The use of circular permutation allowed for the identification of a key structural region of the I domain, the allosteric region, providing some rational design guidance in subsequent directed evolution work. Through focused mutagenesis on the switch-allostery region of the protein a complementary set of mutations both in the allosteric region of the I domain

and the EF4 hand of calmodulin contribute to a near 20-fold improvement over the original, fusion construct. Particular detail was paid to identifying mutations that provided the switch an improved ‘off-state’, a key in engineering any protein switch, and identification of the L289R mutation was critical to improving the off-state of the protein and providing an improved dynamic range of switch activity over the original switch. The ability to use directed evolution to identify mutations to rapidly improve desired functionality of any protein, as well as to probe the structure-function relationship and provide guidance for protein engineering, further illustrates the power of this approach.

Chapter 4 : Directed evolution for altered peptide specificity

4.1 Introduction

Protein engineering has evolved from trying to better understand enzymes to a field focused on not only gaining insight into biological processes but also developing proteins for industrial and therapeutic applications.[104; 105] Protein allostery, the reversible conformational change of a protein by regulation of an effector molecule at a site distinct from the active site of the protein, has been a major theme in understanding gene regulation, metabolic pathways, and protein structure and function over the last 50 years. [10; 106; 107] In fact, a number of current drugs, such as valium, exploit allosteric perturbations, thereby 'locking-in' a specific conformation and providing a therapeutic effect to the cell.[108] Allosteric drugs have the ability to modulate protein activity without completely eliminating it, can be used in combination with drugs that target other sites to provide an increased, beneficial effect, and tend to be highly specific for their binding site.[109] Specifically, drug discovery research efforts have demonstrated that G-protein-coupled receptors (GPCRs) can be positively and negatively modulated by synthetic ligands.[110; 111] Additionally, the ability to provide requisite effector molecules for an allosteric protein to function provides another layer of specificity that could be advantageous in a drug delivery or molecular imaging capacity.

Our lab has previously demonstrated the first example of a chimeric, adhesive switch that increases affinity in the presence of an activating peptide.[89] The I domain, located at the membrane distal end of the α subunit of the integrin LFA-1, contains the entire ligand binding site of the macromolecule. Through inside-out signaling the integrin is activated and begins to refold from its low-affinity, bent conformation into an extended conformation resulting in the axially downward displacement of the C-terminal, α7 helix of the I domain by about 10 Å. The downward movement of the α7 helix of the I domain acts as a pulley causing a rearrangement of the metal ion-dependent adhesion

site (MIDAS) drastically altering the affinity for the ligand, ICAM-1.[9; 11; 92; 95] The I domain served as an ideal candidate for a protein switch as the natural allosteric movement of the protein, typically induced by the surrounding domains of the LFA-1 integrin, could be induced by a second protein fused to the I domain. This was accomplished by fusion of the EF3 and EF4 hands of calmodulin to the N- and C-termini of the I domain, respectively.

Calmodulin (CaM) is an ubiquitous Ca^{2+} sensing protein that binds a number of different target proteins, including smooth muscle myosin light chain kinase (smMLCK). The TR2C domain of CaM, composed of EF3 and EF4 hands, binds smMLCK peptide with an affinity of 80 nM.[98] The EF3 and EF4 hands have also demonstrated the ability to recombine in the presence of a 20 residue peptide of smMLCK in solution.[98] The binding of smMLCK by CaM, along with a number of other peptides, has been well characterized and classified based on their sequence motif.[112; 113] The smMLCK peptide used in these studies resides in the 1-14 class of CaM target proteins and more specifically the subclass, basic 1-8-14.[112] Target peptide classes are defined by the location of consensus hydrophobic residues. In the basic 1-8-14 subclass, three basic residues precede a consensus hydrophobic residue. This is followed by another hydrophobic residue seven amino acids later, followed by a basic residue five amino acids later which is immediately followed by the final consensus hydrophobic amino acid. In Chapter 3, a scrambled version of smMLCK was used as a negative control. Importantly, the scrambled version of smMLCK contained significantly different residues at the key locations within the basic 1-8-14 motif (Table 4.1). The differences in amino acid composition at these locations between the two peptides allowed for the binding of smMLCK and the inability to bind scrambled peptide, and it is these differences that will be exploited to alter the peptide specificity of our switches.

Table 4.1 Peptide sequences

Peptide Name	Sequence*
smMLCK	ARRK**W**QKTGHA**V**RAIGR**L**SS
scrambled smMLCK ("scram")	LVAR**H**IKSSTR**R**AGGKA**W**RQ

* Basic 1-8-14 sequence (β β β φ – – – – – – φ – – – – β φ – – –), where β represents a basic residue and φ represents a hydrophobic residue. The key hydrophobic locations are shown in **bold**.

The initial fusion construct, referred to as wild-type switch, had a near 2-fold change in ICAM-1 binding in the presence of the smMLCK peptide compared to ICAM-1 binding in the absence of peptide.[89] The protein switch was then further engineered using directed evolution in Chapter 3 to improve the switch activity of the protein 20-fold over the original switch. To make an adhesive switch of this nature more viable for eventual applications the ability to alter the specificity of the effector peptide is critical. To this end, we show that it is possible via directed evolution to alter the peptide specificity of a previously engineered I domain-EF hands switch such that the switch now responds to peptide that had previously served as a negative control. Demonstrating the ability to manipulate a complicated protein switch to respond to different effector peptides makes it amenable for a wide array of applications and environments where the specificity of the effector peptide could be tailored to a specific application.

4.2 Materials and Methods

4.2.1 Yeast strains and plasmids

S. cerevisiae EBY100 (*MAT***a** *AGA1*::*GAL1-AGA1*::*URA3 ur*a*3-52 trp1 l*e*u2*Δ*1 his3*Δ*200 pep4*::*HIS*3 *prb1_1.6R ca*n*1 GAL*) cells were transformed with pCT302-derived vectors allowing for selection under Trp- and Ura- conditions and surface expression of proteins under the galactose inducible GAL1,10 promoter.[99] pCT302 was used as the expression vector and fusion constructs were cloned in using NheI and BamHI restriction sites. Use of NheI and BamHI restriction sites allows two epitope tags (HA or c-myc) on either side of the gene of interested to be expressed in the ORF. Yeast were grown in the semi-defined media SD-CAA (20g/L dextrose, 6.7 g/L yeast nitrogen base without amino acids, 5.4 g/L Na_2HPO_4, 9.7 g/L $NaH_2PO_4{\cdot}H_2O$, and 5 g/L CAA) at 30°C overnight. The following day, expression was induced by subculturing at an OD_{600} of 1 and grown overnight in SGR-CAA (same as SD-CAA

except 2% (w/v) dextrose is replaced with 2% (w/v) galactose and 2% (w/v) D-Raffinose).

4.2.2 Flow cytometry

Yeast were grown and induced as previously described in section 3.2.1. Following overnight induction in SGR-CAA at 30°C, $3x10^6$ cells (OD_{600} of 0.3) were transferred to microcentrifuge tubes for washing and immunofluorescent labeling. Yeast were washed twice with 500 µL Labeling Buffer (137 mM NaCl, 2.7 mM KCl, 10mM Na_2HPO_4, 2mM KH_2PO_4, 10 mM $MgCl_2$, and 1 mM $CaCl_2$). Following the washes, cells were incubated with mouse anti-c-myc antibody 9E10 (ascites fluid, Covance, Princeton, New Jersey) at a 1:100 dilution in Labeling Buffer for 45 to 60 minutes at 4°C along with either smMLCK peptide or a scrambled version of smMLCKp (henceforth referred to as 'scrampep') at a concentration of 10 µM. Yeast were washed once with 500 µL of Labeling Buffer then resuspended in 30 µL for labeling with fluorescent secondary antibodies and precomplexed ICAM-Fc/anti-Fc AF647. 1.5 µL of ICAM-Fc per sample (200 µg/mL) (R&D Systems, Minneapolis,MN) was precomplexed with Alexa Fluor 647-conjugated goat anti-human anti-Fc secondary antibody (Jackson ImmunoResearch, West Grove, PA) at a ratio of 1:1.6 at 4°C for at least 60 minutes. Yeast were costained with precomplexed ICAM-Fc/AF647 anti-Fc, at a final ICAM-Fc concentration of 10 µg/mL, goat anti-mouse Alexa Fluor 488 (Invitrogen) at 1:100 and smMLCKp or scrampep at room temperature, protected from light for 1-2 hrs. Following the secondary incubation, yeast were spun down at 10,000xg for 1 minute and resuspended in 1 mL of fresh Labeling Buffer. Samples were analyzed on an Accuri C6 flow cytometer (BD Biosciences, Ann Arbor, MI). 10,000 events were collected per sample and gated based on size. Background-corrected mean AF647 fluorescence (i.e., ICAM-1 binding) was normalized by background-corrected mean AF488 fluorescence (i.e., full length protein surface expression). Normalized mean fluorescence of ICAM-binding of samples incubated with smMLCKp was divided by

normalized mean fluorescence of ICAM-binding of samples incubated with scrampep to calculate a fold-change in ICAM binding. The following equations outline the fold-change calculations:

$$NMFI = \frac{MFI647_{sample} - MFI647_{bkgd}}{MFI488_{sample} - MFI488_{bkgd}}$$

$$Fold \quad Change = \frac{NMFI_{smMLCKp}}{NMFI_{scrampep}}$$

where NMFI is the normalized mean fluorescence intensity, MFI 647_{sample} is the mean fluorescence intensity due to ICAM-binding by the sample protein , MFI647_{bkgd} is the mean fluorescence intensity of ICAM binding by a negative control population expressing a disparate protein on the yeast surface, MFI 488_{sample} is the mean fluorescence intensity of protein expression by the sample protein and MFI 488_{bkgd} is the mean fluorescence intensity of protein expression by uninduced yeast.

4.2.3 Peptide specificity library construction

Saturation mutagenesis via overlap extension PCR was used to change seven residues thought to participate in the binding of smMLCK peptide. NNK codons were used to randomize the seven amino acids in the EF3 and EF4 hands. Four primers were used to create the library (Table 4.2). Briefly, the primers encoded for mutations at key residues using NNK codons and both primers in each direction contained 20 bp of homology with each other to ensure efficient overlap of the gene PCR products. The 'outside' primers also contained 50 bp of homology to the pCT302 vector for

Table 4.2 Peptide specificity library primers

Primer Name	Mutations Encoded	Sequence*
Inside Forward	L105X	5’-cgagaggcgttccgtgtctttgacaaggatgggaatggctacatcagcgccgcagagNNKcgtcaccgaatgacgaac-3’
Outside Forward	E84X	5’-gctagtggtggtggtggttctggtggtggtggttctggtggtggtggttctgctagcagtgaggagNNKatccgagaggcgttccgtgtcttt-3’
Inside Reverse	M124X E127X A128X	5’-aaactcttcataattgacctggccatctccatcgatgtcKNNKNNcctgatKNNctcatccacctcctcatc-3’
Outside Reverse	M144X K148X	5’-ttatcagatctcgagctattacaagtcctcttcagaaataagcttttgttcggatccKNNtgcagtcatKNNctgtacaaactcttcataattgacctg-3’

* Homologous sequences between both ‘inside’ primers and both ‘outside’ primers are underlined.

homologous recombination in yeast. The 'inside' primers were used to carry out the initial PCR reaction using 50 ng of the M12 switch gene as the template DNA with 2.5 units of *Taq* DNA polymerase (NEB, Ipswich, MA), primers at a final concentration of 200 nM each, 5 µL of 10X Standard *Taq* Reaction Buffer (NEB), 1 µL of dNTPs (10 mM), 2 µL of $MgCl_2$ (25 mM), and dH_2O to a final reaction volume of 50 µL. The PCR reaction had the following steps: 1) 68°C for 2 min 2) 95°C for 30 seconds 3) 68°C for 30 seconds 4) 68°C for 50 seconds, 5) repeated steps 2-4 30 times 6) 68°C for 5 minutes 7) hold at 4°C. The PCR product was analyzed on a 1% agarose gel and purified using the Wizard SV gel and PCR clean-up system (Promega, Madison, WI) following the manufacturer's instructions. The purified PCR product was then used as the DNA template for the second PCR and was performed as just described with the only exception being the total volume of the reaction was increased to 1 mL. The second, full length PCR product was again analyzed and purified as described. Purified DNA was then concentrated using Pellet Paint Co-precipitant (EMD Millipore, Billerica, MA) following the manufacturer's instructions. Homologous recombination using 50-bp of homology on each side of the NheI and BamHI restriction sites allowed for efficient recombination of the library into the linearized pCT302 expression vector. Seven electroporations, each with 3 µg of library DNA and 125 ng of linearized vector, were pooled together and transformed yeast were allowed to recover in YPD (2% w/v dextrose, 2% w/v peptone, and 1% w/v yeast extract) for 1 hr at 30°C. The transformed yeast library was then centrifuged at 1,000xg for 3 minutes and resuspended in 10mL of SD-CAA and serial dilutions were plated. Following three days growth at 30°C colonies were counted on the serial dilution plates and library size was estimated to be $2x10^6$

4.2.4 Fluorescence-activated cell sorting

The library was sorted a total of four times using alternating rounds of positive and negative selections to enrich for yeast that displayed switching behavior in response to the presence of the scrambled peptide. The first two rounds of library

sorting were positive selections to enrich for a population that bound ICAM-1 in the presence of 100 μM scrambled peptide. The third round was a negative sort to deplete the library of yeast that bound ICAM-1 in the presence of smMLKC peptide at 100 μM. The fourth round of sorting was a positive sort to re-enrich for ICAM-1 binders in the presence of 100 μM scrambled peptide. Library expressing yeast were grown and induced as previously described. Following overnight induction at 30°C for at least 20 hours, $3.5x10^7$ cells (OD_{600} = 3.5) were aliquoted and immunostained as previously described. Briefly, cells were washed with Labeling Buffer then labeled with anti-c-myc mouse monoclonal antibody, 9E10 (Developmental Studies Hybridoma Bank, University of Iowa) at a dilution of 1:50 for 1 hour on ice. During the primary incubation, library expressing yeast were incubated with either 10 μM smMLCK peptide or 100 μM scrambled smMLCK peptide. Following the primary incubation, cells were washed with Labeling Buffer then labeled with precomplexed ICAM-1-Fc/anti-Fc AlexaFluor 647, smMLCKp at 10 μM or scrampep at 100 μM, and goat anti-mouse AlexaFluor 488 at a 1:100 dilution.

4.3 Results and Discussion

4.3.1 Library specificity construction

The key interacting residues of smMLCK peptide are Arg-797, Lys-799 and Arg-812, which are involved in the formation of salt bridges, and Trp-800, which has its large hydrophobic side chain buried in a hydrophobic pocket created by the EF3 and EF4 hands.[113; 114] Multiple residues of EF3 and EF4 make contacts with the charged and hydrophobic residues of smMLCK providing attractive electrostatic interactions and a hydrophobic pocket for Trp-800_{CaM}. Residues Leu-105_{MLCK}, Met-124_{MLCK}, Ala-128_{MLCK}, $M144_{MLCK}$ all interact with Trp-800_{CaM} via hydrophobic interactions (Figure 4.1). Glu-127_{MLCK} is thought to make salt links with Arg-797_{CaM} and the α-amonium group of Ala-796. Met-124_{MLCK} also provides interactions to the side chain of Ala-796_{CaM} and Lys-799_{CaM}, while $M144_{MLCK}$ has van der Waals contacts with Gln-801_{CaM}. Glu-84_{MLCK} of

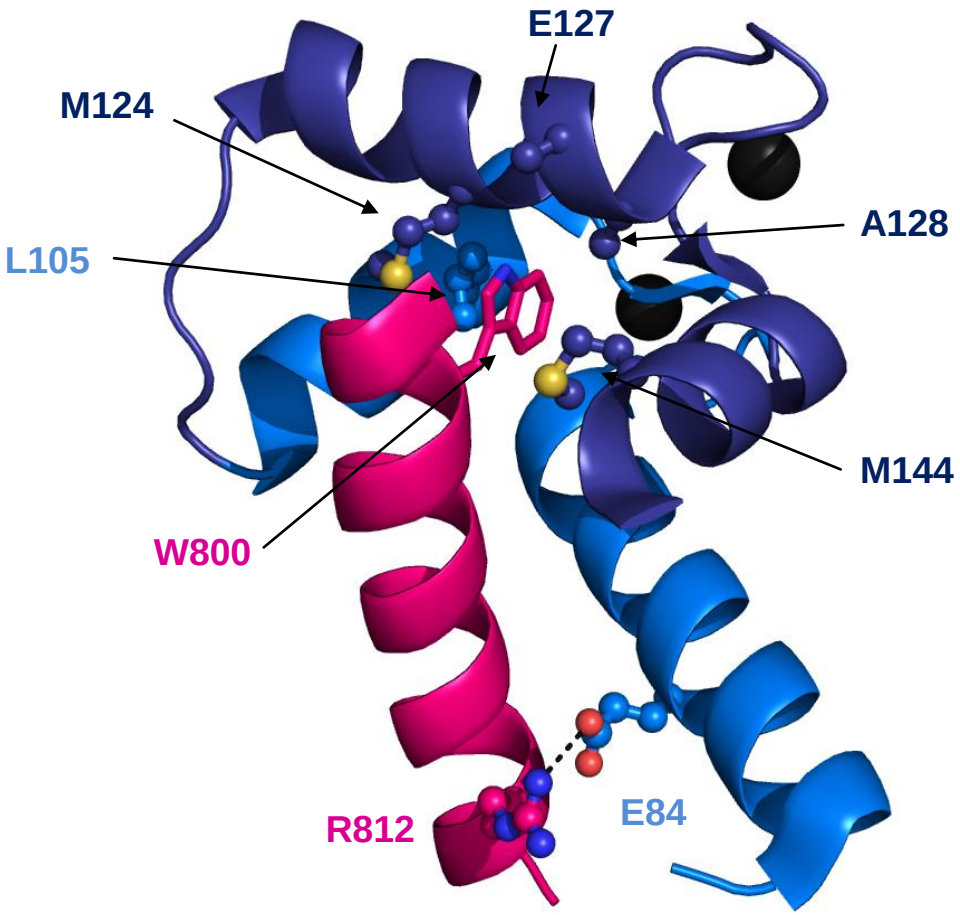


Figure 4.1 Key residues of the smMLCK and CaM interaction. Key positions chosen for site-saturation mutagenesis were chosen based on literature[112; 114] as well as analysis of the structural rendering of the interaction between the two proteins. Here, key residues thought to be important for smMLCK (magenta) binding by calmodulin (PDB 1CDL) EF3 (light blue) and EF4 (dark blue) are shown as ball and stick. Ca^{2+} bound by the EF hands are shown as black spheres. W800 of smMLCK is buried in a hydrophobic pocket created by EF3 and EF4. R812 of smMLCK forms an ion pair with E84 of EF3 of CaM.

EF3 forms a salt bridge with Arg-812 $_{CaM}$ of the peptide providing additional attractive forces for the binding of the smMLCK peptide to the EF hands of calmodulin. Finally, in previous work using random mutagenesis to improve the switch activity of our EF-I domain protein switch, Lys-148 $_{MLCK}$ was found to be a beneficial mutation in a number of clones. While Lys-148 $_{MLCK}$ is located at the extreme C-terminal end of the EF4 hand its frequency of mutation in previous experiments led to the conclusion that it likely provides some importance to binding of smMLCK peptide and was therefore included in the mutagenesis strategy presented here.

The large number of interactions contributing to the binding of smMLCK by the EF hands due to just seven amino acids, out of 64 total residues in the binding cleft created by the EF hands, seemed to indicate that heavy mutagenesis of these residues could provide sufficient enough change to be able to switch the peptide specificity of the calmodulin domain. It was reasoned that by essentially abolishing binding to the smMLCK peptide, a set of mutations allowing for binding to a scrambled version of smMLCK peptide, which had previously served as a negative control in switch experiments, could come to light. Site-saturation mutagenesis using NNK degenerative codons and nested PCR (Table 4.2) was used to try and sample all possible amino acids at each of these seven positions since it was impossible to know which residues in the EF hands would impart new specificity. The theoretical size of the library is $1.3x10^9$ which is on the cusp of yeast transformation efficiency limits.[115] Factoring the biased representation of amino acids due to the NNK codon, the actual theoretical size of the library would be greater than $1.3x10^9$ pushing the attainable size of a library outside the window of yeast transformation efficiency. The estimated size of the library produced for screening changes in peptide specificity was calculated to be $2x10^6$ therefore under-sampling the entire sequence space that would like to be explored, but likely large enough to produce mutants with the desired characteristic of altered peptide

specificity. Further directed evolution and affinity maturation could be used later to produce additional diversity and identify a set of mutations that improve the specificity of the EF hands for the scrambled peptide.

4.3.2 Library Screening

Similar to the sorting performed in Chapter 3 for improving the EF3-I domain-EF4 switch, four rounds of sorting were used to screen the library for mutants with altered peptide specificity. Alternating positive (in the presence of 100 μM scrambled peptide) or negative (in the presence of 30 μM smMLCK peptide) sorts were performed to drive the library population towards a phenotype that selectively bound ICAM-1 in the presence of scrambled peptide and didn't binding ICAM-1 in the presence of smMLCK peptide. The naïve library showed negligibly better ICAM-1 binding in the presence of scrambled peptide (31.6%) than in the presence of smMLCK peptide (30.6%) (Figure 4.2 A). Given the shift in FL4 fluorescence (X-axis) in the non-expressing population in the presence of 100 μM scrambled peptide, the slight one percent increase in ICAM-1 binding is likely due to background. Therefore, it can be assumed that the naïve library, as a whole, had no specific preference for smMLCK or scrambled peptide at the outset of sorting. The first round of fluorescence-activated cell sorting (FACS) resulted in a 5.5% increase in cells that bound ICAM-1 when incubated with 100 μM scrambled peptide compared to cells expressing switches that bound ICAM-1 when incubated with 10 μM smMLCK peptide (Figure 4.2 B) Following the second sort collecting switches that had been incubated with 100 μM scrambled peptide the library population was enriched by only about 2% for mutants that preferred scrambled peptide over switches that preferred smMLCK indicating that non-specific ICAM-1 binders had also been selected (Figure 4.2 C). A third, negative sort in the presence of 30 μM smMLCK was used to remove library members that either bound ICAM-1 regardless of the peptide present or bound ICAM-1 in the presence of smMLCK. A higher concentration of 30 μM of smMLCK peptide, compared to 10 μM which is the concentration routinely used in

Figure 4.2 Peptide specificity library at different stages of FACS. Two color immunofluorescent staining of the peptide specificity library where fluorescence intensity on the vertical axis represents the amount of full-length surface expressed library protein and fluorescence intensity on the horizontal axis represents the amount of bound ICAM-1 in the presence of the indicated peptide. A) The naïve library with no peptide, 10 μM smMLCK peptide (PEP), or 100 μM scrambled smMLCK (SCRAM). B and C) The library following the first and second rounds of sorting, respectively, in the presence of 100 μM scrambled peptide. D) The library population following a third sort in the presence of 30 μM smMLCK peptide to deselect non-switching binders and a fourth round of sorting in the presence of 100 μM scrambled peptide.

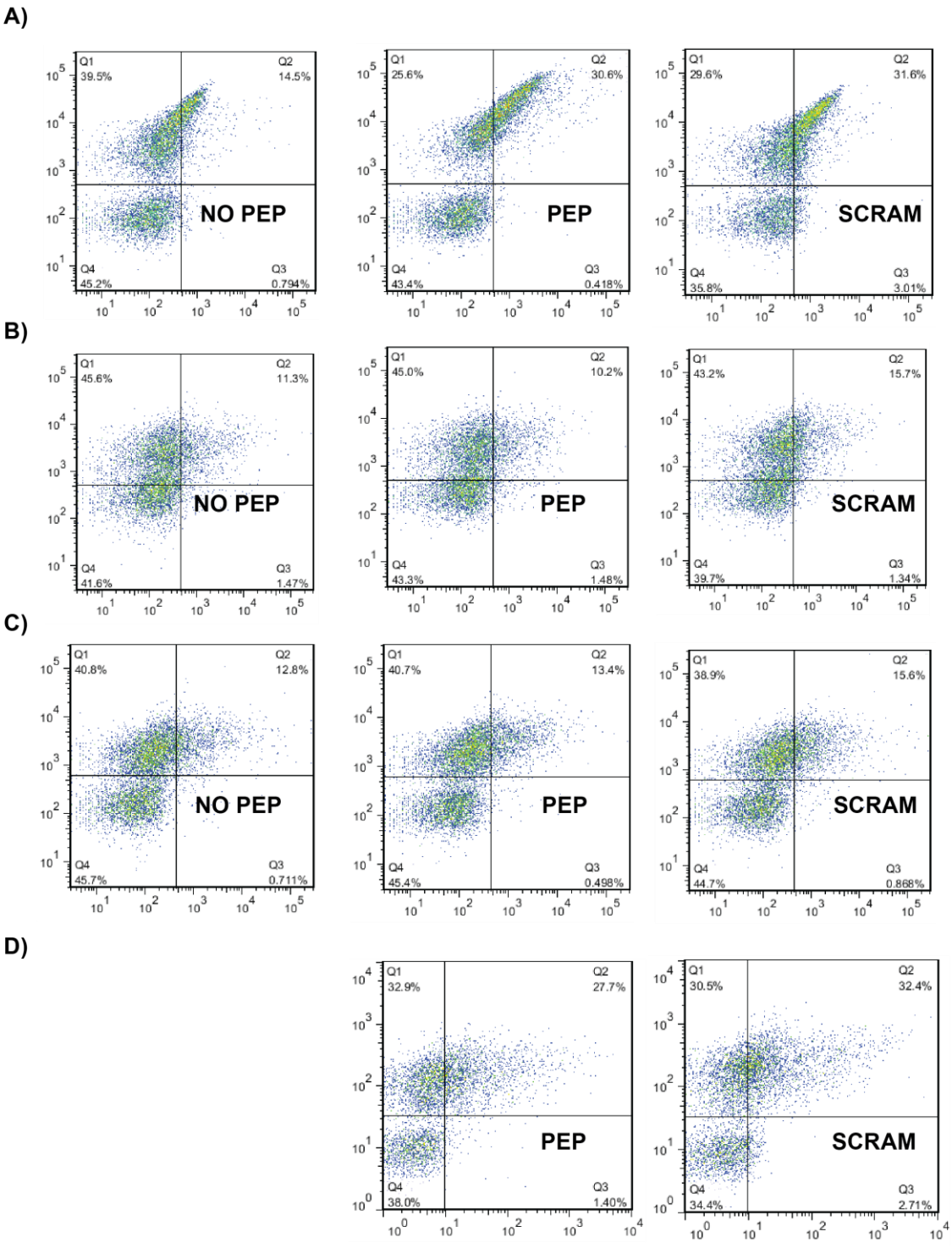


Figure 4.2 Continued

our switch assays, was used to increased the selective pressure for switches that selectively bound scrambled peptide. A final, positive sort in the presence of 100 μM scrambled peptide was used to re-enrich the library population that responded to the presence of scrambled peptide by binding ICAM-1. This final sort resulted in a near 5% increase in ICAM-1 binding when the mutant switches are incubated with scrambled peptide compare to smMLCK peptide (Figure 4.2 D). Following the fourth round of sorting, the library was plated on SD-CAA medium and random colonies were screened for switch activity in the presence of scrambled peptide.

4.3.3 Screening of individual mutants for altered specificity

A total of 58 mutants were screened for altered peptide specificity by calculating the raw (no background correction) fold change in ICAM-1 binding in the presence of scrambled peptide over the presence of smMLCK peptide (Figure 4.3). No background correction was applied in these calculations since there is minimal ICAM-1 binding in the presence of smMLCK peptide and subtraction of background fluorescence can result in negative values skewing the effectiveness of the switch. The average fold change in ICAM binding in the presence of scrambled peptide versus smMLCK was calculated for the wild-type switch and found to be 1.48. Since the wild-type switch has been shown repeatedly to respond to smMLCK peptide the expected value for switch activity for this switch should be less than unity (i.e., specificity for smMLCK over scrambled peptide). The elevated switch activity of the wild-type protein is likely due to some non-specific interactions due to the 10-fold high concentration of scrambled peptide compared to smMLCK. In other words, had both peptide concentrations been increased to 100 μM for the wild-type switch, the switch would still respond selectively to smMLCK as was seen in the peptide titration experiments from Chapter 3. Nevertheless, any switches that demonstrated a fold change greater than 1.48 were considered to have altered peptide specificity.

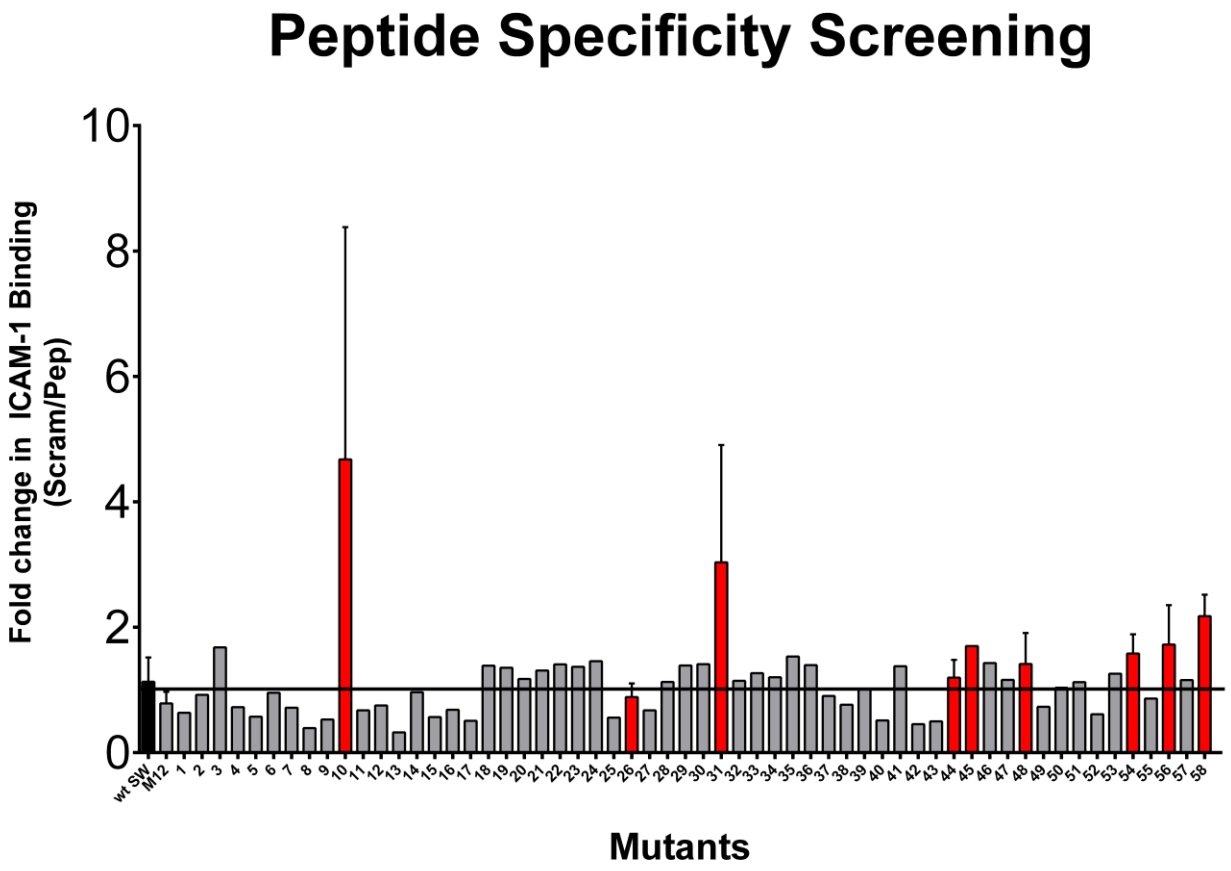


Figure 4.3 Screen of peptide specificity mutants for switch activity. Fifty-eight (58) mutants were randomly screened at least once to determine their switch activity in response to scrambled peptide. Mutants that proved promising after the initial screen (red) had their plasmids rescued from yeast and were retransformed into yeast to confirm that the phenotype transferred.

Eight switches, mutants 10, 26, 31, 44, 45, 48, 54, 56, 58, met or exceeded this threshold for switch activity and a ninth, mutant 26, looked qualitatively similar to these mutants and was also chosen for further examination. Curiously, mutants 10, 26, and 31 lacked non-expressing yeast populations (i.e., the population always located in the lower left quadrant of two-color flow cytometry plots due to plasmid loss) meaning that the gene was likely integrated into yeast genomic DNA complicating plasmid recovery. Plasmids were recovered for mutants that exhibited the negative population and phenotypes were confirmed to transfer with the plasmid.

4.3.4 Mutations affecting peptide specificity

Sequencing analysis revealed a total of 17 unique mutations in the EF3 and EF4 hands of five different mutants whose plasmids were successfully recovered. The majority (13/17) of these mutations were located at targeted residues (E84, L105, M124, E127, A128, M144, and K148) (Table 4.3). Two mutants, 54 and 58, had identical sets of eight mutations; E84D, R106C, D118N, M124F, E127L, A128H, M144I and K148Y. Additionally, four spontaneous mutations arose through the rounds of directed evolution and were identified by sequencing: E87Q, R106C, D118N and Q143R. Three out of the four spontaneous mutations are charged residues indicating the importance of electrostatic interactions in peptide binding. More interesting, however, is that both R106C and D118N always occurred together. In the native calmodulin-smMLCK complex, these two side chains which are located on EF3 and EF4, respectively, form an ion pair with each other stabilizing the complex and presumably assist in effecting the switch activity of the molecule (Figure 4.4). Taking either mutant 54 or 58 as an example of a set of mutations needed for altered specificity some general conclusions about altering specificity can be made. Both mutants had five charged residues and three hydrophobic residues at the outset of the directed evolution study. Following the four rounds of sorting for altered specificity, these mutants retained a negatively charged residue (E84D) and two hydrophobic residues (M124F and M144I) at the same

Table 4.3 Mutations found in altered peptide specificity mutants

Mutant	EF3 Mutations	EF4 Mutations
10	E87Q	Q143R
48	E84C	M124S A128G M144T K148Y
54	E84D R106C	D118N M124F E127L A128H M144I K148Y
56	E84G L105V	M144T K148R
58	E84D R106C	D118N M124F E127L A128H M144I K148Y

Note: Mutations highlighted in red arose spontaneously

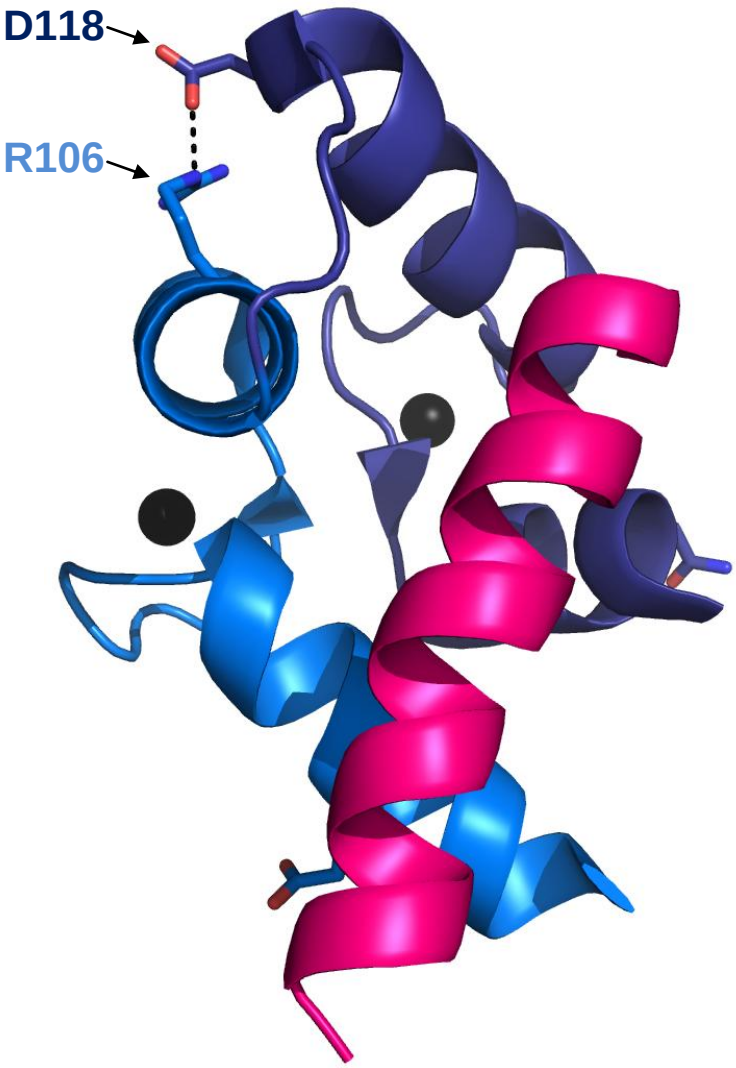


Figure 4.4 Stabilizing effect of Arg-106 and Asp-118. In native calmodulin (PDB 1CDL) complexed with smMLCK peptide (magenta), Arg-106 located on the C-terminal half of EF3 (light blue) and Asp-118 on the C-terminal half of EF4 (dark blue) form a salt bridge.

positions. A fourth mutation (D118N) by itself is a relatively conservative change, but coupled with the R106C mutation could have a significant role in altering specificity. Namely, without the additional stabilizing interaction of the Arg-106-Asp-118 salt bridge the EF3 and EF4 hands may be able to associate more freely and bind the scrambled peptide in a different orientation than the smMLCK peptide. Finally, three positions (E127L, A128H, and K148Y) experienced drastic changes going from charged residues to hydrophobic residues or vice versa. These mutations could provide adequate enough rearrangement of hydrophobic and charged residues within the EF hands to accompany the changes in distribution of charged and hydrophobic residues of the scrambled peptide. Together, the rearrangement of charged and hydrophobic residues along with the possible additional flexibility afforded by the double R106C and D118N mutations may provide the right combination of changes for the switch to bind ICAM-1 in response to the scrambled peptide.

4.4 Conclusion

To expand this technology further towards being commercially viable, it is important to demonstrate the design flexibility available in constructing a protein switch of this nature. The ability to engineer a protein switch to respond to a new effector molecule has been attempted many different times with varying degrees of success.[17; 85; 116; 117] Previously, this fusion protein had been engineered to demonstrate about a 70-fold increase in ICAM-1 binding in the presence of smMLCK peptide compared to a scrambled version of smMLCK. In this study, as a proof-of-concept, we flipped the specificity of the effector peptide for the switch protein from smMLCK to the scrambled peptide. Five mutants isolated after four rounds of directed evolution and fluorescence-activated cell sorting (FACS) showed a roughly 2-fold increase in ICAM-1 binding in the presence of scrambled peptide compared to smMLCK. While this is a very modest switching activity, additional directed evolution could be used to further improve the activity of the switch by improving the affinity of the EF3 and EF4 hands for the

scrambled peptide. Tighter binding of the scrambled peptide by the calmodulin domains would likely improve the switch function of the protein by providing a greater perturbation in the allosteric region of the I domain.

Chapter 5 : Summary and Recommendations

5.1 Summary

The ability of proteins to change their shape allows them to perform their unique functions. Here, we have studied two proteins which undergo conformational rearrangements in response to external stimuli. Hemagglutinin, a viral fusion protein found on the surface of influenza, refolds into an extended conformation at low pH. The conformational refolding of hemagglutinin results in the fusion peptide becoming exposed and likely inserting into the target membrane leading to membrane fusion. Identification of mutations that stabilize or destabilize the neutral pH structure of hemagglutinin may lead to better understanding of the structure-function relationship of the protein. Previous directed evolution work produced a series of mutants that refolded at a different pH than wild-type fowl-plague virus (FPV) HA. The mutants were analyzed to identify individual residues that were responsible for the altered conformational stability. The results indicate that the HA1 head group is vitally important in controlling acid-induced refolding as four out of the six mutations that altered stability were found in the HA1 headgroup.

The entire ligand binding site for the lymphocyte function-associated antigen-1 (LFA-1 integrin αLβ2) is contained solely within the inserted (I) domain of the α chain of the integrin. Affinity for the I domain ligand, ICAM-1, which is overexpressed on inflamed endothelium is allosterically regulated. Downward displacement of the C-terminal α7 helix causes a disruption in the hydrophobic core of the I domain leading to rearrangement of residues in the metal ion-dependent adhesion site (MIDAS) resulting in higher affinity. This allosteric rearrangement was exploited to create a chimeric fusion protein whose affinity for ICAM-1 was regulated by the presence of an exogenous peptide. Directed evolution was used to improve the functional activity of this protein switch and to alter the specificity of the activating peptide

Both proteins, hemagglutinin and the I domain-calmodulin fusion, represent the importance of conformational changes in controlling protein function. The ability to manipulate these changes underscores the power of protein engineering and directed evolution and can lead to the development of complex biomolecules that can be used in a variety of applications from biosensing to molecular imaging to drug delivery.

5.2 Engineering Protein Switches

The first aim of this dissertation was to engineer improved molecular switches via circular permutation. Circular permutation has proven a successful approach for identifying non-obvious sites for domain insertion to more closely couple the functions of two disparate protein domains in order to effect greater switch activity. Circular permutation was used to identify sites within the I domain amenable to domain insertion by sorting for permuted I domains that retained ICAM-1 binding. A population of circularly permuted I domains was enriched 17-fold for ICAM-1 binders; however, upon sequence analysis of clones from this population it was apparent that the only site that was tolerant of new termini was the allosteric region of the I domain. This information was used to then focus mutagenesis of a subsequent library on the allosteric region of protein. By focusing the mutagenesis efforts in this region, along with alternating between positive and negative sorts, the likelihood of identifying mutations that would be beneficial to improved switch activity was increased. Mutations identified through cell sorting were combined to determine a complementary set of amino acid changes that simultaneously improved the off-state of the switch (i.e., minimal ICAM-1 binding in the presence of the scrambled peptide) and retained or improved the on-state of the protein (i.e., retained or improved ICAM-1 affinity in the presence of smMLCK peptide). Mutant, M12, emerged for the directed evolution studies with a near twenty-fold increase in switch activity. This mutant contained a total of three mutations: L289R, Y307S, and M337V. The most significant being L289R, helping to stabilize the off-state of the switch. This mutation results in a salt bridge between the β6-α7 loop and the α7

helix stabilizing the closed conformation of the I domain until a sufficient enough strain on the α7 helix, resulting from the binding of smMLCK peptide, breaks this interaction allowing the I domain to bind ICAM-1 with higher affinity. The other two mutations, Y307S and M337V, likely contribute to greater conformational flexibility (Y307S) and increased peptide affinity (M337V).

Demonstrating that both domains of an adhesive protein switch can be engineered expands the potential of this technology. To that end, we demonstrated that it was possible to alter the specificity of the peptide to which the switch responds. In the original and improved constructs, smMLCK peptide acted as the agonist peptide increasing the I domain affinity for ICAM when the peptide was present. Conversely, when a scrambled version of the smMLCK peptide was present as a negative control, the I domain remained in a lower affinity state as the EF hands did not bind the peptide therefore not inducing the high affinity state. As a proof of concept, directed evolution was used to change the protein switch's response to each of these peptides. In essence, smMLCK peptide now served as a negative control while the scrambled peptide served as the agonist peptide inducing a higher affinity conformation of I domain. Directed evolution was used to isolate a number of different mutants that responded specifically to scrambled peptide.

5.3 Characterizing Protein Switches

The two proteins in this study required various degrees of characterization. Hemagglutinin, which undergoes an acid-induced conformational switch, had individual point mutants introduced to identify residues that were responsible for the altered pH stabilities. Fifteen of the sixteen mutations found through a previous directed evolution experiment and characterized in this work had never before been found highlighting the power of directed evolution studies. More importantly to this work, examination of these individual residues highlighted the importance of the HA1 headgroup in pH sensing and

conformational stability. Four mutations (I96M, G237R, W243S, and G365C) were found to destabilize the HA trimer with three of the four residues being located in the HA1 headgroup. Additionally, W243S and G365C are highly conserved residues across all HA subtypes. Two mutations (D247N and S429P) stabilizing the neutral pH structure were also identified. Again, one of the mutations, D247N, was found at the interface between two HA1 headgroups and is believed to alleviate the native charge repulsion present between D247 and E414. The majority of these pH-sensing residues are located in the HA1 headgroup signifying its importance in the conformational switching activity of hemagglutinin. The majority of previously identified pH sensing residues have been located in the HA2 domain of hemagglutinin, implicating this subunit as the critically involved in pH sensing. Our results indicate that it is very likely that the HA1 headgroup controls pH sensing and thus triggers HA refolding since it needs to be displaced from its neutral location in order for HA2 to refold to its extended low pH conformation.

5.3 Recommendations

5.3.1 Characterization of altered specificity switches

While we have shown the ability to alter the peptide specificity of the EF3 and EF4 hands of our chimeric protein switch further engineering should be pursued to improve the activity of the switch. While decreased affinity for ICAM-1 wasn't unexpected, likely due to low affinity binding of the calmodulin domains to scrambled peptide, it would need to be improved to fully exhibit the potential of engineering a chimeric, adhesive switch. Further mutagenesis using a loop shuffling approach, similar to that used to engineer high affinity fibronectin type III (Fn3) binders to lysozyme,[118] would take advantage of the mutations already found to be beneficial while expanding the diversity to identify mutants with higher affinity for scrambled peptide thus improving the switch activity. If loop shuffling failed to improve the altered peptide specificity

switches a true DNA shuffling approach, where the gene is randomly fragmented, could then be undertaken to improve the new generation of switches. Following identification of an improved, scrambled peptide specific switch, site-directed mutagenesis could be used to try and gain further insight into the new binding mechanism. Finally, binding kinetics using soluble proteins could be used for characterization consistent with previously developed switch mutants in our lab.

5.3.2 Crystal structure of the protein switch

Structural determination of wild-type switch, switch M12 and possibly an altered peptide specificity switch could greatly benefit this work. While it would be an arduous process and would require collaboration with another lab, structures of these switches in both the bound and unbound states would provide the final atomic details and complete our understanding of how these switches are functioning. Specifically, a better understanding of how peptide binding gets allosterically translated to the MIDAS to alter ICAM-1 binding could be discovered. Both x-ray crystallography and NMR would need to be used to resolve the actual structure in detail and to gain insight into the dynamics of the switch upon peptide binding.

5.3.3 Expansion of Protein switch work

Engineering an adhesive protein switch that has essentially two orthogonal states dependent upon the presence of peptide is a not only a significant achievement but also unique in the approach. To date, the vast majority of protein switches have been developed to respond to small ligands (e.g., sugars, ions, light, etc.)[17; 18; 19; 97; 119] making their use in a therapeutic application in the near future unlikely. Given that the I domain of LFA-1 is already critical to the inflammatory cascade, demonstrating the ability to engineer clinically relevant domains as fusions to the I domain to effect control of ICAM-1 binding would make this approach very attractive for a number of different applications. Previously, the αL I domain has been conjugated to the surface of

nanoparticles for targeted delivery of drug and gene therapies to cells over expressing ICAM-1 both *in vitro* and *in vivo*.[25; 26; 27] These molecules use previously engineered high affinity I domain[28] to selectively target ICAM-1 making these nanoparticles highly efficient at targeting tumors where drug delivery by the permeability and retention (EPR) effect is less efficient. Designing a protein switch where I domain affinity is regulated by the presence of exogenous peptide would provide an additional layer of specificity in targeted imaging and delivery systems. Further engineering of peptide specificity would enhance the targeting capabilities.

5.6 Conclusions

This dissertation undertook the task of engineering and characterizing two distinct proteins both of which had distinct structural rearrangements, or switching actions, which contributed to their function. Understanding and manipulating protein switches is a worthwhile endeavor for several reasons. Manipulation of protein conformational changes can lead to a greater understanding of the structure-function relationship of a protein and the identification of key structural features can provide a good starting point for future protein engineering work. Also, the ability to combine two distinct proteins and couple their functions holds tremendous potential for the design of future biomaterials. The number of strategies or applications where the coupling of two or more proteins could be exploited for more precise control of an enzymatic reaction or protein interaction is likely limitless making engineering of protein switches a powerful tool.

References

1. Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H. & Phillips, D. C. (1958). A Three-Dimensional Model for the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature* **181**, 662-666.

2. Davies, D. R. & Metzger, H. (1983). Structural Basis of Antibody Function. *Annual Review of Immunology* **1**, 87-115.

3. Perutz, M. F., Rossman, M. G., Cullis, A. F., Muirhead, H., Will, G. & North, C. T. (1960). Structure of Hæmoglobin: A Three-Dimensional Fourier Synthesis at 5.5-Å Resolution, Obtained by X-Ray Analysis. *Nature* **185**, 416-422.

4. Monod, J., Wyman, J. & Changeux, J.-P. (1965). On the Nature of Allosteric Transitions: A Plausible Model. *Journal of Molecular Biology* **12**, 88-118.

5. Baldwin, J. & Chothia, C. (1979). Haemoglobin: The structural changes related to ligand binding and its allosteric mechanism. *Journal of Molecular Biology* **129**, 175-220.

6. Jelsch, C., Mourey, L., Masson, J.-M. & Samama, J.-P. (1993). Crystal structure of Escherichia coli TEM1 β-lactamase at 1.8 Å resolution. *Proteins: Structure, Function, and Bioinformatics* **16**, 364-383.

7. Saibil, H. R., Fenton, W. A., Clare, D. K. & Horwich, A. L. (2013). Structure and Allostery of the Chaperonin GroEL. *Journal of Molecular Biology* **425**, 1476-1487.

8. Jahn, R., Lang, T. & Südhof, T. C. (2003). Membrane Fusion. *Cell* **112**, 519-533.

9. Shimaoka, M., Takagi, J. & Springer, T. A. (2002). Conformational Regulation of Integrin Structure and Function. *Annual Reviews of Biophysics and Biomolecular Structure* **31**, 485-516.

10. Monod, J., Changeux, J.-P. & Jacob, F. (1963). Allosteric Protein and Cellular Control Systems. *J. Mol. Biol.* **6**, 306-329.

11. Schürpf, T. & Springer, T. A. (2011). Regulation of integrin affinity on cell surfaces. *EMBO*, 1-16.

12. Eckert, D. M. & Kim, P. S. (2001). Mechanisms of Viral Membrane Fusion and Its Inhibition. *Annual Reviews of Biochemistry* **70**, 777-810.

13. Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J. & Wiley, D. C. (1997). Atomic structure of the ectodomain from HIV-1 gp41. *Nature* **387**, 426-430.

14. Ostermeier, M. (2009). Designing switchable enzymes. *Current Opinion in Structural Biology* **19**, 442-448.

15. Golynskiy, M. V., Koay, M. S., Vinkenborg, J. L. & Merkx, M. (2011). Engineering Protein Switches: Sensors, Regulators, and Spare Parts for Biology and Biotechnology. *ChemBioChem* **12**, 353-361.

16. Baird, G. S., Zacharias, D. A. & Tsien, R. Y. (1999). Circular permutation and receptor insertion within green fluorescent proteins. *PNAS* **96**, 11241-11246.

17. Guntas, G., Mansell, T. J., Kim, J. R. & Ostermeier, M. (2005). Directed evolution of protein switches and their application to the creation of ligand-binding proteins. *PNAS* **102**, 11224-11229.

18. Meister, G. E. & Joshi, N. S. (2013). An Engineered Calmodulin-Based Allosteric Switch for Peptide Biosensing. *ChemBioChem* **14**, 1460-1467.

19. Mathieu, V., Fastrez, J. & Soumillion, P. (2010). Engineering allosteric regulation into the hinge region of a circularly permuted TEM-1 B–lactamase. *Protein Engineering, Design & Selection* **23**, 699-709.

20. Radley, T. L., Markowska, A. I., Bettinger, B. T., Ha, J.-H. & Loh, S. N. (2003). Allosteric Switching by Mutually Exclusive Folding of Protein Domains. *Journal of Molecular Biology* **332**, 529-536.

21. O'Neill, H. C. (1992). The diversity of retroviral diseases of the immune system. *Immunology and Cell Biology* **70**, 193-199.

22. Barnes, P. J. & Karin, M. (1997). Nuclear Factor-κB - A Pivotal Transcription Factor in Chronic Inflammatory Disease. *The New England Journal of Medicine* **336**, 1066-1071.

23. Bertrand, N. & Leroux, J.-C. (2012). The journey of a drug-carrier in the body: An anatomo-physiological perspective. *J. Control. Release* **161**, 152-163.

24. Torchilin, V. P. & Lukyanov, A. N. (2003). Peptide and protein drug delivery to and into tumors: challenges and solutions. *Drug Discovery Today* **8**, 2003.

25. Kang, S., Lu, K., Leelawattanachai, J., Hu, X., Park, S., Park, T., Min, I. M. & Jin, M. (2013). Virus-mimetic polyplex particles for systemic and inflammation-specific targeted delivery of large genetic contents. *Gene Ther*, 1-11.

26. Kang, S., Park, T., Chen, X., Dickens, G., Lee, B., Lu, K., Rakhilin, N., Daniel, S. & Jin, M. M. (2011). Tunable physiologic interactions of adhesion molecules for inflamed cell-selective drug delivery. *Biomaterials* **32**, 3487-3498.

27. Park, S., Kang, S., Chen, X., Kim, E. J., Kim, J., Kim, N., Kim, J. & Jin, M. M. (2013). Tumor suppression via paclitaxel-loaded drug carriers that target inflammation marker upregulated in tumor vasculature and macrophages. *Biomaterials* **34**, 598-605.

28. Jin, M., Song, G., Carman, C. V., Kim, Y.-S., Astrof, N. S., Shimaoka, M., Wittrup, K. D. & Springer, T. A. (2006). Directed evolution to probe protein allostery and integrin I domains of 200,000-fold higher affinity. *PNAS* **103**, 5758-5763.

29. Luo, D. & Saltzman, W. M. (2000). Synthetic DNA delivery systems. *Nat Biotech* **18**, 33-37.

30. Gerasimov, O. V., Boomer, J. A., Qualls, M. M. & Thompson, D. H. (1999). Cytosolic drug delivery using pH- and light-sensitive liposomes. *Advanced Drug Delivery Reviews* **38**, 317-338.

31. White, J. M. (1990). Viral and Cellular Membrane Fusion Proteins. *Annual Reviews of Physiology* **52**, 675-697.

32. Harrison, S. C. (2008). Viral membrane fusion. *Nature Structural & Molecular Biology* **15**, 690-698.

33. Chernomordik, L. V. & Kozlov, M. M. (2008). Mechanics of membrane fusion. *Nature Structural & Molecular Biology* **15**, 675-683.

34. Colman, P. M. & Lawerence, M. C. (2003). The Structural Biology of Type I Viral Membrane Fusion. *Nature Reviews Molecular Cell Biology* **4**, 309-319.

35. Gruenke, J. A., Armstrong, R. T., Newcomb, W. W., Brown, J. C. & White, J. M. (2002). New Insights into the Spring-Loaded Conformational Change of Influenza Virus Hemagglutinin. *Virology* **76**, 4456-4466.

36. Wilson, I. A., Skehel, J. J. & Wiley, D. C. (1981). Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3Å resolution. *Nature* **289**, 366-373.

37. Armstrong, R. T., Kushnir, A. S. & White, J. M. (2000). The Transmembrane Domain of Influenza Hemagglutinin Exhibits a Stringent Length Requirement to Support the Hemifusion to Fusion Transition. *J. Cell. Biol.* **151**, 425-437.

38. Carr, C. M. & Kim, P. S. (1993). A Spring-Loaded Mechanism for the Conformational Change of Influenza Hemagglutinin. *Cell* **73**, 823-832.

39. Skehel, J. J. & Wiley, D. C. (2000). Receptor Binding and Membrane Fusion in Virus Entry: The Influenza Hemagglutinin. *Annual Reviews of Biochemistry* **69**, 531-569.

40. Copeland, C. S., Doms, R. W., Bolzau, E. M., Webster, R. G. & Helenius, A. (1986). Assembly of Influenza Hemagglutinin Trimers and Its Role in Intracellular Transport. *Journal of Cell Biology* **103**, 1179-1191.

41. Zhirnov, O. P., Vorobjeva, I. V., Ovcharenko, A. V. & Klenck, H. D. (2003). Intracellular Cleavage of Human Influenza A Virus Hemagglutinin and Its Inhibition. *Biochemistry (Moscow)* **68**, 1020-1026.

42. Russell, R. J., Gamblin, S. J., Haire, L. F., Stevens, D. J., Xiao, B., Ha, Y. & Skehel, J. J. (2004). H1 and H7 influenza haemagglutinin structures extend a structural classification of haemagglutinin subtypes. *Virology* **325**, 287-296.

43. Rachakonda, P. S., Veit, M., Korte, T., Ludwig, K., Böttcher, C., Huang, Q., Schmidt, M. F. G. & Herrmann, A. (2007). The relevance of salt bridges for the stability of the influenza virus hemagglutinin. *The FASEB Journal* **21**, 995-1002.

44. Cross, K. J., Langley, W. A., Russell, R. J., Skehel, J. J. & Steinhauer, D. A. (2009). Composition and Functions of the Influenza Fusion Peptide. *Protein and Peptide Letters* **16**, 766-778.

45. Huang, Q., Opitz, R., Knapp, E.-W. & Herrmann, A. (2002). Protonation and Stability of the Globular Domain of Influenza Virus Hemagglutinin. *Biophysical Journal* **82**, 1050-1058.

46. Thoennes, S., Li, Z.-N., Lee, B.-J., Langley, W. A., Skehel, J. J., Russell, R. J. & Steinhauer, D. A. (2008). Analysis of residues near the fusion peptide in the influenza hemagglutinin structure for roles in triggering membrane fusion. *Virology* **370**, 403-414.

47. Reed, M. L., Yen, H.-L., DuBois, R. M., Bridges, O. A., Salomon, R., Webster, R. G. & Russell, C. J. (2009). Amino Acid Residues in the Fusion Peptide Pocket Regulate the pH of Activation of the H5N1 Influenza Virus Hemagglutinin Protein. *Virology* **83**, 3568-3580.

48. Lai, A. L. & Tamm, L. K. (2010). Shallow Boomerang-shaped Influenza Hemagglutinin G13A Mutant Structure Promotes Leaky Membrane Fusion. *Journal of Biological Chemistry* **285**, 37467-37475.

49. Langley, W. A., Thoennes, S., Bradley, K. C., Galloway, S. E., Talekar, G. R., Cummings, S. F., Varecková, E., Russell, R. J. & Steinhauer, D. A. (2009). Single residue deletions along the length of the influenza HA fusion peptide lead to inhibition of membrane fusion function

Virology **394**, 321-330.

50. Tamm, L. K., Lai, A. L. & Li, Y. (2007). Combined NMR and EPR spectroscopy to determine structures of viral fusion domains in membranes. *Biochimica et Biophysica Acta* **1768**, 3052-3060.

51. Tamm, L. K. (2003). Hypothesis: spring-loaded boomerang mechanism of influenza hemagglutinin-mediated membrane fusion. *Biochimica et Biophysica Acta* **1614**, 14-23.

52. Bentz, J. (2000). Membrane Fusion Mediated by Coiled Coils: A Hypothesis. *Biophysical Journal* **78**, 886-900.

53. Bentz, J. (2000). Minimal Aggregate Size and Minimal Fusion Unit for the First Fusion Pore of Influenza Hemagglutinin-Mediated Membrane Fusion. *Biophysical Journal* **78**, 227-245.

54. Floyd, D., Ragains, J. R., Skehel, J. J., Harrison, S. C. & van Oijen, A. M. (2008). Single-particle kinetics of influenza virus membrane fusion. *PNAS* **105**, 15382-15387.

55. Daneili, T., Pelletier, S. L., Henis, Y. I. & White, J. M. (1996). Membrane Fusion Mediated by the Influenza Virus Hemagglutinin Requires the Concerted Action of at Least Three Hemagglutinin Trimers. *Journal of Cell Biology* **133**, 559-569.

56. Hess, S. T., Gould, T. J., Gudheti, M. V., Maas, S. A., Mills, K. D. & Zimmerberg, J. (2007). Dynamic clustered distribution of hemagglutinin resolved at 40 nm in living cell membranes discriminates between raft theories. *PNAS* **104**, 17370-17375.

57. Markovic, I., Leikina, E., Zhukovsky, M., Zimmerberg, J. & Chernomordik, L. V. (2001). Sychronized activation and refolding of influenza hemagglutinin in multimeric fusion machines. *Journal of Cell Biology* **155**, 833-843.

58. Lee, J. H., Goulian, M. & Boder, E. T. (2006). Autocatalytic Activation of Influenza Hemagglutinin. *Journal of Molecular Biology* **364**, 275-282.

59. Carr, C. M., Chaudhry, C. & Kim, P. S. (1997). Influenza hemagglutinin is spring-loaded by a metastable native conformation. *PNAS* **94**, 14306-14313.

60. Xu, R. & Wilson, I. A. (2011). Structural Characterization of an Early Fusion Intermediate of Influenza Virus Hemagglutinin. *Virology* **85**, 5172-5182.

61. Godley, L., Pfeifer, J., Steinhauer, D., Ely, B., Shaw, G., Kaufmann, R., Suchanek, E., Pabo, C., Skehel, J. J., Wiley, D. C. & Wharton, S. A. (1992). Introduction of Intersubunit Disulfide Bonds in the Membrane-Distal Region of the Influenza Hemagglutinin Abolishes Membrane Fusion Activity. *Cell* **68**, 635-645.

62. Kemble, G. W., Bodian, D. L., Rosé, J., Wilson, I. A. & White, J. M. (1992). Intermonomer disulfide bonds impair the fusion activity of influenza virus hemagglutinin. *Journal of Virology* **66**, 4940-4950.

63. Madhusoodan, M. & Lazaridis, T. (2003). Investigation of Pathways for the Low-pH Conformational Transition in Influenza Hemagglutinin. *Biophysical Journal* **84**, 1926-1939.

64. Daniels, R. S., Downie, J. C., Hay, A. J., Knossow, M., Skehel, J. J., Wang, M. L. & Wiley, D. C. (1985). Fusion Mutants of the Influenza Virus Hemagglutinin Glycoprotein. *Cell* **40**, 431-439.

65. Steinhauer, D. A., Martín, J., Lin, Y. P., Wharton, S. A., Oldstone, M. B. A., Skehel, J. J. & Wiley, D. C. (1996). Studies using double mutants of the conformational transitions in influenza hemagglutinin required for its membrane fusion activity. *PNAS* **93**, 12873-12878.

66. Sugrue, R. J., Bahadur, G., Zambon, M. C., Hall-Smith, M., Douglas, A. R. & Hay, A. J. (1990). Specific structural alteration of the influenza haemagglutinin by amantadine. *The EMBO Journal* **9**, 3469-3476.

67. Lee, J. H. (2007). Engineering and Characterization of Influenza Hemagglutinin Properties in Mammalian System, University of Pennsylvania.

68. Steinhauer, D. A., Wharton, S. A., Skehel, J. J. & Wiley, D. C. (1995). Studies of the Membrane Fusion Activities of Fusion Peptide Mutants of Influenza Virus Hemagglutinin. *Virology* **69**, 6643-6651.

69. Daniels, R. S., Jeffries, S., Yates, P., Schild, G. C., Rogers, G. N., Paulson, J. C., Wharton, S. A., Douglas, A. R., Skehel, J. J. & Wiley, D. C. (1987). The receptor-binding and membrane-fusion properties of influenza virus variants selected using ant-haemagglutinin monoclonal antibodies. *The EMBO Journal* **6**, 1459-1465.

70. Steinhauer, D. A., Wharton, S. A., Skehel, J. J., Wiley, D. C. & Hay, A. J. (1991). Amantadine selection of a mutant influenza virus containing an acid-stable hemagglutinin glycoprotein: Evidence for virus-specific regulation of the pH of glycoprotein transport vesicles. *PNAS* **88**, 11525-11529.

71. Ilyushina, N. A., Govorkova, E. A., Russell, C. J., Hoffmann, E. & Webster, R. G. (2007). Contribution of H7 haemagglutinin to amantadine resistance and infectivity of influenza virus. *Journal of General Virology* **88**, 1266-1274.

72. Cross, K. J., Wharton, S. A., Skehel, J. J., Wiley, D. C. & Steinhauer, D. A. (2001). Studies on influenza haemagglutinin fusion peptide mutants generated by reverse genetics. *EMBO* **20**, 4432-4442.

73. Boder, E. T., Midelfort, K. S. & Wittrup, K. D. (2000). Directed evolution of anitbody fragments with monovalent femtomolar antigen-binding affinity. *PNAS* **97**, 10701-10705.

74. Turner, N. J. (2009). Directed evolution drives the next generation of biocatalysts. *Nature Chemical Biology* **5**, 567-573.

75. Kováčová, A., Ruttkay-Nedecký, G., Haverlík, I. K. & Janeček, Š. (2002). Sequence Similarities and Evolutionary Relationships of Influenza Virus A Hemagglutinins. *Virus Genes* **24**, 57-63.

76. Fouchier, R. A. M., Munster, V., Wallensten, A., Bestebroer, T. M., Herfst, S., Smith, D., Rimmelzwaan, G. F., Olsen, B. & Osterhaus, A. D. M. E. (2005). Characterization of a Novel Influenza A Virus Hemagglutinin Subtype (H16) obtained from Black-Headed Gulls. *Virology* **79**, 2814-2822.

77. Sauter, N. K., Hanson, J. E., Glick, G. D., Brown, J. H., Crowther, R. L., Park, S. J., Skehel, J. J. & Wiley, D. C. (1992). Binding of influenza virus hemagglutinin to analogs of its cell-surface receptor, sialic acid: analysis by proton nuclear magnetic resonance spectroscopy and x-ray crystallography. *Biochemistry* **31**, 9609-9621.

78. Mittal, A., Shangguan, T. & Bentz, J. (2002). Measuring pKa of Activation and pKi of Inactivation for Influenza Hemagglutinin from Kinetics of Membrane Fusion of Virions and of HA Expressing Cells. *Biophysical Journal* **83**, 2652-2666.

79. Park, S., Xu, Y., Stowell, X., Gai, F., Saven, J. G. & Boder, E. T. (2006). Limitations of yeast surface display in engineering proteins of high thermostability. *Protein Engineering, Design & Selection* **19**, 211-217.

80. Chen, J., Lee, K. H., Steinhauer, D. A., Stevens, D. J., Skehel, J. J. & Wiley, D. C. (1998). Structure of the Hemagglutinin Precursor Cleavage Site, a Determinant of Influenza Pathogenicity and the Origin of the Labile Conformation. *Cell* **95**, 409-417.

81. Segal, M. S., Bye, J. M., Sambrook, J. F. & Gething, M. J. (1992). Disulfide bond formation during the folding of influenza virus hemagglutinin. *The Journal of Cell Biology* **118**, 227-244.

82. Biedermannova, L., Riley, K. E., Berka, K., Hobza, P. & Vondrasek, J. (2008). Another role of proline: stabilization interactions in proteins and protein complexes concerning proline and tryptophane. *Physical Chemistry Chemical Physics* **10**, 6350-6359.

83. Weber, W. & Fussenegger, M. (2006). Pharmacologic transgene control systems for gene therapy. *The Journal of Gene Medicine* **8**, 535-556.

84. Ryther, R. C. C., Flynt, A. S., Phillips, J. A. & Patton, J. G. (2004). siRNA therapeutics: big potential from small RNAs. *Gene Ther* **12**, 5-11.

85. Chockalingam, K., Chen, Z., Katzenellenbogen, J. A. & Zhao, H. (2005). Directed evolution of specific receptor-ligand pairs for use in the creation of gene switches. *PNAS* **102**, 5691-5696.

86. Guntas, G., Mitchell, S. F. & Ostermeier, M. (2004). A Molecular Switch Created by In Vitro Recombination of Nonhomologous Genes. *Chemistry and Biology* **11**, 1483-1487.

87. Strickland, D., Moffat, K. & Sosnick, T. R. (2008). Light-activated DNA binding in a designed allosteric protein. *Proceedings of the National Academy of Sciences* **105**, 10709-10714.

88. Hamorsky, K. T., Ensor, C. M., Pasini, P. & Daunert, S. (2012). A protein switch sensing system for the quantification of sulfate. *Analytical Biochemistry* **421**, 172-180.

89. Pepper, L. R., Hammer, D. A. & Boder, E. T. (N/A). Engineering the αL I Domain as an Adhesive Allosteric Switch. *manuscript*.

90. Lawerence, M. B. & Springer, T. A. (1991). Leukocytes Roll on a Selectin at Physiologic Fow Rates: Distinction from and Prerequisite for Adhesion through Integrins. *Cell* **65**, 859-873.

91. Ley, K., Laudanna, C., Cybulsky, M. I. & Nourshargh, S. (2007). Getting to the site of inflammation: the leukocyte adhesion cascade updated. *Nature Reviews Immunology* **7**, 678-689.

92. Carman, C. V. & Springer, T. A. (2003). Integrin avidity regulation: are changes in affinity and conformation underemphasized? *Current Opinion in Cell Biology* **15**, 547-556.

93. Shimaoka, M., Xiao, T., Liu, J.-h., Yang, Y., Dong, Y., Jun, C.-D., McCormack, A., Zhang, R., Joachimiak, A., Takagi, J., Wang, J.-h. & Springer, T. A. (2003). Structures of the αL I Domain and Its Complex with ICAM-1 Reveal a Shape-Shifting Pathway for Integrin Regulation. *Cell* **112**, 99-111.

94. Lee, J.-O., Rieu, P., Arnaout, M. A. & Liddington, R. (1995). Crystal Structure of the A Domain from the α Subunit of Integrin CR3 (CD11b/CD18). *Cell* **80**, 631-638.

95. Qu, A. & Leahy, D. J. (1995). Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, αLβ2) integrin. *PNAS* **92**, 10277-10281.

96. Zhang, H., Casasnovas, J. M., Jin, M., Liu, J.-h., Gahmberg, C. G., Springer, T. A. & Wang, J.-h. (2008). An Unusual Allosteric Mobility of the C-Terminal Helix of a High-Affinity αL Integrin I Domain Variant Bound to ICAM-5 $\mathcal{M}o\lambda\varepsilon\chi\upsilon\lambda\alpha\rho\ \mathcal{X}\varepsilon\lambda\lambda$ **31**, 432-437.

97. Guntas, G. & Ostermeier, M. (2004). Creation of an Allosteric Enzyme by Domain Insertion. *J. Mol. Biol.* **336**, 263-273.

98. Shuman, C. F., Jiji, R., Åkerfeldt, K. S. & Linse, S. (2006). Reconstitution of Calmodulin from Domains and Subdomains: Influence of Target Peptide. *Journal of Molecular Biology* **358**, 870-881.

99. Boder, E. T. & Wittrup, K. D. (2000). Yeast Surface Display for Directed Evolution of Protein Expression, Affinity, and Stability *Methods in Enzymology* **328**, 430-444.

100. Gietz, D. R. & Schiestl, R. H. (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols* **2**, 31-34.

101. McCusker, E. & Robinson, A. S. (2008). Refolding of G protein α subunits from inclusion bodies expressed in *Escherichia coli*. *Protein Expr Purif* **58**, 342-355.

102. Jin, M., Andricioaei, I. & Springer, T. A. (2004). Conversion between Three Conformational States of Integrin I Domains with a C-Terminal Pull Spring Studied with Molecular Dynamics. *Structure* **12**, 2137-2147.

103. Guex, N. & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* **18**, 2714-2723.

104. Brannigan, J. A. & Wilkinson, A. J. (2002). Protein engineering 20 years on *Molecular Cell Biology* **3**, 964-970.

105. Carter, P. J. (2011). Introduction to current and future protein therapuetics: A protein engineering prospective. *Experimental Cell Research* **317**, 1261-1269.

106. Lewis, M. (2013). Allostery and the lac Operon. *J. Mol. Biol.* **425**, 2309-2316.

107. Wrabl, J. O., Gu, J., Liu, T., Schrank, T. P., Whitten, S. T. & Hilser, V. J. (2011). The role of protein conformational fluctuations in allostery, function, and evolution. *Biophysical Chemistry* **159**, 129-141.

108. Hardy, J. A. & Wells, J. A. (2004). Searching for new allosteric sites in enzymes. *Current Opinion in Structural Biology* **14**, 706-715.

109. Nussinov, R. & Tsai, C.-J. (2013). Allostery in Disease and in Drug Discovery. *Cell* **153**, 293-305.

110. Jeffrey Conn, P., Christopoulos, A. & Lindsley, C. W. (2009). Allosteric modulators of GPCRs: a novel approach for the treatment of CNS disorders. *Nat Rev Drug Discov* **8**, 41-54.

111. Wood, M. R., Hopkins, C. R., Brogan, J. T., Conn, P. J. & Lindsley, C. W. (2011). "Molecular Switches" on mGluR Allosteric Ligands That Modulate Modes of Pharmacology. *Biochemistry* **50**, 2403-2410.

112. Yap, K. L., Kim, J., Truong, K., Sherman, M., Yuan, T. & Ikura, M. (2000). Calmodulin Target Database. *Journal of Structural and Functional Genomics* **1**, 8-14.

113. Chin, D. & Means, A. R. (2000). Calmodulin: a prototypical calcium sensor. *Cell Biology* **10**, 322-328.

114. Meador, W. E., Means, A. R. & Quiocho, F. A. (1992). Target Enzyme Recognition by Calmodulin: 2.4Å Structure of a Calmodulin-Peptide Complex. *Science* **257**, 1251-1255.

115. Chao, G., Lau, W. L., Hackel, B. J., Sazinksy, S. L., Lippow, S. M. & Wittrup, K. D. (2006). Isolating and engineering human antibodies using yeast surface display. *Nature Protocols* **1**, 755-768.

116. Whelan, J. & Miller, N. (1996). Generation of estrogen receptor mutants with altered ligand specificity for use in establishing a regulatable gene expression system. *The Journal of Steroid Biochemistry and Molecular Biology* **58**, 3-12.

117. Liang, J., Kim, J. R., Boock, J. T., Mansell, T. J. & Ostermeier, M. (2007). Ligand binding and allostery can emerge simulatenously. *Protein Science* **16**, 929-937.

118. Hackel, B. J., Kapila, A. & Wittrup, K. D. (2008). Picomolar Affinity Fibronectin Domains Engineered Utilizing Loop Length Diversity, Recursive Mutagenesis, and Loop Shuffling. *Journal of Molecular Biology* **381**, 1238-1252.

119. Brakemann, T., Stiel, A. C., Weber, G., Andresen, M., Testa, I., Grotjohann, T., Leutenegger, M., Plessmann, U., Urlaub, H., Eggeling, C., Wahl, M. C., Hell, S. W. & Jakobs, S. (2011). A reversibly photoswitchable GFP-like protein with fluorescence excitation decoupled from switching. *Nat Biotech* **29**, 942-947.

Appendices

Appendix 1. Site directed mutagenesis of HA primers

Table A1. 1 Primers for sited directed mutagenesis of FPV HA

Primer	Sequence
D93S	5’ – ATGCGACCAATTTCTAGAATTTTCAGCTAGTCTAATAATCGAGAGACGAG – 3’
D93S_antisense	5’ - CTCGTCTCTCGATTATTAGACTAGCTGAAAATTCTAGAAATTGGTCGCAT – 3’
I96M	5’ – TCAGCTGATCTAATAATGGAGAGACGAGAAGGA – 3’
I96M_anti	5’ – TCCTTCTCGTCTCTCCATTATTAGATCAGCTGA – 3’
G142D	5’ – GGAATAAGGACCAACGACACAACTAGTGCATGT – 3’
G142D_anti	5’ – ACATGCACTAGTTGTGTCGTTGGTCCTTATTCC – 3’
FPV HA Y155F	5’ – GGGTCTTCATTCTTTGCAGAAATGGAGTGG – 3’
FPV HA Y155F_anti	5’ – CCACTCCATTTCTGCAAAGAATGAAGACCC – 3’
D166G	5’ – CTGTCAAATACAGGCAATGCTTCTTTCCCA – 3’
D166G_anti	5’ – TGGGAAAGAAGCATTGCCTGTATTTGACAG – 3’
FPV HA A168V	5’ – GACAATACAGACAATGTTTCTTTCCCACAA – 3’
FPV HA A168V_anti	5’ – TTGTGGGAAAGAAACATTGTCTGTATTGTC – 3’
V188G	5’ – GAATCAGCTCTGATAGGATGGGGAATCCACCAT – 3’

Table A1.1 Continued

Primer	Sequence
V188G_anti	5’ – ATGGTGGATTCCCCATCCTATCAGAGCTGATTC – 3’
Y218H	5’ – GGGAGTTCCAAACATCATCAATCTTTTGTG – 3’
Y218H_anti	5’ – CACAAAAGATTGATGATGTTTGGAACTCCC – 3’
D240Y	5’ – GCCAGAGCGGACGGATTTATTTTCATTGGTTGATCT – 3’
D240Y_antisense	5’ – AGATCAACCAATGAAAATAAATCCGTCCGCTCTGGC – 3’
D247N	5’ – GATTGATTTTCATTGGTTGATCTTGAATCCCAATGATACAGTTACTTTTAG – 3’
D247N_anti	5’ – CTAAAAGTAACTGTATCATTGGGATTCAAGATCAACCAATGAAAATCAATC – 3’
G237R	5’ – AATGGCCAGTCCCGACGGATTGATTTTCAT – 3’
G237R_anti	5’ – ATGAAAATCAATCCGTCGGGACTGGCCATT – 3’
W243S forward	5’ – CGGATTGATTTTCATTCGTTGATCTTGGATCCC – 3’
W243S reverse	5’ – GGGATCCAAGATCAACGAATGAAAATCAATCCG – 3’
G365C	5’ – GAAGAAGACTTGTGGTACTGTTTCAGGCATCAGAATG – 3’
G365C_antisense	5’ – CATTCTGATGCCTGAAACAGTACCACAAGTCTTCTTC – 3’
L407A	5’ – ACCAACCAGCAATTTGAGGCAATAGATAATGAA – 3’
L407A_anti	5’ – TTCATTATCTATTGCCTCAAATTGCTGGTTGGT – 3’

Table A1.1 Continued

Primer	Sequenc
S429P	5’ – AACTGGACCAAAGACCCCATCACAGAAGTATGG – 3’
S429P_anti	5’ – CCATACTTCTGTGATGGGGTCTTTGGTCCAGTT – 3’
E462Q	5’ – TCAGAGATGAACAGGCTGTATCAGCGAGTGAGG – 3’
E462Q_antisense	5’ – CCTCACTCGCTGATACAGCCTGTTCATCTCTGA – 3’

Appendix 2. pH transition curves for pH 4.8 selected mutants

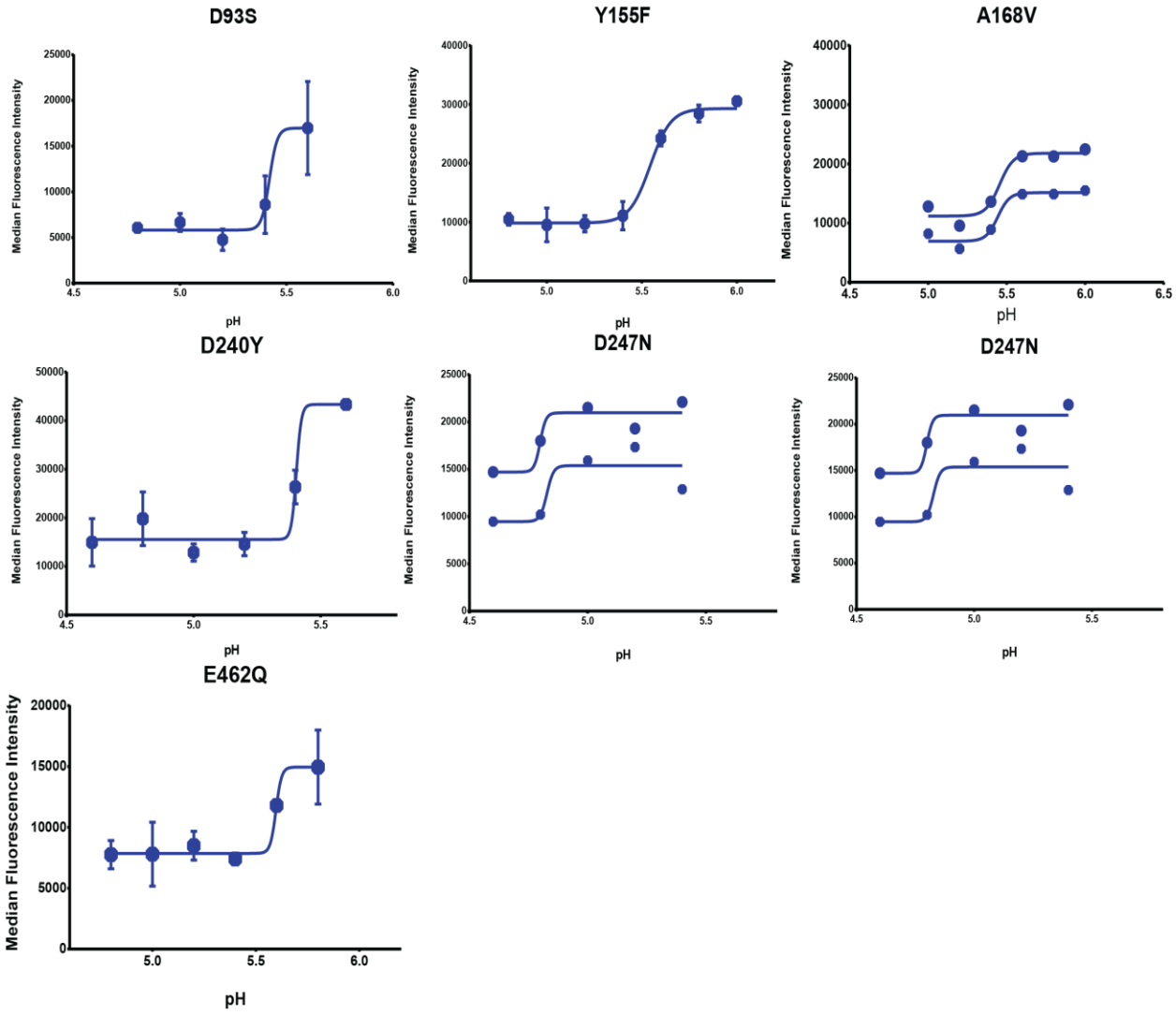


Figure A2. 1 pH transition curves for pH 4.8 selected mutants

Appendix 3. pH transition curves for pH 5.6 selected mutants

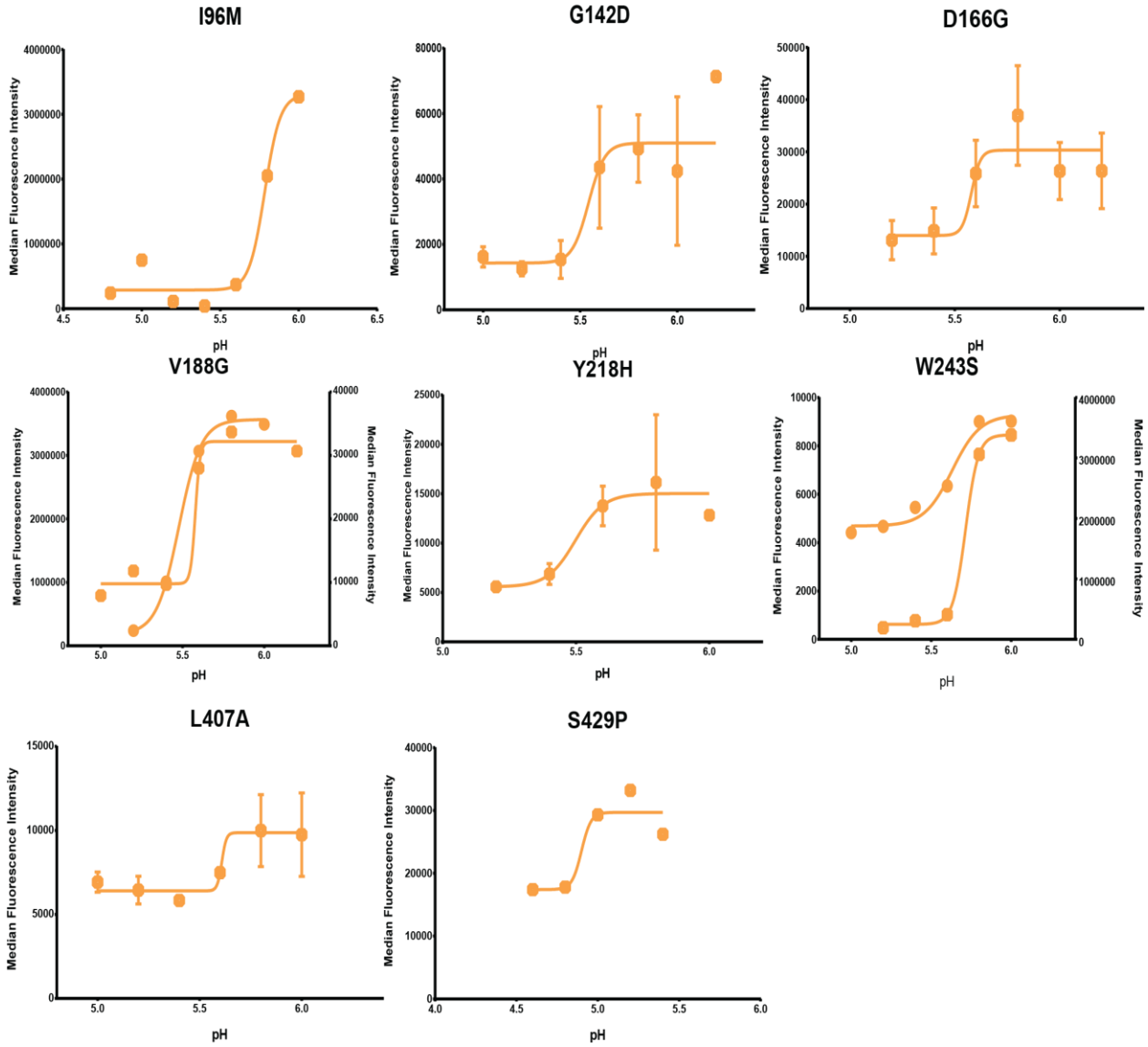


Figure A3. 1 pH transition curves for pH 5.6 selected mutants.

Appendix 4. pH transition curves for pH 6.0 selected mutants

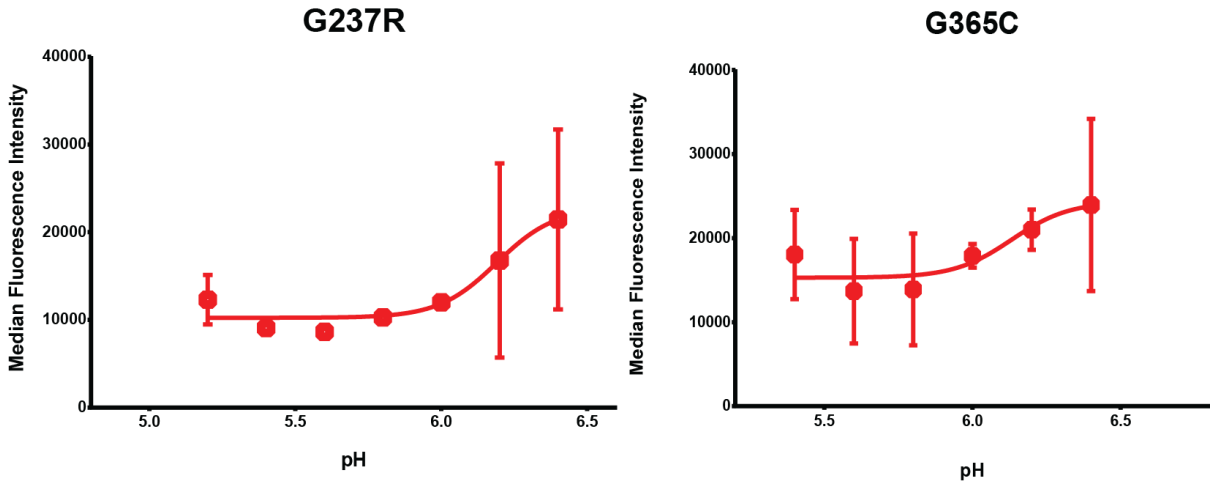


Figure A4. 1 pH transition curves for pH 6.0 selected mutants

Appendix 5. Yeast surface display construct

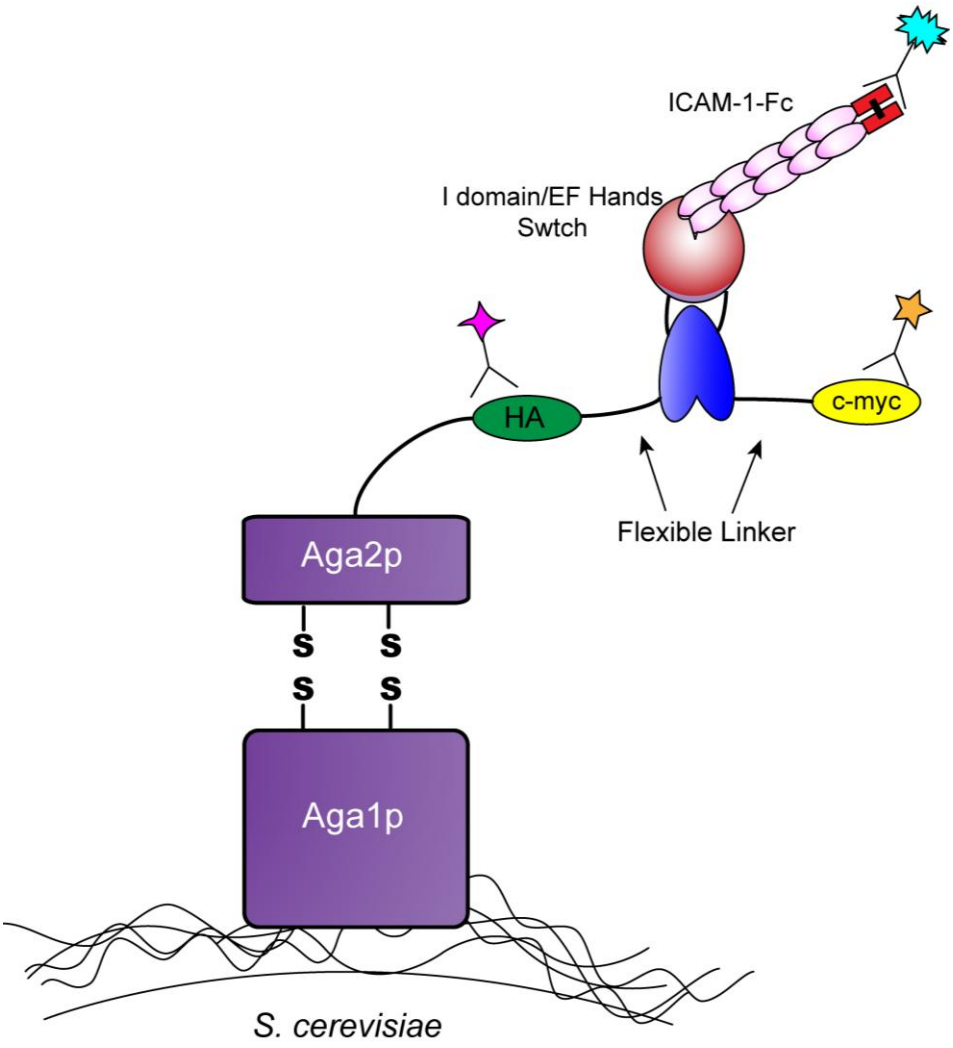


Figure A5. 1 Yeast surface display construct showing a protein switch. Proteins are displayed as C-terminal fusions to Aga2p with two epitope tags (HA and c-myc) on either side of the protein of interest. Commercial antibodies against either epitope can be used to assay surface expression level. For ICAM binding, an ICAM-Fc fusion (R&D Systems) was precomplexed with an anti-Fc AlexaFluor 647 antibody (Jackson ImmunoResearch) and the precomplexed ICAM/Anti-Fc AF647 was used to assess I domain or switch affinity.

Appendix 6. Circular permutation primers

Table A6. 3 Primers for addition of AfeI restriction site

Primer Name	Sequence
AfeI - rc A4 for	**5’ - GCTAGCAGCGCTTGTATCAAGGGCAACGTAG – 3’**
AfeI – rc A4 rev	**5’ – GGATCCAGCGCTGCCCTCAATGACATAGATCT – 3’**

Note: Primers for insertion of AfeI restriction site for blunt-end ligation into pCT302 to create pCT-JVP

Table A6. 4 Primers for the addition of GGGSG linker for circular permutation

Primer Name	Sequence
N-Id 6 Linker Cp	**5’ – GGATCCGGTGGTTGTATCAAGGGCAACGTA – 3’**
C-Id 5&6 Linker Cp	**5’ – GGATCCACCACCGCCCTCAATGACATA – 3’**

Note: Primers used to add short Gly-Ser linker when circularized with BamHI (GGATCC) restriction enzyme.

Appendix 7. Circular permutation of I domain

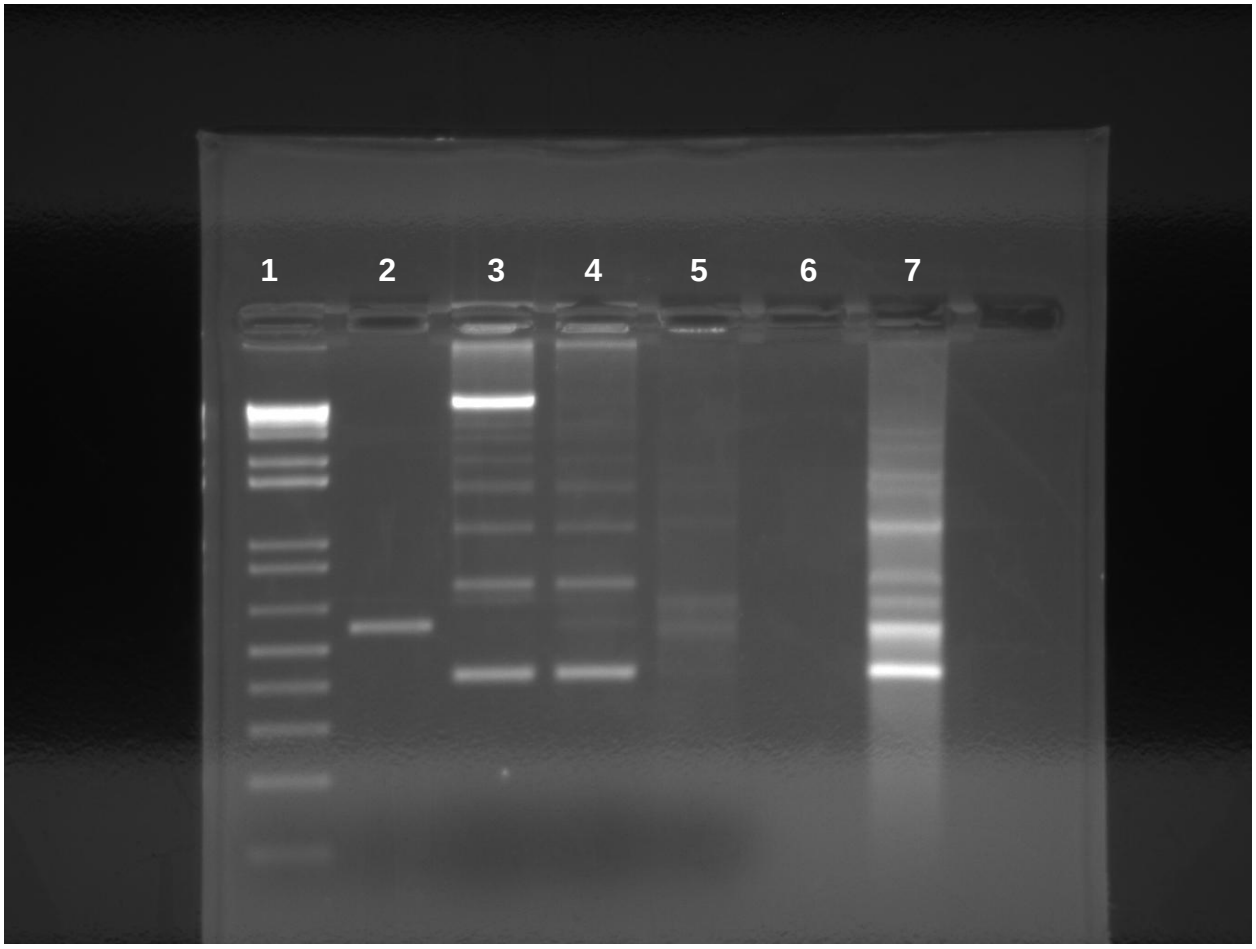


Figure A7. 1 Circular permutation of I domain. Lane 1: 1 kb+ DNA ladder, lane 2: linearized I domain, lane3: circularly ligated I domain, lane 4: ExoIII digested ligation, lane 5: DNaseI digested I domain, lane 6: empty, lane 7: T4 polymerase/ligase repair.

Appendix 8. pCT Switch plasmid

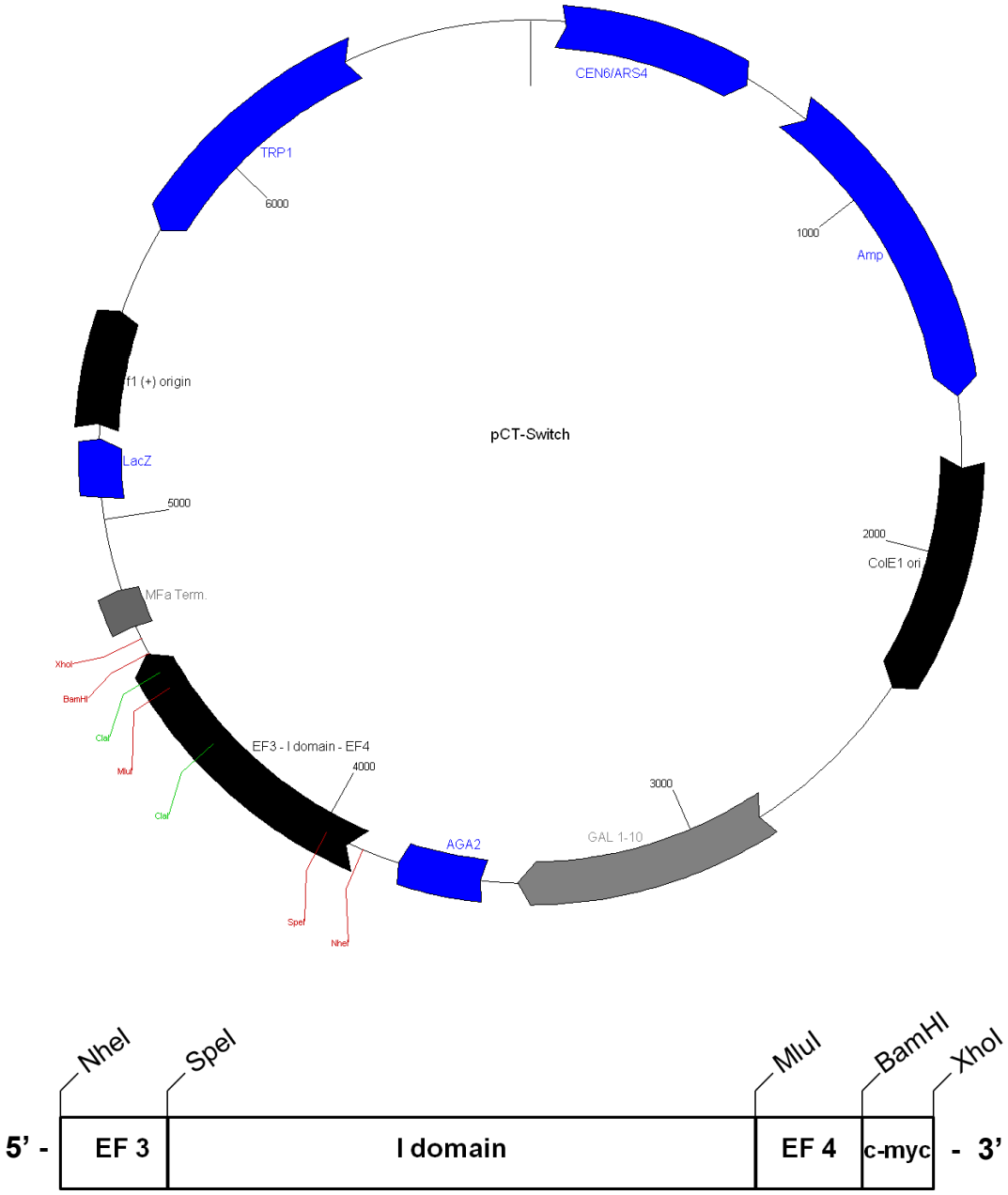


Figure A8. 1 pCT302 showing the switch gene located between NheI and BamHI restriction sites.

Appendix 9. Two color flow cytometry plots of engineered switches.

Figure A9. 1 Two color flow cytometry plots of engineered switches. In all plots, populations in blue represent switches that were incubated with scrambled peptide and populations shown in red represent switches incubated with smMLCK peptide. The vertical axis represents the amount of surface expressed protein and the horizontal axis represents the amount of bound ICAM-1. Each dot represents an individual cell. Cells located above the horizontal gate are considered to contain full length protein on the yeast surface and anything to the right of the vertical gate represents cells with bound ICAM-1.

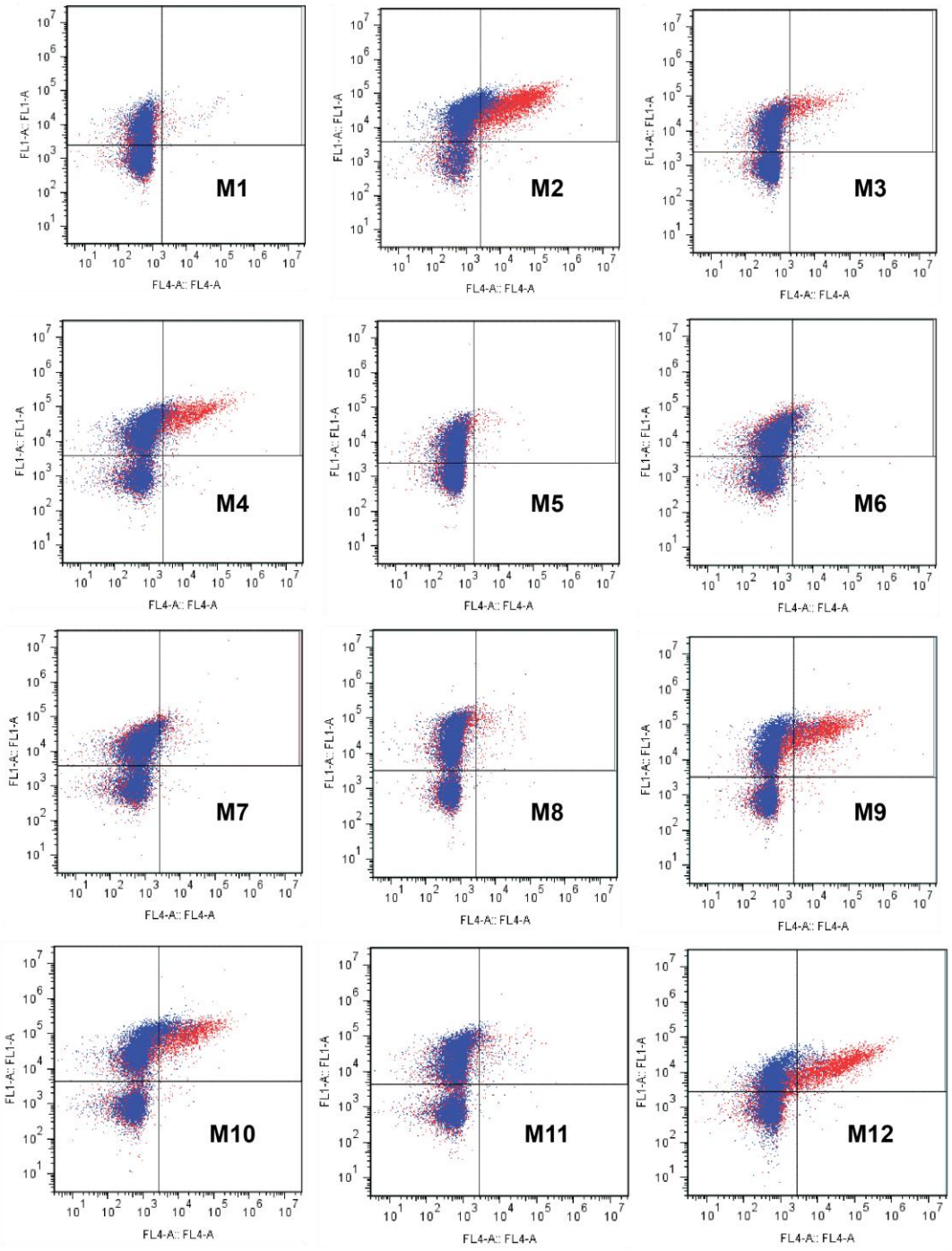


Figure A9. 2 Continued

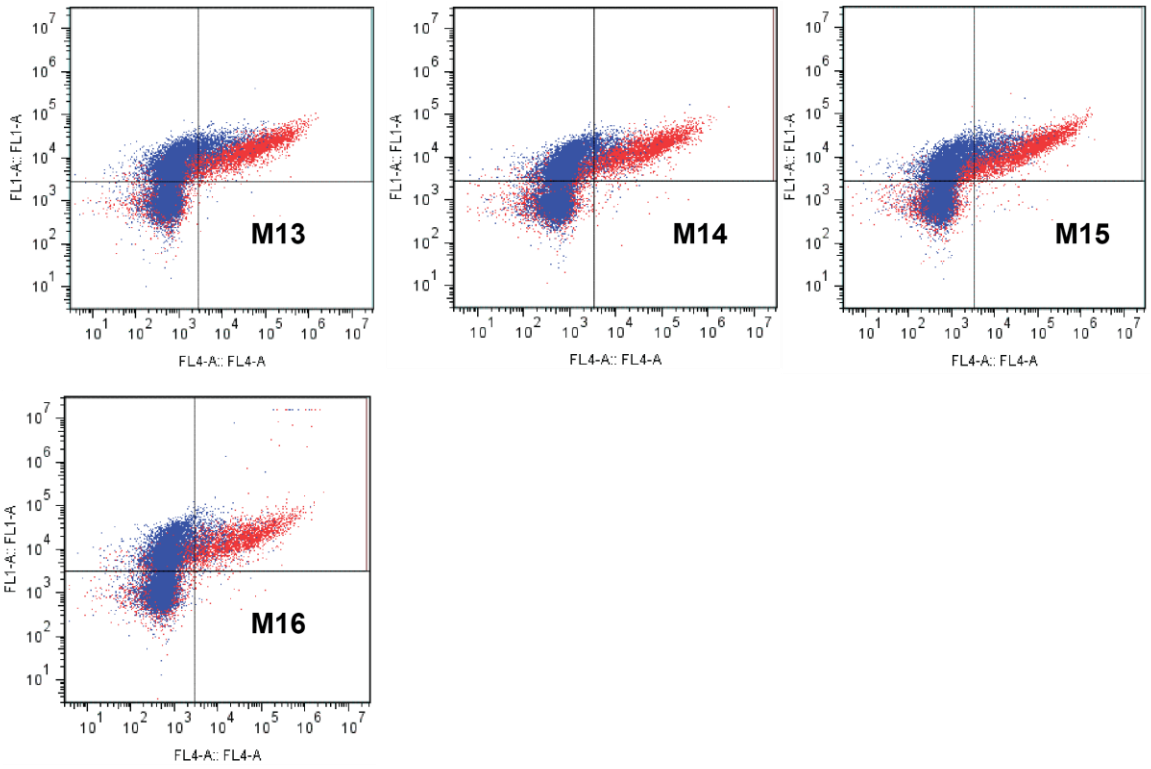


Figure A9.1 Continued

Appendix 10. Representative sensorgrams for I domain/EF hands switches.

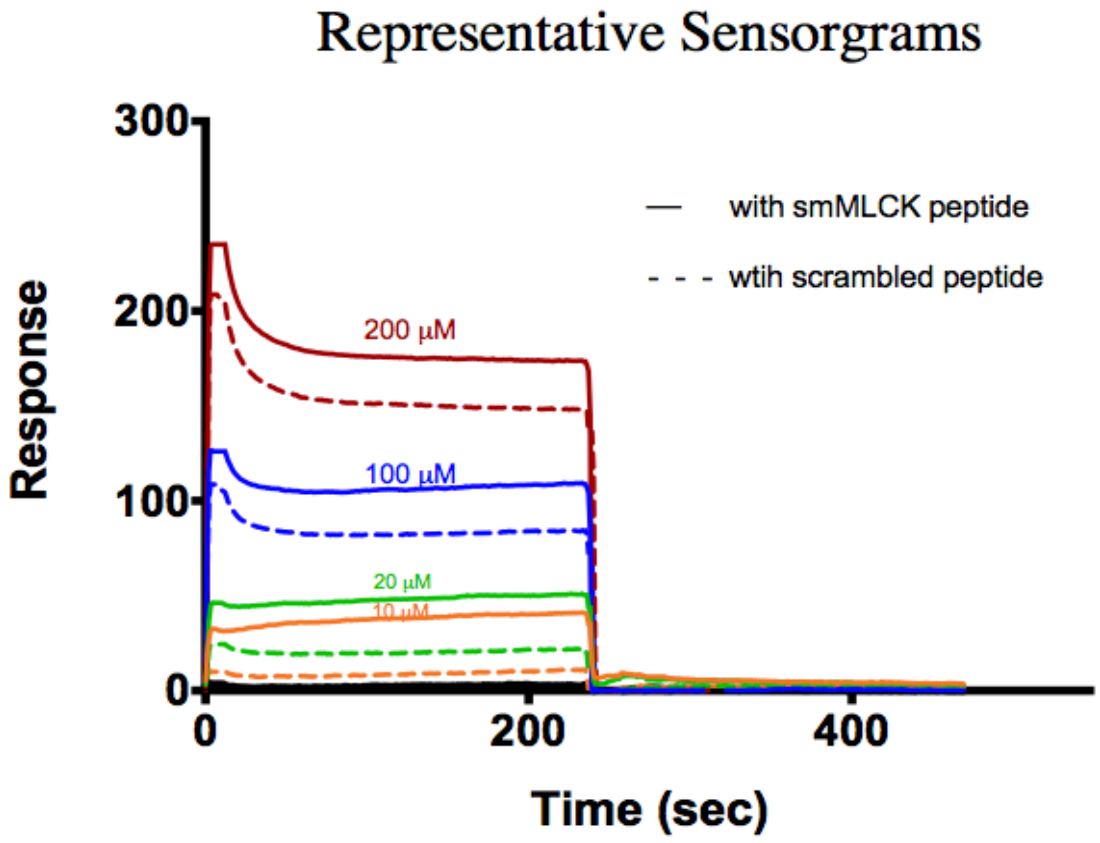


Figure A10. 1 Representative sensorgram of a first generation I domain/EF hands switch. Sensorgrams were generated using a Biacore 3000 where ICAM-1-Fc was immobilized on a CM5 sensor chip using amine coupling. First generation switches (SW1 and SW9) were incubated for at least 1 hour with either smMLCK peptide or scrambled peptide and then injected. Switches were allowed to associate for four minutes followed by a four minute dissociation phase.

Vita

James Vincent Price was born December 17, 1984 in Denver, Colorado to Roger and Kathy Price. Vince has a twin sister, Erin, born two minutes before him and a younger brother, Victor. The Price children attended St. Vincent de Paul catholic grade school from 1st to 8th grade where they made some lifelong friendships. All three then attended and participated in the International Baccalaureate (IB) Program at Littleton High School in Littleton, Colorado. While in high school, Vince was a three-sport athlete earning all-conference honors in basketball his senior year while also excelling at science and math. His aptitude for the hard sciences led Vince to enroll in the College of Engineering at the University of Colorado, Boulder. He graduated in 5 years earning both his bachelor's and master's degree in Chemical Engineering concurrently. During college, Vince interned with the process development group of Amgen in Longmont, Colorado. This experienced helped solidify Vince's desire to attend graduate school and earn his Ph.D. Ultimately, Vince decided to attend the University of Tennessee and work under the tutelage of Dr. Eric Boder whose lab focuses primarily on protein engineering. Vince has accepted a postdoctoral research position in the lab of Wilfred Chen at the University of Delaware where he will gain additional training prior to beginning a successful career in the sciences.