# STRUCTURAL MODELING OF REGULATION IN α-ACTININ/F-ACTIN INTERACTIONS

by

**Timothy Niño S. Travers**

B.S. Molecular Biology and Biotechnology

University of the Philippines Diliman Campus, 2002

Submitted to the Graduate Faculty of

School of Medicine in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2014

UNIVERSITY OF PITTSBURGH

SCHOOL OF MEDICINE

This dissertation was presented

by

Timothy Niño S. Travers

It was defended on

July 25, 2014

and approved by

Dr. James R. Faeder, Associate Professor, Department of Computational and Systems Biology, University of Pittsburgh

Dr. Daniel M. Zuckerman, Associate Professor, Department of Computational and Systems Biology, University of Pittsburgh

Dr. Yu-li Wang, Professor, Department of Biomedical Engineering, Carnegie Mellon University

Dissertation Advisor: Dr. Carlos J. Camacho, Associate Professor, Department of Computational and Systems Biology, University of Pittsburgh

Dissertation Co-Advisor: Dr. Alan Wells, Department of Pathology, University of Pittsburgh

# STRUCTURAL MODELING OF REGULATION IN α-ACTININ/F-ACTIN INTERACTIONS

Timothy Niño S. Travers, Ph.D.

University of Pittsburgh, 2014

The α-actinins (ACTNs) are a highly conserved family of actin-crosslinking proteins that are critical to various fundamental biological processes in eukaryotes, ranging from cell motility and surface remodeling to muscle contraction. Binding of ACTNs to actin filaments is regulated by several mechanisms: epidermal growth factor (EGF)-induced tyrosine phosphorylation, binding of calcium, limited proteolysis by calpain enzymes, and binding of phosphoinositide moieties. The molecular mechanisms by which these external cues drive the regulation of ACTN function are still not understood, however, largely because there is currently no high-resolution experimental structure that brings together the multiple domains that comprise ACTNs. An understanding of these molecular mechanisms should provide us with insights into how the cell is able to modulate actin cytoskeletal remodeling and give rise to complex cellular phenomena.

In this thesis, we investigate how these external cues regulate the actin-binding function of human ACTN4, a non-muscle isoform that is essential to cell motility and has been implicated in cancer invasion and metastasis. First, we develop and validate an atomic model of the multi-domain assembly that makes up the full ACTN4 homodimer, with a novel ternary complex between CH2, neck, and CaM2 comprising the core of this assembly. Next, we show that a novel tandem phosphorylation mechanism in the disordered N-terminal region of ACTN4, where phosphorylation of the functional Y31 requires prior phosphorylation at Y4, is responsible for the regulation of ACTN4 function in the presence of EGF. This tandem mechanism can work in conjunction with m-calpain cleavage of the N-terminal to generate varied actin-binding responses at the front and rear ends of the cell during motility. Using our full structural model, we also show that: (i) Y265 phosphorylation eases ABD opening: (ii) binding of calcium may break the CaM2/neck complex; (iii) CaM2 protects the neck region from m-calpain cleavage; and (iv) binding of phosphoinositides to CH2 allows ACTN4 to crosslink actin filaments at the inner membrane. Finally, we bring together these structural insights to develop a preliminary network-level model that can serve as a computational tool for predicting the actin-binding response of ACTN4 in the presence of multiple external cues.

# ACKNOWLEDGMENTS

This thesis would not be possible without many amazing people.

To my graduate mentor, Dr. Carlos J. Camacho, I sincerely thank you for all your guidance and support these past six years. Your passion for doing research, and enthusiasm in searching for answers to a broad assortment of biological problems, is simply inspirational. You never failed to encourage me, and you always have the best interests of the people you work with at heart. I am truly privileged to have learned from you.

To our collaborators on this project, Dr. Alan Wells and Dr. Hanshuang Shao, I am grateful for the years of having worked with you. Our discussions on the various aspects of this research are always thought-provoking and insightful, and it has been a wonderful back-and-forth between modeling and experiments. I am fortunate to have worked on a challenging yet exciting project.

Dr. Alan Wells also served on my thesis committee, along with Dr. James Faeder, Dr. Daniel Zuckerman, and Dr. Yu-li Wang. Before moving to new opportunities, Dr. Hagai Meirovitch and Dr. Ivan Maly were also members on my committee. All of you provided constructive critiques and feedback that contributed greatly in shaping this thesis, and to all of you, I truly thank you.

To past and present members of the Camacho lab whom I had the pleasure of working with, our discussions were always stimulating, assistance and advice were never more than a step

away, and everyone helped in maintaining a positive work environment that Carlos fostered. To Alpay, Marta, Jintao, Weiyi, Gustavo, David, Matt, Nick, and Xanthe, thank you so much and I wish you the very best in all your future endeavors.

To my family back home in the Philippines, I am very grateful for having such dear and wonderful people in my life. To my mother Marilou and my uncle Philip, whose constant and unyielding love and support helped me get to this point. To my late grandfather and grandmother, Restituto and Natalia, who were remarkable examples of how to live a modest life that is rich in spirit. To my late father George, who always placed the happiness of our family before everything else.

And to my beloved and amazing wife Charmine, thank you. Thank you for always being by my side, for constantly sharing in all the ups and downs, for never allowing me fall into despair, for letting me be a part of your life. Please don’t ever forget that in everything, through thick and thin, I’m gonna be around.

## TABLE OF CONTENTS

# LIST OF TABLES

## LIST OF FIGURES

# CHAPTER 1

# INTRODUCTION

## 1.1 BACKGROUND

One of the most abundant proteins found in eukaryotic cells is actin, which is able to form polymers comprising long and stable filamentous structures called microfilaments [1]. These microfilaments make up one of the three main kinds of cellular cytoskeleton (along with microtubules and intermediate filaments), and so it comes as no surprise that actin is essential for a variety of fundamental biological processes in eukaryotes, from cell motility and surface remodeling to muscle contraction [2, 3]. These processes require crosslinking and/or bundling of actin filaments [4], as well as their linking to other proteins, through interactions with various actin-binding proteins. At focal adhesions, for instance, actin-binding proteins such as talin, vinculin, zyxin, and α-actinin form a link between microfilaments and transmembrane integrins, thereby allowing force transmission from the actin cytoskeleton to the extracellular matrix [5, 6]. A complete picture of how actin dynamics allows for these cellular processes to take place thus requires an understanding of the precise roles that actin-binding proteins play.

The α-actinins (ACTNs) are in fact a highly conserved family of actin-binding proteins (Table 1.1), with four isoforms identified to date in mammals: ACTN2 and ACTN3 are specific to muscle cells while ACTN1 and ACTN4 are ubiquitously expressed in non-muscle cells [7-9]. All four mammalian isoforms are multi-domain proteins, and each are comprised of the following domains: an unstructured N-terminal region, an actin-binding domain (ABD), a "neck" region", a rod domain, and a C-terminal calmodulin (CaM)-like domain (Fig 1.1). Dimerization is essential for the ACTNs in order to crosslink actin filaments, and the formation of antiparallel homodimers is facilitated through the sturdy rod domain [10-12]. The rod domain is composed of a series of four spectrin-like repeats, and has a $K_D$ for dimer formation at the picomolar level [13, 14]. In addition to actin filament crosslinking, recent findings suggest that the non-muscle

**Table 1.1: Percent identity matrix for human ACTN protein sequences.**

| | **ACTN1** | **ACTN2** | **ACTN3** | **ACTN4** |
|---|---|---|---|---|
| **ACTN1** | 100 | | | |
| **ACTN2** | 79.93 | 100 | | |
| **ACTN3** | 77.90 | 79.75 | 100 | |
| **ACTN4** | 85.76 | 77.29 | 73.36 | 100 |

Full human ACTN protein sequences were taken from UniProtKB/Swiss-Prot [15], with the following UniProt IDs: P12814 (ACTN1), P35609 (ACTN2), Q08043 (ACTN3), and O43707 (ACTN4). Percent identities between the sequences were computed from a multiple sequence alignment built using the Clustal Omega alignment software v1.2.1 [16], using default parameters at the web interface provided by EMBL-EBI [17].

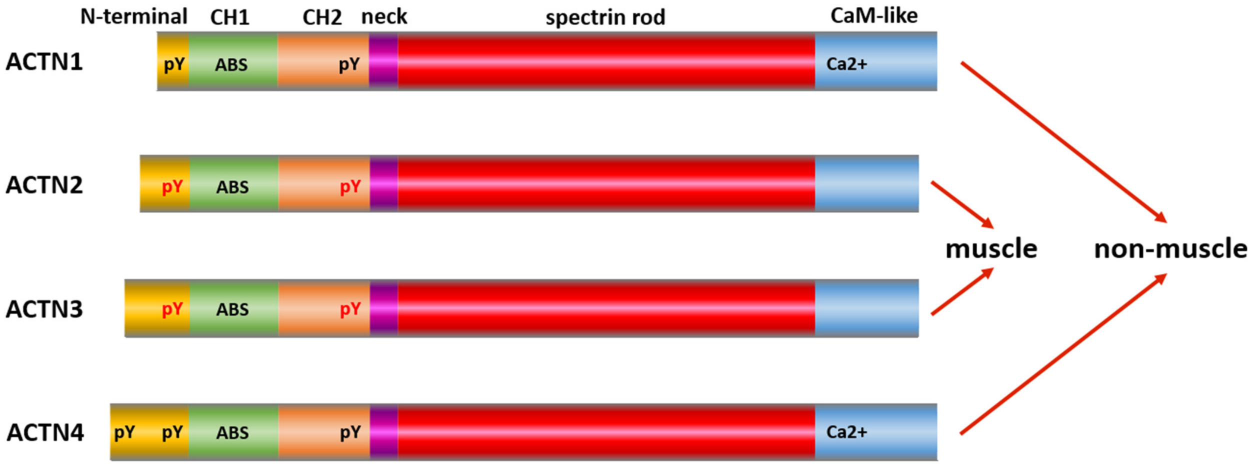


**Figure 1.1: Domain structure of ACTNs.**

Each ACTN monomer contains (from left): a disordered N-terminal region, two calponin-homology (CH) domains that comprise the actin-binding domain (ABD), a "neck" region that connects CH2 to the first spectrin-like repeat of the rod domain, the rod domains (made of four spectrin-like repeats), and a calmodulin (CaM)-like domain. The actin-binding site (ABS) is located within the CH1 domain [18]. Only the CaM-like domains of the non-muscle isoforms contain binding sites for calcium ions [19, 20]; the corresponding domains in the muscle isoforms contain point and deletion mutations that abrogate these binding sites. Tyrosine phosphorylation sites in the N-terminal region and CH2 have been confirmed experimentally for the non-muscle isoforms [21-24]; corresponding phosphorylation sites in the muscle isoforms are inferred based on sequence homology.

ACTN isoforms bridge the actin cytoskeleton to the inner cell membrane, with ACTN4 in particular being essential for the maintenance of normal cell motility [25-29] and implicated in cancer invasion and metastasis [30-35].

The binding of ACTNs to actin filaments has been shown to be regulated by several mechanisms: (i) phosphorylation at particular tyrosine residues in the unstructured N-terminal region inhibits actin binding [21, 22], (ii) phosphorylation at a tyrosine residue in the CH2 domain promotes actin binding [21], (iii) binding of calcium ions to the CaM-like domain in the calcium-sensitive non-muscle isoforms inhibits actin binding [19, 20], (iv) limited proteolysis by the ubiquitous intracellular enzymes calpain-1 and calpain-2 [36-38], and (v) binding of phosphatidylinositol moieties (e.g. PIP2 and PIP3) to a putative binding site on the ABD [39] promotes actin binding in muscle isoforms [40] but inhibits binding in non-muscle isoforms [41]. The existence of these multiple mechanisms clearly points to the complexity of regulating the actin binding function of ACTNs, especially for non-muscle cells where extracellular signals control the dynamics of the actin cytoskeleton. Unraveling the molecular mechanisms behind these various regulatory events for the ACTNs should thus allow us to attain a better understanding on how covalent and non-covalent modifications in this family of actin-binding proteins allows cellular functions such as cell spreading and motility to take place.

The research presented in this dissertation focuses on the use computational approaches, primarily through structural modeling and molecular dynamics (MD) simulation studies, to investigate the molecular mechanisms by which the binding of α-actinins to actin filaments is differentially regulated by the various regulatory events mentioned earlier. Our use of these computational techniques is driven by previous experimental data, and in turn any new concepts and models that emerge from these techniques are experimentally verified. Our experimental collaborators for this research are Dr. Alan Wells and Dr. Hanshuang Shao from the Department of Pathology, University of Pittsburgh.

## 1.2 COMPUTATIONAL METHODS

In this section, a brief survey of the various computational modeling techniques used in this research is provided.

### 1.2.1 Molecular dynamics simulations

For biomolecules such as proteins, the function carried out is closely tied to the dynamics of the molecule. While various experimental techniques can provide information on the dynamics of these molecules, each technique has its own limitations in terms of spatial and/or temporal resolution [42]. An alternative approach is to generate a computational model of the atomic-level motions of the molecule; ideally the model should be based on first-principles physics, however a direct solution to the time-dependent Schrödinger equation in order to describe the quantum mechanical behavior is in fact computationally infeasible for protein-sized molecules. To circumvent this, the standard approach currently used is to model all the atoms in the molecule as point particles, and to calculate the evolution of the positions and velocities of these point particles using the laws of classical physics. This is the basis for the computational method referred to as all-atom molecular dynamics (MD) simulations.

In MD simulations, a force field is needed to calculate the forces acting on each atom at every time step. Force fields are comprised of two components: (i) a potential energy function for computing the total force exerted on an atom by the system, with each term in the function describing a particular interaction type and based on approximations to first-principles physics, and (ii) a set of parameters that are used in this function and that are based on fitting to experimental data and smaller-scale quantum mechanical computations. For most force fields in common use today, the following interaction types are accounted for in the potential energy function: covalent bonds (two-atom) and angles (three-atom) using harmonic potentials around equilibrium values, torsional/dihedral angles (four-atom) using a cosine expansion function, nonbonded van Der Waals (VDW) interactions using a Lennard-Jones potential, and nonbonded electrostatic interactions using a Coulombic potential. Each of these parts relies on the use of atom-specific parameters that are also provided by the force field. Note that current force fields generally use the same form for describing the potential energy function, and mostly differ in the optimization protocol for deriving the atomic parameters.

One of the limitations of MD simulations is that in practice only nanosecond timescales are typically accessible, while the biomolecular events of interest (such as protein folding, ligand binding, allosteric changes) usually occur on timescales of microseconds to milliseconds. Major advances to access longer timescales with MD simulations have been made both from the algorithmic and hardware perspectives. On the hardware side, the use of graphical processing units (GPUs) has become increasingly popular for scientific computation by allowing massive parallelization of computations on a single chip. Simulations that used to require multiple CPUs/nodes on CPU-based computer clusters can now be performed on a single GPU in a desktop machine, with matching or even better performance [43]. Leveraging the parallelization capabilities of GPUs requires clever porting of existing MD implementations; one example is the development of a new precision model for the GPU implementation of the Amber MD software, where the replacement of double precision floating point arithmetic with 64-bit fixed point integer arithmetic for the accumulation of force components was found to provide higher computational throughput on GPUs as well as reducing the required memory footprint, while affording equivalent numerical stability of the calculations [44]. In this research, all MD simulations are accelerated using the Amber implementation for GPUs.

Two other limitations of MD simulations must be kept in mind when forming hypotheses based on these or when comparing these to experimental data. First is the issue of whether or not sufficient equilibrium sampling has been achieved in the simulation, such that different regions of configuration space have been accessed with the correct relative probabilities and distributed according to the Boltzmann-factor distribution [45]. Second is the accuracy of the force field model used in the simulation. By being able to now access longer timescales during the simulations, extensive validation against experimental data has allowed for substantial improvements in the parameterization of various force fields. One major and anticipated improvement is the development of polarizable force fields, which can model the redistribution of electrons around an atom in response to changes in its environment, thereby possibly providing more accurate models of hydrogen bond networks within biomolecules and with solvent [46, 47].

In Section 4.3, we use a particular MD variant called targeted molecular dynamics (TMD), in which a subset of atoms is guided towards a target structure during the course of the simulation through the use of steering forces. The rationale for using this approach, as well as its limitations, are discussed in that section.

### 1.2.2 Virtual docking

Cellular processes come about due to the biomolecular interactions that occur with other biomolecules or with a variety of small molecules. Virtual molecular docking tools aim to predict the how two molecules orient with respect to one another in order to form a stable complex, with predicted poses ranked based on some scoring function to approximate the binding affinity of the complex. Docking protocols can thus be thought of as a combination of a search algorithm and a scoring function [48]. In this research, we use protein-protein virtual docking for predicting if and how some of the multiple domains from ACTN4 interact with one another, and also protein-ligand docking to predict the binding mode of the hydrophilic head group of phosphatidylinositol moieties (e.g. PIP2 and PIP3) to the ABD.

Docking search algorithms can be classified based on how much flexibility is tolerated in the ligand and/or the protein side chains. Rigid-body search algorithms were first developed and still remain popular today, and are based on an exhaustive sampling of the rotational and translational degrees of freedom for the molecules being docked [49]. In protein-protein docking, fast Fourier transform (FFT)-based approaches are commonly used to enhance the speed of the search [50, 51]. Modeling of flexibility is commonplace today for exploring the conformational space of small molecule ligands; while modeling of flexibility in protein receptors, even for only the side chain motions, would be preferred, this can have a significant impact on the search time due to the increase in number of protein configurations that need to be considered. Perhaps one of the biggest limitations in docking searches is that role of solvent molecules and ions is often disregarded, due to the huge number of degrees of freedom associated with these that would make such searches computationally impractical if accounted for. It may be possible to address this by implicitly modeling the solvent effect within the scoring functions, although this will still not provide any useful insights on the interactions made by explicit waters during binding.

Development of scoring functions used for ranking predictions remains a key research area [52], with current approaches being either physics-based (using a classical physical treatment of interactions), knowledge-based (using observations based on statistical physics from a large number of complexes with known structures), or empirical (based on counting certain types of interactions) [53]. Just as in the force fields used in MD simulations, inaccuracies in scoring functions are still a major limiting factor, and as such the improvement of existing scoring functions or the developments of new ones continues to be a major topic of research.

## 1.3 OUTLINE

This dissertation is organized as follows:

In Chapter 2, we present an atomic model for the full homodimer of WT ACTN4. This model shows the assembly of the multiple domains of ACTN4, and provides a starting point for understanding how external cues can lead regulation of the actin-binding function via allosteric changes in the assembly.

In Chapter 3, we show that a tandem phosphorylation mechanism is involved in the EGF-induced dual phosphorylation of the ACTN4 N-terminal region, such that phosphorylation of the functional Y31 requires prior phosphorylation of Y4. We also show that this tandem mechanism, in conjunction with m-calpain cleavage, allows for differential regulation of actin binding between the front and rear of the cell during cell migration.

In Chapter 4, we use the atomic model of WT ACTN4 homodimer that was developed in Chapter 2 to gain insight on how the mechanisms by which the following external cues lead to regulation of actin binding: phosphorylation of Y265, binding of calcium ions, m-calpain cleavage at the neck region, and binding of phosphoinositide moieties.

In Chapter 5, we bring together the structural insights from the previous chapters to develop a preliminary network-level model that can serve as a computational tool for predicting the actin-binding response under multiple external cues.

In Chapter 6, we give the main conclusions of this research and discuss possible directions for future research.

In Appendix A, we describe the protocol used here for running GPU-accelerated MD simulations.

# CHAPTER 2

# STRUCTURAL MODELING OF THE FULL MULTI-DOMAIN ACTN4 HOMODIMER

## 2.1 BACKGROUND

The majority of human proteins are composed of multiple domains such that their assemblies and interactions [54], including those in response to various extracellular signals, are still an active research topic. This holds true for the isoforms belonging to the ACTN family of actin-binding proteins, and an additional complication is that each ACTN monomer is a large protein such that only parts of the full structure have been solved experimentally. Being able to understand how regulatory events such as tyrosine phosphorylation, which likely induce allosteric changes in the protein, are able to regulate protein function, however, can be made more tractable if one is first able to map the positions of the different domains relative to one another.

The most studied among the multiple domains present in the ACTNs (Fig 1.1) in terms of the number of solved structures is the actin-binding domain (ABD) that is found towards the N-terminal of the monomer sequence, since it is this domain that directly performs the function of binding to actin filaments. To date, there are three solved structures for the ABD of various

ACTNs: human ACTN1 (PDB IDs 2EYI and 2EYN [55]), human ACTN3 (PDB IDs 1TJT and 1WKU [56]), and human ACTN4 (PDB ID 2R0O [57]). The ABD is comprised of two adjacent calponin-homology (CH) domains, which is a recurrent domain in several actin-binding proteins [58, 59], and all the solved structures show both CH domains juxtaposed to each other with an extensive binding interface between them (Fig 2.1), in what is referred to as the "closed conformation" of the ABD. Results from several studies suggest a total of three putative actin-binding sites (ABSs) found in the ABD, with two found on the CH1 domain and the other on the CH2 domain [60-63]. However, a recent low-resolution cryo-EM structure of the complex between the ABD and an actin filament shows that binding occurs primarily through ABS2 on the CH1 domain only. ABS2 binds to a hydrophobic cleft between actin subdomains 1 and 3, which is believed to be binding "hot spot" for actin-binding proteins [64]. Interestingly, this

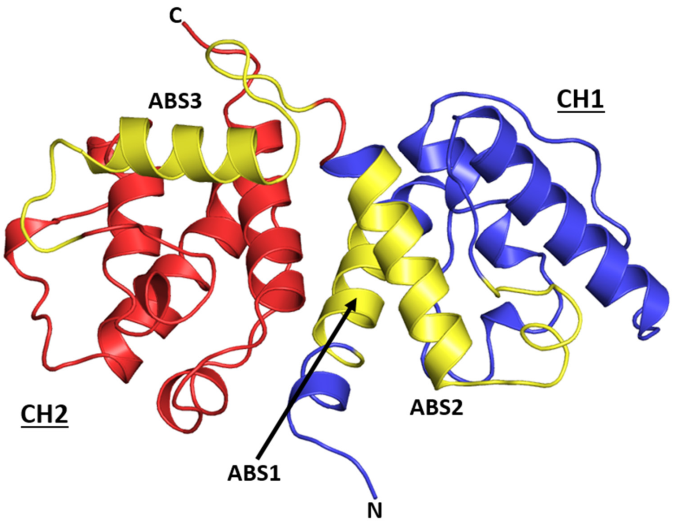


**Figure 2.1: Closed conformation of the ACTN ABD.**

The CH1 and CH2 domains are colored in blue and red, respectively. The three putative actin-binding sites (ABS1-3) are colored in yellow. The N-terminal of the CH1 domain (labeled "N") connects to the disordered N-terminal region, while the C-terminal of the CH2 domain (labeled "C") connects to the neck region. Coordinates are from PDB ID 2R0O [57].

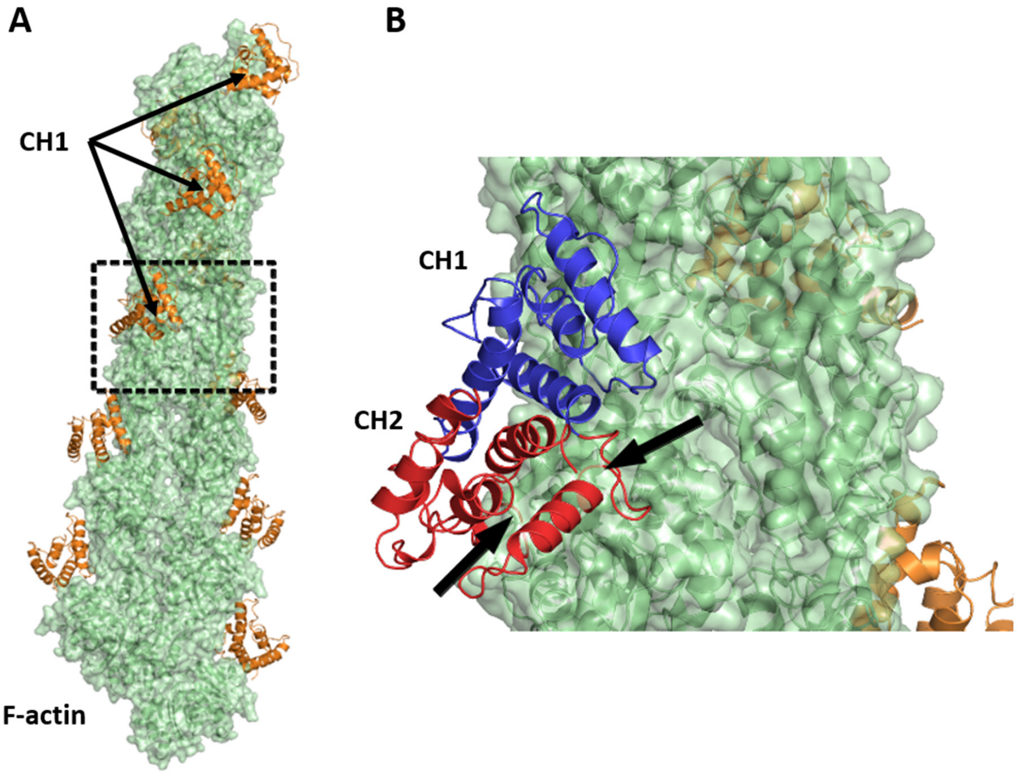


**Figure 2.2: Low-resolution cryo-EM structure of ACTN CH1 bound to F-actin.**

**(A)** An actin filament (green) was "decorated" with ACTN ABD prior to solving the cryo-EM structure, however reconstruction of only the CH1 domains (orange) was possible at the chosen resolution (16 Å) [18]. **(B)** Zoom-in of boxed region in **(A)**. In this view, the closed conformation of the ABD is structurally superimposed on one of the bound CH1 domains. Block arrows point to clashes between the CH2 domain and F-actin if the ABD binds in its closed conformation. Coordinates are from PDB ID 3LUE [18].

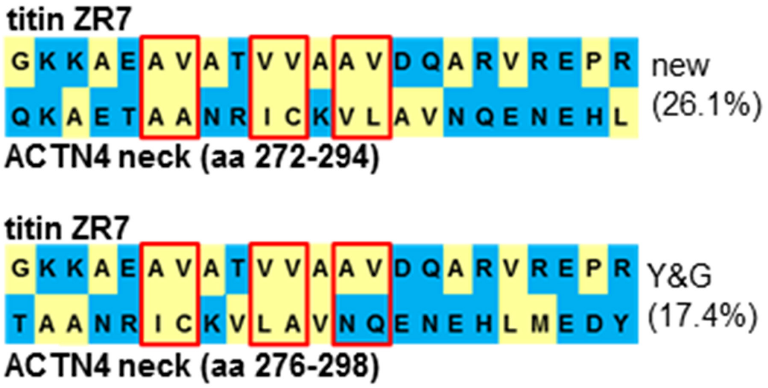
cryo-EM structure also shows that steric clashes will occur between the actin filament and CH2 domain if the ABD binds in a closed conformation (Fig. 2.2) [18]. This strongly suggests that prior opening of the ABD (via separation of the two CH domains) is required for actin binding to occur, and so a full structural model of the ACTN homodimer must allow for this ABD opening to take place.

Because of the antiparallel nature of the homodimer, an interesting question is whether or not there are important intermolecular interactions between the two monomers, in addition to those involving the rod domains that lead to dimer formation in the first place. Experimental evidence in the form of isothermal calorimetry measurements support the formation of a complex between the neck region and the second globular head (CaM2) of the CaM-like domain [65]. No structure has been experimentally resolved for this complex, however, and this complex was not found in a 20 Å cryo-EM structure of chicken smooth muscle ACTN (PDB ID 1SJJ [66]). In this chapter, we first build a homology model for the complex between the neck region and CaM2, then proceed to develop and validate an atomic model for the full human ACTN4 homodimer, which we found involves a novel ternary complex between the CH2 domain from the ABD, the neck region, and CaM2 [67]. This structural model should provide a starting point for understanding how allosteric changes in the multi-domain assembly, as a result of external cues, lead to regulation of ACTN4 function.

## 2.2 STRUCTURAL MODEL OF THE CAM2/NECK COMPLEX

For the muscle-specific ACTN isoforms, CaM2 has been shown to bind titin, with the interactions on the latter localized to the N-terminal Z-repeat motifs [68-71]. Titin is a giant protein that acts as a scaffold for binding of various proteins to assemble the sarcomere [72], in the case of muscle ACTN to be targeted specifically to the Z-lines. Upon binding to ACTN CaM2, the Z-repeat motifs of titin have been shown via circular dichroism and heteronuclear NMR spectroscopy to gain helical secondary structure [73, 74], and this was confirmed when a solution structure of CaM2/Z-repeat complex was solved (PDB ID 1H8B [75]). Interestingly, a multiple sequence alignment of human ACTN2 with several Z-repeat variants suggested a putative binding site within the ACTN neck region, which was confirmed through isothermal calorimetry to bind CaM2 with nanomolar affinity [65].

In the solution structure of the CaM2/Z-repeat complex, hydrophobic residues cover the side of the helical Z-repeat that is interacting with binding groove of CaM2 [75]. However, in the alignment of Young and Gautel [65] between the ACTN neck region and Z-repeat sequences, a polar NQ sequence would be found on the third helical turn and get buried in the hydrophobic interface of CaM2 (Fig. 2.3, top alignment). By shifting the alignment by one helical turn, all


titin ZR7
G K K A E A V A T V V A A V D Q A R V R E P R new
Q K A E T A A N R I C K V L A V N Q E N E H L (26.1%)
ACTN4 neck (aa 272-294)

titin ZR7
G K K A E A V A T V V A A V D Q A R V R E P R Y&G
T A A N R I C K V L A V N Q E N E H L M E D Y (17.4%)
ACTN4 neck (aa 276-298)

**Figure 2.3: Sequence alignments of titin Z-repeat 7 (ZR7) with ACTN4 neck region.**

In both alignments, polar and nonpolar residues are colored in blue and yellow, respectively. Top alignment shows our pairwise sequence alignment between titin repeat ZR7 and the human ACTN4 neck region (covering residues 272-294) that maximizes the hydrophobic contacts in contact with CaM2. Red boxes denote the matching of neck residues with those in ZR7 that are in contact with CaM2 based on the solution structure in PDB 1H8B [75]. Bottom alignment shows that suggested by Young and Gautel (covering residues 276-298), which has a polar sequence in the third helical turn of the neck [65]. The numbers in parentheses at the right of each alignment give the percent sequence similarity for the alignment as computed by the EMBOSS stretcher algorithm [76] using default parameters at the web interface provided EMBL-EBI [17]. Adapted from Figure 2B in [67].

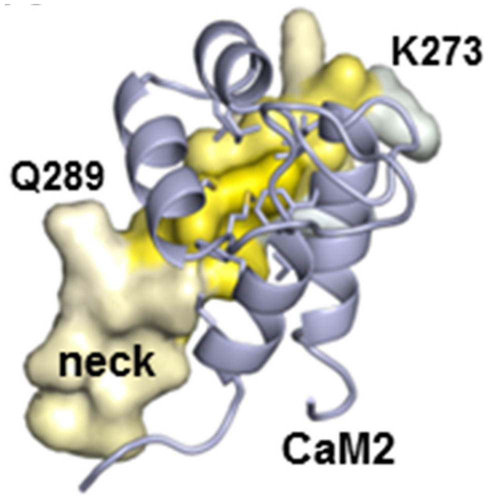


**Figure 2.4: Homology model of ACTN4 CaM2/neck complex.**

The helical neck is shown in surface representation here, with coloring based on side chain hydrophobicity using the Eisenberg scale [77]. Colors go from white to yellow, in the direction of increased side chain hydrophobicity. ACTN4 CaM2 is shown in cartoon representation (blue). Note that the most nonpolar region of the neck region (i.e., deepest yellow color) is directly interacting with the hydrophobic pocket of CaM2. Adapted from Figure 2A in [67].

three helical turns of the neck that face CaM2 are now populated with nonpolar residues (Fig. 2.3, bottom alignment), similar to what is seen in the solution structure, thereby maximizing the hydrophobic interactions at the binding interface. Using the CaM2/Z-repeat complex solution structure, we were then able to build a homology model for the ACTN4 CaM2/neck complex (Fig. 2.4), introducing *in silico* mutations as needed using PyMOL [78], that reflects the predominantly nonpolar binding interface for the neck as shown in our alignment.

For the CaM2/Z-repeat complex solution structure, the thirty models have an average change in solvent accessible surface area (ΔSASA) of 1326 ± 88 Å$^2$ upon forming the complex, which is very similar to our CaM2/neck homology model ΔSASA of 1318 Å$^2$ (Table 2.1). The relative binding free energy upon formation of the complex in our homology model is estimated

**Table 2.1: Comparison of ΔSASA and relative binding ΔG for complexes involving CaM2.**

| "Receptor" | "Receptor" SASA (Å$^2$) | "Ligand" | "Ligand" SASA (Å$^2$) | Complex SASA (Å$^2$) | ΔSASA (Å$^2$) | ΔG (kcal/mol)[a] |
|---|---|---|---|---|---|---|
| CaM2[b] | 5010 | Z-repeat | 2597 | 6281 | 1326 | -23.3 |
| CaM2 | 4598 | neck | 2498 | 5760 | 1318 | -19.7 |
| CaM2/neck | 5760 | ABD | 11578 | 15298 | 2040 | -24.7 |

[a] Binding free energies were estimated using FastContact [79].

b For the CaM2/Z-repeat complex, all values are averages over the 30 models in PDB ID 1H8B [75].

at -19.7 kcal/mol using the FastContact software [79], while that for formation of CaM2/Z-repeat complex is estimated at -23.3 ± 5.0 kcal/mol (error based on estimates from each NMR model). This indicates that our CaM2/neck model is predicted to form a tight complex similar to CaM2/Z-repeat, consistent with isothermal calorimetry results [65]. In addition to the hydrophobic interactions that primarily drive the binding, several polar contacts also contribute in stabilizing out CaM2/neck model (Fig. 2.5). A molecular dynamics (MD) simulation confirms the stability of the complex formed in our homology model (Fig. 2.6, blue curve). A similar MD simulation for a model based on the Young and Gautel alignment, on the other hand, shows a large increase in backbone RMSD in the last 20 ns of the simulation (Fig. 2.6, maroon curve); in fact, the helical neck in this simulation twisted by one helical turn such that it matches our homology model. (Fig. 2.6, green curve).

## 2.3 ACTN4 CH2, NECK, AND CAM2 FORM A TERNARY COMPLEX

Our next step was to use our homology model for the ACTN4 CaM2/neck complex as a starting point for building an atomic model of the full ACTN4 homodimer. In a low-resolution cryo-EM

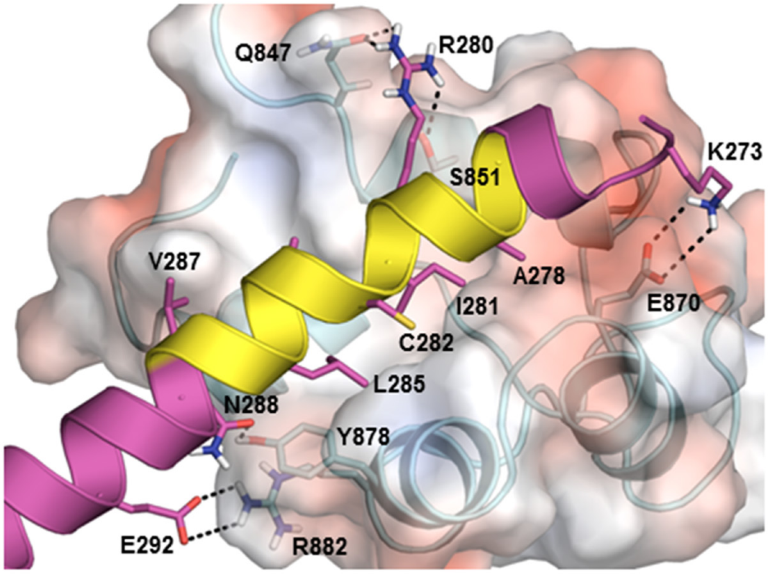


**Figure 2.5: The CaM2/neck complex is also stabilized by polar contacts.**

ACTN4 CaM2 is shown in surface representation, and colored according to electrostatic potential using APBS (red to white in the direction of increasing hydrophobicity). The helical ACTN4 neck is shown in cartoon, and colored in magenta except for the three helical turns in yellow that make nonpolar interactions with the CaM2 binding interface. Also shown are several polar contacts (dashed lines) that stabilize neck-CaM2. Adapted from Figure 2C in [67].

structure of chicken smooth muscle ACTN (PDB ID 1SJJ [66]), one end of the antiparallel dimer has a closed ABD interacting with CaM2 while the other end has an open ABD. The CaM2/neck complex was not resolved in this cryo-EM structure, however, even though this complex has been shown to form [65]. Still, due to the lack of symmetry at both ends of the ACTN dimer and shortcomings during freezing that could weaken the hydrophobic interactions [80] that are primarily responsible for formation of the CaM2/neck complex, interactions that are not seen in the cryo-EM structure (such as between CaM2/neck) can still be possible. Indeed, we found that the structure of the one end of the dimer (that has closed ABD interacting with CaM2) from the cryo-EM is fully consistent with our homology model of the CaM2/neck complex.

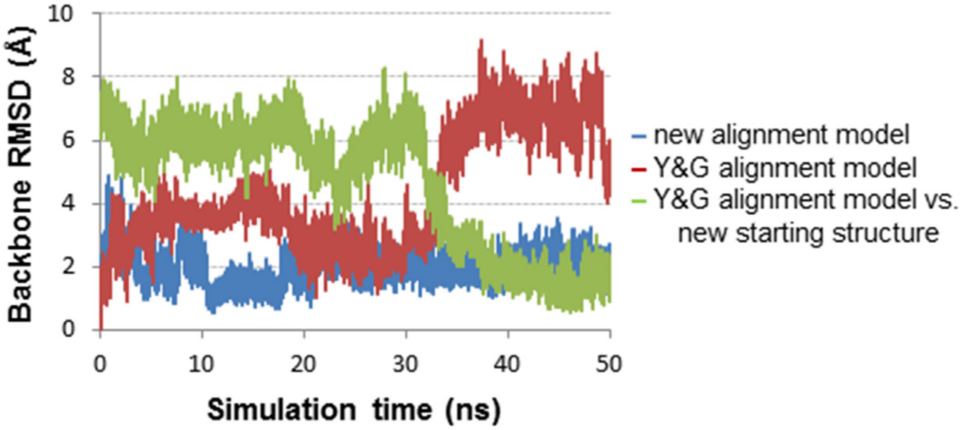


**Figure 2.6: Molecular dynamics confirm the stability of our CaM2/neck model.**

Unrestrained explicit solvent MD simulations were run for the homology models based on our alignment and that of Young and Gautel (Y&G) shown in Fig. 2.3. Backbone RMSD plots of the neck region were then computed from the trajectories using VMD [81]. The blue curve shows that the simulation of the model based on our alignment was stable throughout the trajectory; the model based on the Y&G alignment (magenta curve) shows a large increase in backbone RMSD at around 34 ns into the simulation, when the helical neck twisted by a helical turn and adopted a position similar to that seen in our homology model (green curve; backbone RMSD of Y&G alignment model relative to our model). Adapted from Figure 2D in [67].

A structural superimposition between the closed ABD end of the cryo-EM structure and our CaM2/neck model (Fig. 2.7), using CaM2 as the reference for the superimposition, shows that without the need for any further rearrangements, the helical neck can in fact fit between CaM2 and the closed ABD (CH2 in particular). Several contacts are formed between the neck region and CH2 domain, and compared to the mainly hydrophobic interface that stabilizes the CaM2/neck complex (Fig. 2.3 and Fig. 2.5), the CH2/neck interface is primarily characterized by contacts that are polar in nature (Fig. 2.8). We estimated a relative binding free energy of -15.1 kcal/mol for CH2/neck using FastContact [79]. In addition, several polar contacts appear to

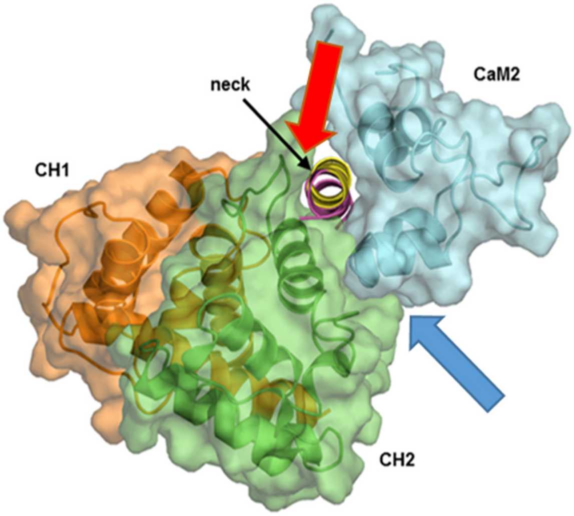


**Figure 2.7: The ACTN4 neck fits in between CH2 and CaM2.**

The CH2/CaM2 complex from the closed ABD end of cryo-EM structure PDB 1SJJ [66] is shown in combination of surface and cartoon representation. CaM2, CH1, and CH2 are colored blue, orange, and green, respectively. CH1 at the back makes no interactions with CaM2. The neck region (shown in cartoon only) is viewed through its helical axis, and can fit between CH2 and CaM2. The blue and red arrows point to interfaces with polar contacts between CH2/CaM2 (Fig. 2.8) and CH2/neck (Fig. 2.9), respectively. Adapted from Figure 3A in [67].

stabilize the interface between CaM2 and CH2 (Fig. 2.9) with an estimated relative binding energy of -9.6 kcal/mol. This leads to a total relative binding energy of -24.7 kcal/mol between the CaM2/neck complex and the CH2 domain (Table 2.1). All these interactions are fully consistent with a structural model involving a ternary complex between CH2, the neck region, and CaM2.

To this ternary complex, we can then add the rest of the ACTN4 domains using homology models from other experimental structures. For the rod domains, we used a crystal structure of the human ACTN4 rod domains homodimer (PDB ID 1HCI [12]), which contains

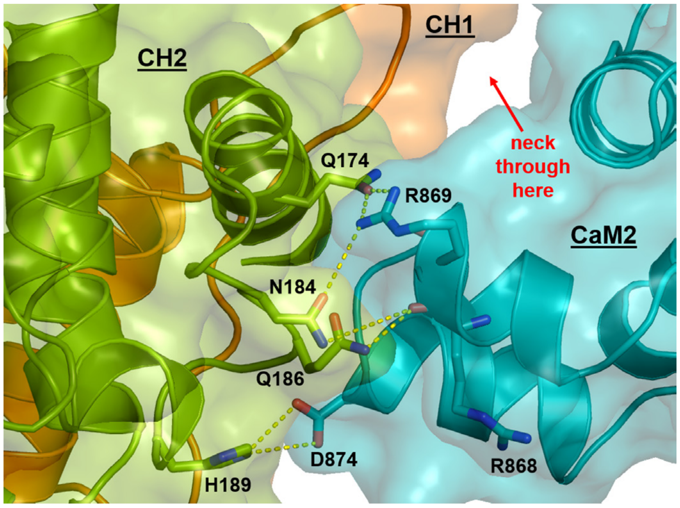


**Figure 2.8: ACTN4 CH2/neck binding interface.**

Interactions between CH2 (shown in green surface and cartoon) and the helical neck region (shown in cartoon only) are predominantly polar in nature. Residues that form polar interactions are shown in sticks. Several nonpolar residues that are part of the binding interface with CaM2 are also shown as extending from three helical turns (yellow cartoon) and away from CH2. Adapted from Figure 3B in [67].

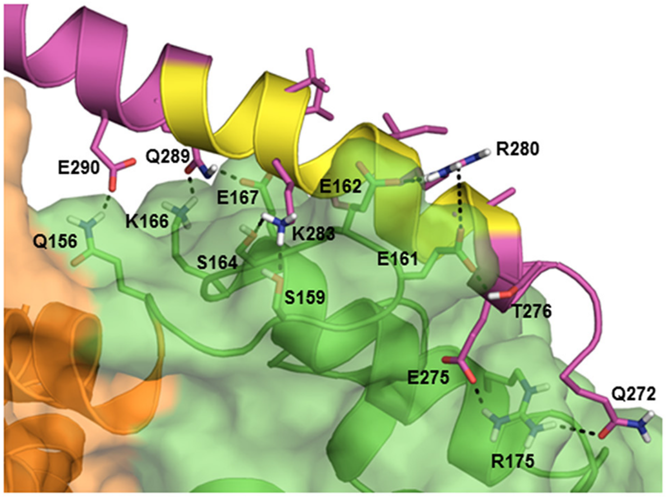


**Figure 2.9: ACTN4 CH2/CaM2 binding interface.**

Interactions between CH2 (shown in green surface and cartoon) and the CaM2 (shown in blue surface and cartoon) are predominantly polar in nature. Residues that form polar interactions are shown in sticks. Red arrow points to the pocket where the helical neck is located in the ternary complex with CH2 and CaM2. CH1 (colored orange) is shown behind the CH2/CaM2 complex.

two rod domain monomers (with four spectrin-like repeats each) bound in an antiparallel manner. The helical neck region from the ternary complex was connected to the first spectrin repeat (SR1) from one of these monomers. For the CaM-like domain, we built a homology model based on the solution structure of calcium-free calmodulin (PDB ID 1CFC [82]), which shares 49.4% sequence similarity. CaM2 from this model was structurally aligned to CaM2 from the ternary complex, and it is noteworthy that the resulting position of CaM1 allowed for it to be attached to the fourth spectrin repeat (SR4) from the other rod domain monomer. For the disordered N-terminal region, we chose a snapshot from an MD simulation that had the lowest SASA for the phosphorylatable Y31 (discussed more in Chapter 3). It is interesting that all of these domains could be added to the ternary complex such that the correct connectivity is preserved and there are no structural overlaps with other domains. This gives an atomic model for one of the end regions of the homodimer, which can be replicated at the other end of the rod dimer to provide a full model for the entire ACTN homodimer. This full model is shown at the center of Fig. 2.10, with all the experimental structures of the pairwise complexes that were used for its construction shown in the periphery (shown with the corresponding PDB IDs).

For this full model with the ABD in a closed conformation, the actin-binding site (ABS2) on CH1 (pointed to by black arrow in Fig. 2.10) is blocked by the nearby spectrin domains, thereby requiring the ABD to adopt an open conformation for binding to F-actin to take place. This is consistent with the cryo-EM structure of F-actin decorated with human ACTN3 ABD (PDB ID 3LUE [18]), where only CH1 was found at the binding interface. We also note that in our full model, CH1 does not make interactions that contribute to stabilizing the ternary complex. In fact, CH1 only makes interactions with CH2 in this model, which means that

].

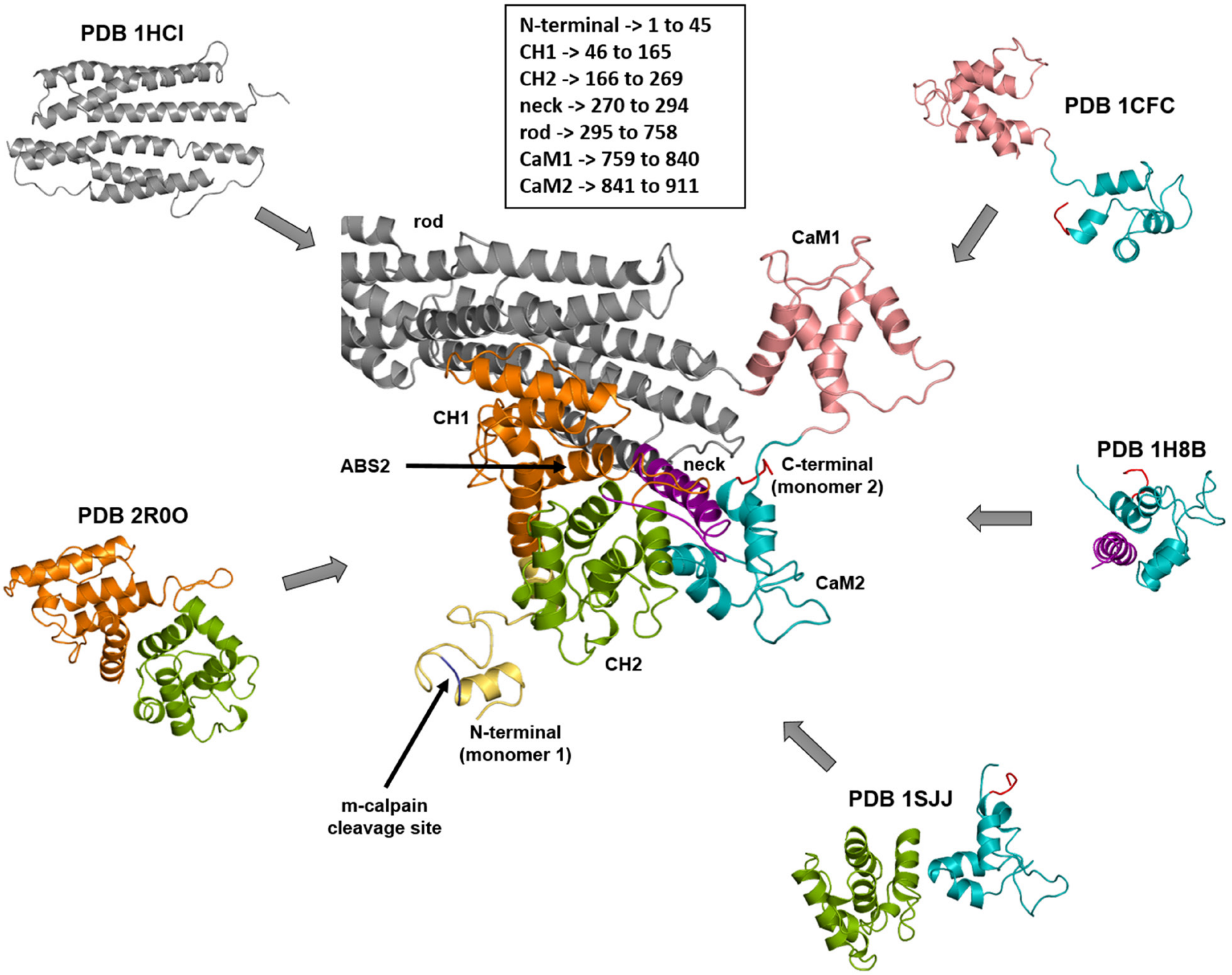


**Figure 2.10: Full atomic model of ACTN4 homodimer end region.**

The full model (shown in the center) was built from homology models to published structures (PDB IDs of known pairwise interactions shown surrounding the central model). For monomer 1, CH1 and CH2 of the ABD are colored orange and green, respectively, and the neck region is colored purple. Also shown is a snapshot from an MD simulation of the disordered N-terminal (colored yellow), to show that this segment can be added to the full model without overlapping other domains. The primary m-calpain cleavage site in ACTN4 is located in this disordered N-terminal region [36]. ABS2 is found on CH1, and in this model its access to F-actin is partially blocked by the rod domains (colored gray for both monomers) while the ABD takes a closed conformation. For monomer 2, CaM2 is colored blue, while CaM1 is colored pink and positioned on top of the CH2/neck/CaM2 ternary complex. Boxed area on top gives the residue numbering for each domain in the human ACTN4 monomer. Adapted from Figure 4 in [67].

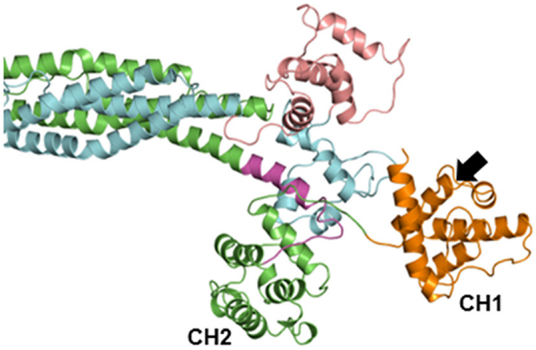


**Figure 2.11: Atomic model of ABD opening in the full ACTN4 homodimer model.**

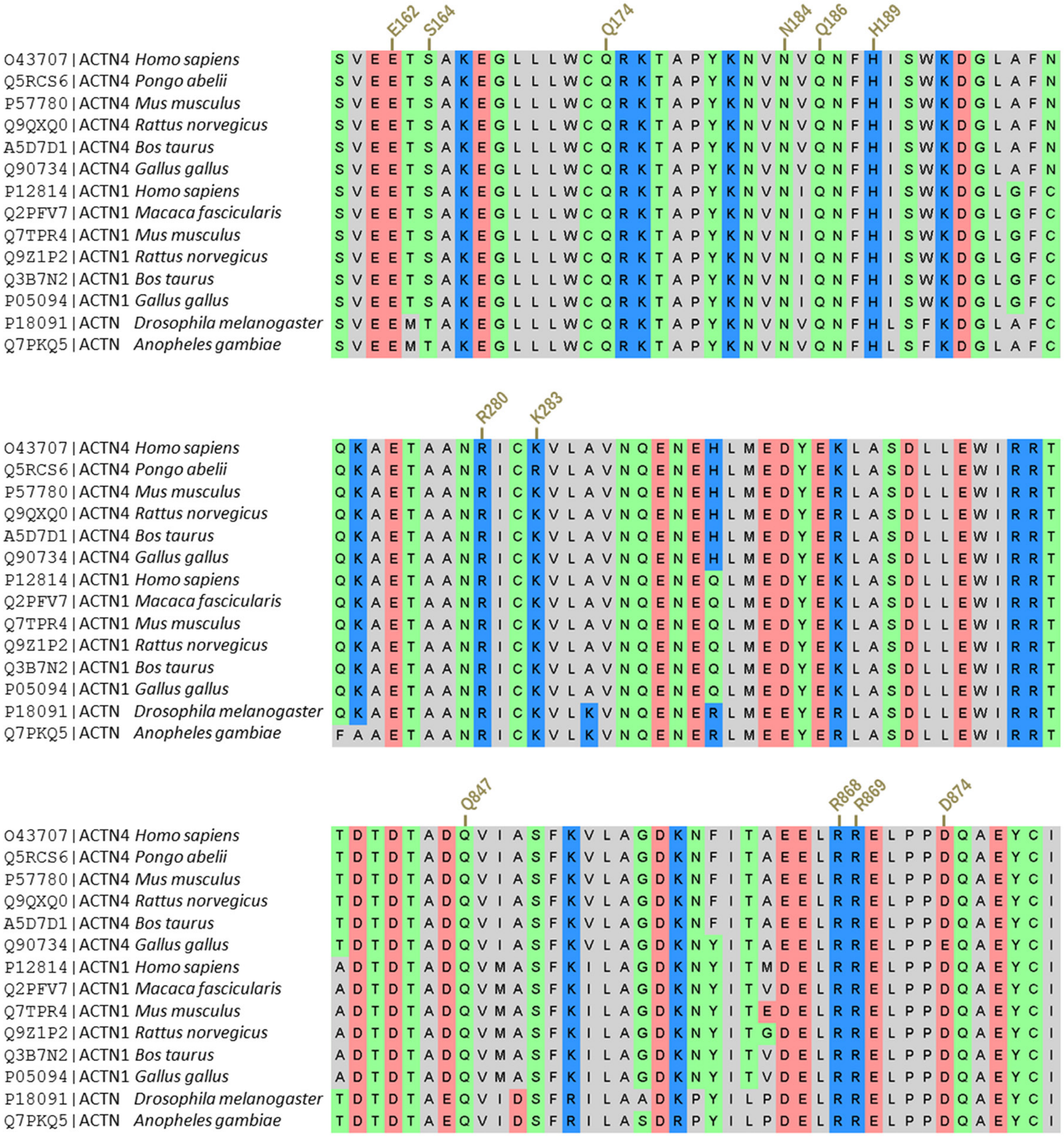
For clarity, the disordered N-terminal is not shown here. Black arrow points to ABS2 on CH1. Adapted from Figure 5 in [67].

breaking the CH1-CH2 interface to open the ABD is all that is needed for actin binding to happen. This also means that the ABD can switch to an open conformation without disrupting the structure of the ternary complex, thereby establishing the latter as the core motif that maintains the structural stability of the ACTN4 end region. Interestingly, by structurally superimposing CH2 from the one end of the chicken smooth muscle cryo-EM structure that has open ABD onto our full model, CH1 gets positioned at the tip of the dimer (Fig. 2.11). This predicted configuration appears optimal for binding F-actin, since ABS2 on CH1 is now fully exposed and accessible.

## 2.4 EXPERIMENTAL VALIDATION OF THE ACTN4 TERNARY COMPLEX

The ternary complex in our full ACTN4 model predicts that CH2 makes stable interactions with both the neck region and CaM2. We chose the following three interfaces to validate the interactions between CH2 and neck (Fig. 2.8) or CaM2 (Fig. 2.9): (a) E162 and S164 in the loop between CH1 and CH2 with R280 and K283 of the neck and Q847 of CaM2, (b) a cluster of polar residues (Q174, N184, and Q186) of CH2 with R868/R869 of CaM2, (b), and (c) H189 of CH2 with D874 of CaM2. All these residues are highly conserved among the non-muscle ACTN isoforms (Fig. 2.12) and so are likely to play a direct role in the stability of CH2 binding to neck and CaM2. Disrupting these contacts should tend to free the ABD from the ternary complex; this would then provide more avenues for ABD opening to occur and thus give more opportunities for ABS2 to bind to F-actin. Gel-based assays using mutant constructs for each of these set of residues indeed validate these predictions. The experiments were performed by Dr. Hanshuang Shao, under the supervision of Dr. Alan Wells, both from the Department of Pathology, University of Pittsburgh.

For the loss-of-function double mutant E162A/S164A, actin binding was found to increase to around 166% relative to that for WT (Fig. 2.13A). This is consistent with our prediction above, and thereby validating that CH2 makes interactions with the helical neck region in the ternary complex. We were also able to validate two sets of interactions between CH2 and CaM2. First, R869 from CaM2 makes interactions with several polar residues from CH2, but the R869A mutation shows no effect on actin binding (Fig. 2.13B, left). Interestingly, the double mutant R868A/R869A increases actin binding to about 187% relative to WT, but the R868A mutation alone has no effect. These results imply a redundancy in function for R868 and R869 in tethering together CH2 and CaM2. This redundancy is apparent from the structure since either arginine can reach to make interactions with the cluster of polar residues on CH2. This type of robustness is not uncommon in regulatory proteins; for example, single loss-of-function mutations of R469 or R470 in the NPR-C receptor have been shown to have no effect on PLC-β



```
O43707|ACTN4 Homo sapiens              SVEETSAKEGLLLWCQRKTAPYKNVNVQNFHISWKDGLAFN
Q5RCS6|ACTN4 Pongo abelii              SVEETSAKEGLLLWCQRKTAPYKNVNVQNFHISWKDGLAFN
P57780|ACTN4 Mus musculus              SVEETSAKEGLLLWCQRKTAPYKNVNVQNFHISWKDGLAFN
Q9QXQ0|ACTN4 Rattus norvegicus         SVEETSAKEGLLLWCQRKTAPYKNVNVQNFHISWKDGLAFN
A5D7D1|ACTN4 Bos taurus                SVEETSAKEGLLLWCQRKTAPYKNVNVQNFHISWKDGLAFN
Q90734|ACTN4 Gallus gallus             SVEETSAKEGLLLWCQRKTAPYKNVNVQNFHISWKDGLAFN
P12814|ACTN1 Homo sapiens              SVEETSAKEGLLLWCQRKTAPYKNVNIQNFHISWKDGLGFC
Q2PFV7|ACTN1 Macaca fascicularis       SVEETSAKEGLLLWCQRKTAPYKNVNIQNFHISWKDGLGFC
Q7TPR4|ACTN1 Mus musculus              SVEETSAKEGLLLWCQRKTAPYKNVNIQNFHISWKDGLGFC
Q9Z1P2|ACTN1 Rattus norvegicus         SVEETSAKEGLLLWCQRKTAPYKNVNIQNFHISWKDGLGFC
Q3B7N2|ACTN1 Bos taurus                SVEETSAKEGLLLWCQRKTAPYKNVNIQNFHISWKDGLGFC
P05094|ACTN1 Gallus gallus             SVEETSAKEGLLLWCQRKTAPYKNVNIQNFHISWKDGLGFC
P18091|ACTN  Drosophila melanogaster   SVEEMTAKEGLLLWCQRKTAPYKNVNVQNFHLSFKDGLAFC
Q7PKQ5|ACTN  Anopheles gambiae         SVEEMTAKEGLLLWCQRKTAPYKNVNVQNFHLSFKDGLAFC

O43707|ACTN4 Homo sapiens              QKAETAANRICKVLAVNQENEHLMEDYEKLASDLLEWIRRT
Q5RCS6|ACTN4 Pongo abelii              QKAETAANRICRVLAVNQENEHLMEDYEKLASDLLEWIRRT
P57780|ACTN4 Mus musculus              QKAETAANRICKVLAVNQENEHLMEDYERLASDLLEWIRRT
Q9QXQ0|ACTN4 Rattus norvegicus         QKAETAANRICKVLAVNQENEHLMEDYERLASDLLEWIRRT
A5D7D1|ACTN4 Bos taurus                QKAETAANRICKVLAVNQENEHLMEDYERLASDLLEWIRRT
Q90734|ACTN4 Gallus gallus             QKAETAANRICKVLAVNQENEHLMEDYEKLASDLLEWIRRT
P12814|ACTN1 Homo sapiens              QKAETAANRICKVLAVNQENEQLMEDYEKLASDLLEWIRRT
Q2PFV7|ACTN1 Macaca fascicularis       QKAETAANRICKVLAVNQENEQLMEDYEKLASDLLEWIRRT
Q7TPR4|ACTN1 Mus musculus              QKAETAANRICKVLAVNQENEQLMEDYEKLASDLLEWIRRT
Q9Z1P2|ACTN1 Rattus norvegicus         QKAETAANRICKVLAVNQENEQLMEDYEKLASDLLEWIRRT
Q3B7N2|ACTN1 Bos taurus                QKAETAANRICKVLAVNQENEQLMEDYEKLASDLLEWIRRT
P05094|ACTN1 Gallus gallus             QKAETAANRICKVLAVNQENEQLMEDYEKLASDLLEWIRRT
P18091|ACTN  Drosophila melanogaster   QKAETAANRICKVLKVNQENERLMEEYERLASDLLEWIRRT
Q7PKQ5|ACTN  Anopheles gambiae         FAAETAANRICKVLKVNQENERLMEEYERLASDLLEWIRRT

O43707|ACTN4 Homo sapiens              TDTDTADQVIASFKVLAGDKNFITAEELRRELPPDQAEYCI
Q5RCS6|ACTN4 Pongo abelii              TDTDTADQVIASFKVLAGDKNFITAEELRRELPPDQAEYCI
P57780|ACTN4 Mus musculus              TDTDTADQVIASFKVLAGDKNFITAEELRRELPPDQAEYCI
Q9QXQ0|ACTN4 Rattus norvegicus         TDTDTADQVIASFKVLAGDKNFITAEELRRELPPDQAEYCI
A5D7D1|ACTN4 Bos taurus                TDTDTADQVIASFKVLAGDKNFITAEELRRELPPDQAEYCI
Q90734|ACTN4 Gallus gallus             TDTDTADQVIASFKVLAGDKNYITAEELRRELPPEQAEYCI
P12814|ACTN1 Homo sapiens              ADTDTADQVMASFKILAGDKNYITMDELRRELPPDQAEYCI
Q2PFV7|ACTN1 Macaca fascicularis       ADTDTADQVMASFKILAGDKNYITVDELRRELPPDQAEYCI
Q7TPR4|ACTN1 Mus musculus              ADTDTADQVMASFKILAGDKNYITEDELRRELPPDQAEYCI
Q9Z1P2|ACTN1 Rattus norvegicus         ADTDTADQVMASFKILAGDKNYITGDELRRELPPDQAEYCI
Q3B7N2|ACTN1 Bos taurus                ADTDTADQVMASFKILAGDKNYITVDELRRELPPDQAEYCI
P05094|ACTN1 Gallus gallus             ADTDTADQVMASFKILAGDKNYITVDELRRELPPDQAEYCI
P18091|ACTN  Drosophila melanogaster   TDTDTAEQVIDSFRILAADKPYILPDELRRELPPDQAEYCI
Q7PKQ5|ACTN  Anopheles gambiae         TDTDTAEQVIDSFRILASDRPYILPDELRRELPPDQAEYCI
```

**Figure 2.12: Conserved residues stabilize the ACTN4 ternary complex.**

Multiple sequence alignment (MSA) of non-muscle ACTNs is shown for only the regions around the contact residues (marked above each alignment, numbering based on human ACTN4 sequence) that are predicted to stabilize the binding of CH2 to neck or CaM2. ACTN sequences are taken from UniProtKB/Swiss-Prot [15], with accession IDs shown on the left. Residues are colored as follows: blue (positively-charged), pink (negatively, charged), green (polar), and gray (non-polar). MSA was built using the Clustal Omega alignment software v1.2.1 [16], using default parameters at the web interface provided by EMBL-EBI [17].

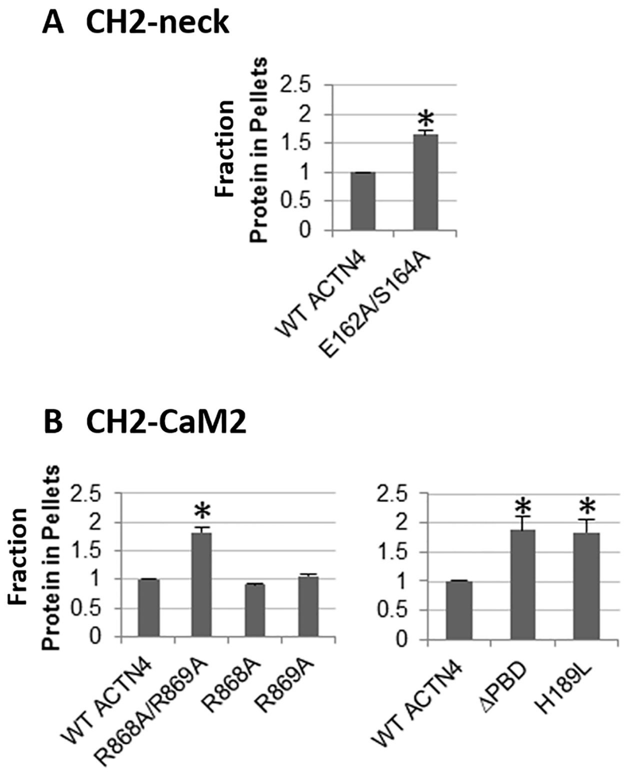


**Figure 2.13: Experimental validation of residues that stabilize the ACTN4 ternary complex.**

Quantitative results of Coomassie Blue G-250 stained polyacrylamide gels for actin-binding assays to validate the following interfaces: E162 and S164 in CH2/neck interface **(A)**, R868 and R869 in CH2/CaM2 interface **(B**, left**)**, and H189 in CH2/CaM2 interface **(B**, right**)**. Loss-of-function mutations for each of these sets of residues were compared to WT protein in terms of amount bound to F-actin. Image analysis and quantitation of three experiments are shown in each plot, with the percent of mutant ACTN4 that co-sedimented with F-actin normalized relative to WT. Student's t test was performed to compare each mutant with WT. Asterisks denote $p < 0.05$ between the labeled mutant and WT. Error bars indicate ± S.D. Adapted from Figures 6 to 8 in [67].

activity but the double mutation of both arginines does have a huge impact [83]. Second, actin-binding assays for the H189L mutation in CH2 also show increased binding of around 187% relative to WT (Fig. 2.13B, right). The H189 residue is in fact part of a trio of residues that are purported to act as the binding site for phosphoinositides [39], and mutations of this trio to hydrophobic residues (referred to as the ΔPBD mutation) also shows actin binding that is comparable to that of the single mutant H189L. Our data therefore suggests a possible mechanism by which phosphoinositides can disrupt the contacts made by H189 with CaM2 and so have an impact on actin binding. Note that our structural model shows that CaM2/neck still forms a stable complex even if mutations in R868/R869, E162/S164, and H189 act to disengage CH2 of the ABD from the ternary complex.

## 2.5 SUMMARY

In this chapter, we developed an atomic model for the end region of the WT ACTN4 homodimer. Starting from a homology model of the CaM2/neck complex (based on the solution structure of the CaM2/titin complex), we incorporated information from a low-resolution cryo-EM structure of chicken smooth muscle ACTN to find that a novel ternary complex can form between CH2/neck/CaM2. From there, homology models of the spectrin-like rod domains and CaM1 were added to the structure, giving a model of the ACTN4 end region. Since both ends of the antiparallel homodimer contain the same domain, this means that a structural model of the full molecule was effectively built. The length of the four repeats in the rod domains make it highly unlikely that allosteric changes in one end will be transmitted to the other end. All the components of the model are based on high-resolution structures except for the CH2/CaM2 interface that is based on the cryo-EM; this interface was experimentally validated, however, by using actin-binding assays to confirm predictions based on the model that mutations in the CH2/CaM2 interface will enhance the binding to actin filaments. Using this full model, the next step would be to see if it can be used to generate any insight into the molecular mechanisms by which various external cues are able to regulate the actin-binding function of ACTN4. But first, in the next chapter we will look at the disordered N-terminus of ACTN4, which is the only region not explicitly developed in our homodimer model due to its lack of structure, and discuss a novel mechanism by which tandem phosphorylation of two tyrosines in this region leads to regulation of actin binding in the presence of EGF.

# CHAPTER 3

# ACTN4 REGULATION VIA A TANDEM PHOSPHORYLATION MECHANISM IN ITS DISORDERED N-TERMINAL REGION

## 3.1 BACKGROUND

The only region for which there is no experimental structure available or a homology model is not straightforward to build is the 45-residue disordered N-terminal region before CH1 of the ABD. This region is believed to be unstructured based on the fact that homologous sequences that are adjacent to the ABDs of the actin-binding proteins plectin and fimbrin show no electron density when solving their structures using X-ray crystallography [84, 85]. Interestingly, tyrosine phosphorylation of the N-terminal region is one mechanism that regulates the binding of the non-muscle ACTNs to actin filaments: dual phosphorylation of ACTN4 at Y4 and Y31 in response to epidermal growth factor (EGF) [21] and phosphorylation of ACTN1 at Y12 by focal adhesion kinase (FAK) [22] both lead to a decrease in actin binding. However, how the phosphorylation of two tyrosine residues in the unstructured N-terminus of ACTN4 can affect its actin-binding function is not yet understood.

Protein phosphorylation has been found to occur with greater frequency in the intrinsically disordered regions (IDRs) of proteins [86], often occurring at multiple neighboring sites [87]. Since disorder content is higher in signaling proteins [88, 89], an understanding of how multiple phosphorylation events in IDRs leads to the regulation of protein function is critical for generating a more complete picture of cellular signaling. Structural changes have been shown to take place in these regions upon phosphorylation; these changes can involve explicit disorder-to-order transitions [90-92] or the opposite [93-95], or less subtle changes such as an increase in the radius of gyration [96, 97]. In this chapter, we first demonstrate that only phosphorylation of Y31 is necessary and sufficient to reduce actin binding of ACTN4. We then show that the ACTN4 N-terminal region adopts a molten globule-like state such that Y4 and Y31 are initially exposed and buried, respectively, in the unphosphorylated protein, and phosphorylation of Y4 causes structural changes that lead to the exposure and eventual phosphorylation of Y31. This tandem phosphorylation mechanism within an IDR of ACTN4, when coupled with m-calpain cleavage that occurs at the rear of motile cells, provides a mechanism for locale-specific regulation of actin remodeling during cell motility.

## 3.2 TANDEM PHOSPHORYLATION OF Y4/Y31 IN THE ACTN4 N-TERMINAL

To gain insight into the role of dual phosphorylation of the ACTN4 N-terminal region, we constructed a multiple sequence alignment (MSA) of ACTN isoforms from various species and found that not only are Y12 of ACTN1 and Y31 of ACTN4 homologous, but this particular tyrosine is conserved in all four mammalian isoforms (Fig. 3.1). This suggests that a similar

**Figure 3.1: Phylogenetic tree of ACTN sequences from various species.**

Full ACTN isoform protein sequences from various species were obtained from the NCBI protein database. Multiple sequence alignment (MSA) and phylogenetic tree construction for these protein sequences were done using the interface provided by the Phylogeny.fr web service [98]. The MUSCLE program [99] was used for building the MSA, using default parameters provided by the web service. Full protein sequences were used in building the alignment; shown here is only the portion of the alignment covering the disordered N-termini. The column colored in red contains the strictly conserved Y31 from human ACTN4 (HOMO_4), which is homologous to Y12 in ACTN1 (HOMO_1). Columns containing Y4 of ACTN4 and human ACTN1 are colored in blue and green, respectively. The maximum likelihood phylogenetic tree was the constructed from the MSA using the PhyML program [100], again using default parameters provided by the web service. Statistical support for the tree branches were computed via bootstrapping with 100 replicates. TreeDyn [101] was used for rendering the phylogenetic tree. Scale bar represents the number of substitutions per site.

regulatory mechanism may exist for the muscle isoforms for inhibition of actin binding. We were in fact able to trace the conservation of human ACTN4 Y31 back to the emergence of Cnidarians (*Hydra spp.*), suggesting that the function of this particular tyrosine was established to be critical early in animal evolution. Interestingly, the MSA shows that only ACTN4 has a second phosphorylation site at Y4; ACTN1 Y4 is not homologous to ACTN4 Y4 and has so far not been found to be phosphorylated. Note that the amino acid sequence between Y4 and Y31 in ACTN4 shows the insertion of a long hydrophobic linker, while ACTN1 has a hydrophilic stretch of residues next to Y12. This suggests that the ACTN4 N-terminus might adopt a more collapsed state than the ACTN1 N-terminus based on charge distribution and length arguments [102, 103].

The broad conservation of ACTN4 Y31 in ACTN isoforms from other animal species indicates that phosphorylation of this tyrosine plays a major role in the function of actin-binding inhibition. This was confirmed by performing actin-binding assays using single phosphomimic mutations to glutamic acid at either Y4 or Y31 of ACTN4. The Y31E mutant shows the same decrease in binding as the double phosphomimic Y4/31E mutant, while the Y4E mutant shows almost no change in binding when compared with WT protein (Fig. 3.2). This shows that the reduced actin binding that was seen previously for the double phosphomimic is entirely due to phosphorylation of Y31, and so the conserved Y31/Y12 is necessary and sufficient to modulate the actin-binding function of ACTN4/ACTN1 [21, 22]. But this still leaves open the question as to the function of the more recently evolved Y4 and its EGF-induced phosphorylation in ACTN4.

To explore the possible role of phosphorylated Y4, we performed explicit-solvent molecular dynamics (MD) simulations of WT and Y4E phosphomimic constructs of the 45-residue long disordered ACTN4 N-terminal region. The simulation for the WT construct showed that the Y31 side chain remains mostly buried in a molten globule packed by two relatively stable anti-parallel helices that each start around Y4 and Y31, with the Y4 side chain

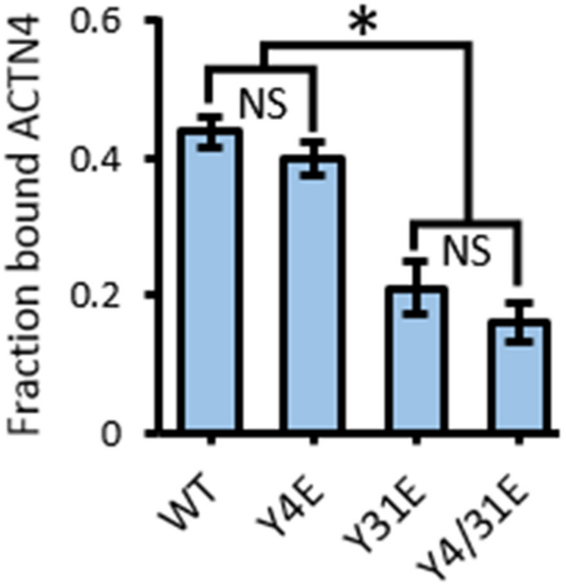


**Figure 3.2: Y31 phosphorylation is responsible for the inhibition actin binding of ACTN4.**

Actin-binding assays were performed for WT, Y4E, Y31E, and Y4/31E constructs of ACTN4, with ultracentrifugation separating unbound ACTN4 (supernatant or S fraction) from actin-bound ACTN4 (pellet or P fraction). The plot shows the quantitation of three experiments for amounts of ACTN4 in the P fraction. The phosphomimic at Y4 has no significant effect on actin binding relative to WT (unphosphorylated) protein, while the phosphomimic at Y31 shows decreased actin binding that is similar to the double phosphomimic mutant. Error bars indicate s.e.m. [* $P < 0.05$ based on student's t-test. NS, not significant.] The experiments were performed by Dr. Hanshuang Shao, under the supervision of Dr. Alan Wells, both from the Department of Pathology, University of Pittsburgh.

significantly solvent exposed (Fig. 3.3 top plot, and Fig. 3.4A). This means that Y4 (but not Y31) is normally accessible for phosphorylation in the unphosphorylated protein. Strikingly, simulations of the Y4E phosphomimic construct show that the additional negative charge alters the packing of the helices into a parallel arrangement, such that the Y31 side chain becomes solvent exposed and accessible to kinase in the process (Fig. 3.3 middle plot, and Fig. 3.4B). Note that this is very different from ACTN1 where only Y12 is phosphorylated, and not

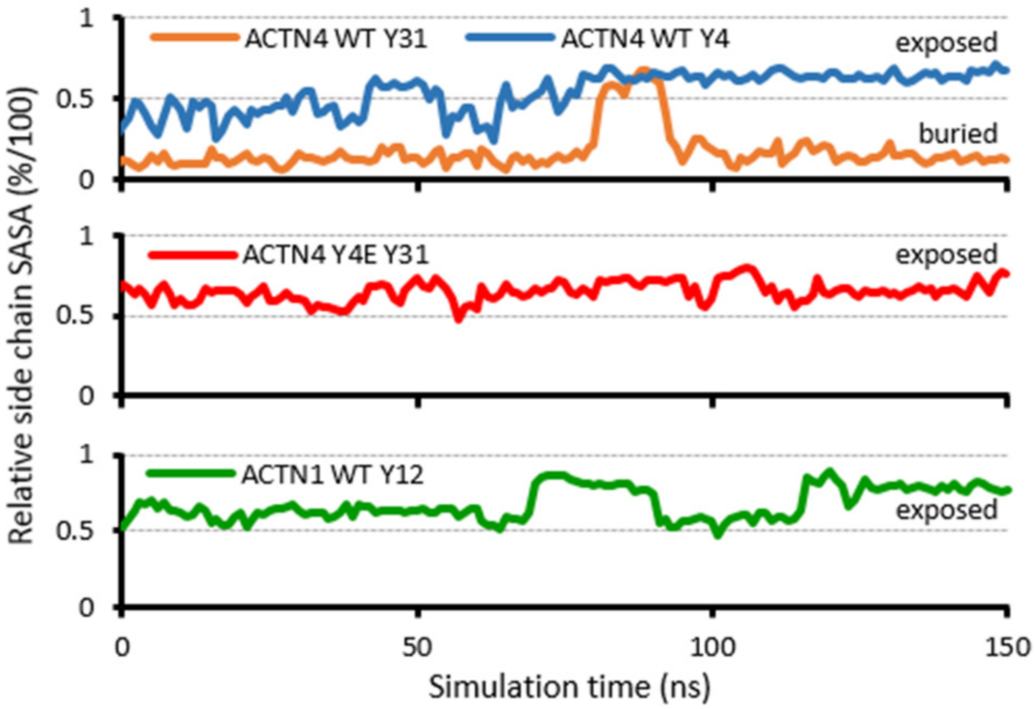


**Figure 3.3: Kinase accessibility of Y31 becomes possible upon phosphorylation of Y4.**

Shown are plots of side chain solvent-accessible surface area (SASA) over 150-ns MDs for Y4 and Y31 from WT ACTN4 (top), Y31 from Y4E ACTN4 (middle), and Y12 from WT ACTN1 (bottom). For WT (unphosphorylated) protein, Y4 is kinase accessible while the Y31 side chain is buried within the rest of the N-terminal region. Upon phosphorylation of Y4 (modeled here using the Y4E phosphomimic), changes in the structural dynamics of the N-terminal region cause Y31 to become solvent exposed and kinase accessible. For WT ACTN1 which only has one phosphorylation site at Y12, the side chain of the latter is solvent exposed as there is no other phosphorylation site to control its accessibility.

surprisingly, MD simulations of the hydrophilic ACTN1 N-terminal region confirm that the Y12 side chain is mostly solvent exposed (Fig. 3.3 bottom plot). To check if Y4 phosphorylation is really needed for Y31 to become accessible for phosphorylation, we performed assays to compare the phosphorylation levels of Y31 between Y4E (phosphomimic) and Y4A (loss-of-function) mutants. These assays showed that the loss of Y4 (via the Y4A mutant) does indeed

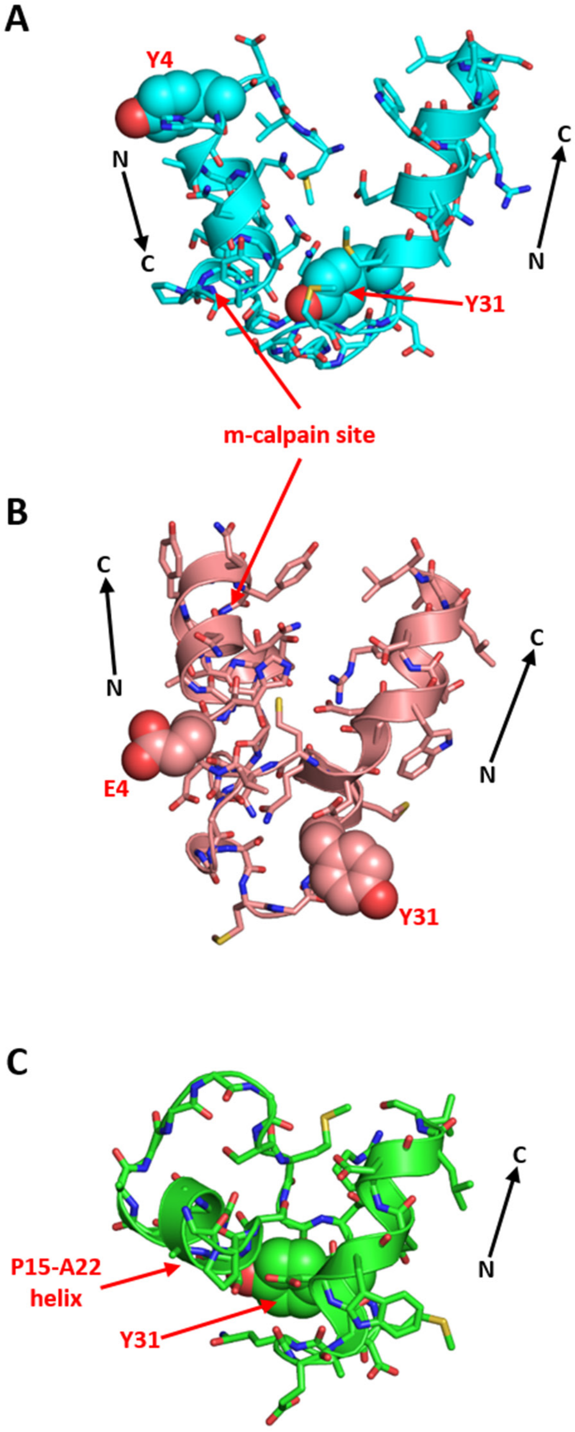


**Figure 3.4: Changes in the helical packing of the ACTN4 N-terminal region.**

For (A) full WT, (B) full Y4E, and (C) cleaved WT, the C-terminus of the peptide is shown at the upper left. (A) Y4 (exposed) and Y31 (buried) are shown in spheres. The helix that starts at Y4 packs in an antiparallel arrangement with the helix that starts at Y31, with the Y31 side chain buried between the two helices. The primary m-calpain cleavage site (Y13-G14 peptide bond) is exposed and accessible to the protease. (B) E4 and Y31 (both exposed) are shown in spheres. The extra charge flips the helix that starts at the E4 phosphomimic in a parallel arrangement with the helix that starts at Y31, which now makes Y31 exposed and accessible to kinase. The m-calpain cleavage site is still accessible to the protease. (C) Y31 (buried) is shown in spheres. Removal of residues 1 to 13 by m-calpain cleavage leads to a helical folding from P15 to A22. This helix helps to keep Y31 buried and inaccessible to kinase.

cause Y31 to be phosphorylated at a much lesser extent, when compared to the phosphomimic mutant that should make Y31 phosphorylatable (Fig. 3.5). These results strongly suggest an elegant mechanism where an alternate phosphorylation switch can be evolved by tuning the collapse propensity of an IDR. This tandem phosphorylation mechanism is sketched in Fig. 3.6:

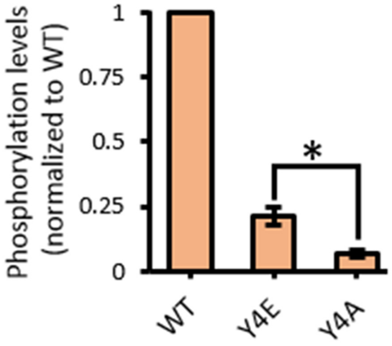


**Figure 3.5: Loss of Y4 inhibits phosphorylation of Y31.**

Phosphorylation detection assays were performed for WT, Y4E, and Y4A constructs of ACTN4. The plot shows the quantitation of three experiments for amount of EGF-induced phosphorylated tyrosine normalized to WT levels. For WT, phosphorylation of Y4 and Y31 is being detected. Both Y4E and Y4A measure Y31 phosphorylation levels only; Y4A shows decreased Y31 phosphorylation compared to Y4E. Error bars indicate s.e.m. [* $P < 0.05$ based on student's t-test.] The experiments were performed by Dr. Hanshuang Shao, under the supervision of Dr. Alan Wells, both from the Department of Pathology, University of Pittsburgh.

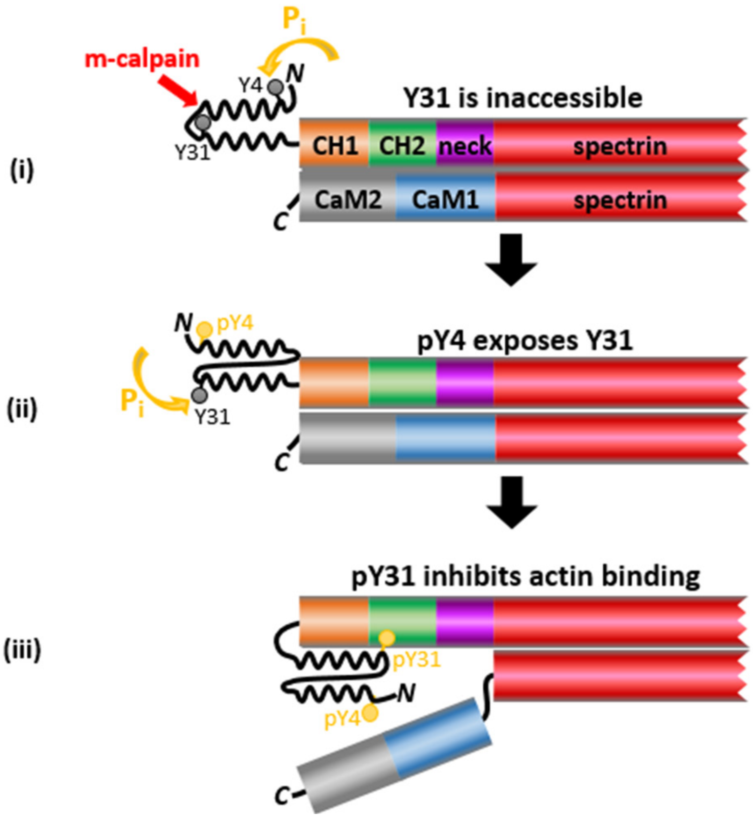


**Figure 3.6: Tandem phosphorylation model for Y4 and Y31 of the ACTN4 N-terminal region.** Black wavy lines denote N-terminal helices. (i) Y4 and Y31 are exposed and buried, respectively, as two helical elements in the N-terminal adopt an antiparallel conformation. (ii) Upon phosphorylation of the exposed Y4, the N-terminal undergoes a transition in its molten globule state such that the two helical elements adopt a parallel conformation. This transition makes Y31 accessible to kinase. (iii) Phosphorylated Y31 ca now act to inhibit actin binding, by latching the ABD in a closed configuration (see Fig. 3.7) The primary m-calpain cleavage site (red arrow) between Y13-G14 [36] is accessible to the protease and removes the first helical element (containing Y4).

(i) Y4 and Y31 are mostly accessible and inaccessible, respectively, for phosphorylation in the unphosphorylated protein; (ii) phosphorylation of Y4 leads to structural changes in the disordered N-terminus that exposes Y31 to its kinase; and, (iii) phosphorylation of Y31 mediates regulation of actin binding by ACTN4. In other words, the phosphorylation state of Y4 acts as a switch that determines whether or not the functionally-relevant Y31 will get phosphorylated.

An important question is how the phosphorylation of Y31 leads to reduced actin binding of ACTN4. Docking experiments of the helical element containing Y31E (phosphomimic of Y31) to the ABD strongly suggests that the helix can bind to a crevice on the surface of CH2 (Fig. 3.7A). This interaction will disrupt the binding of ABD and CaM2 in the ternary complex (Fig. 3.7B), which taken alone would lead to a predicted increase in actin binding similar to that seen for the three sets of artificial mutations R868A/R869A, E162A/S164A, and H189L (see Chapter 2.4) from earlier. More importantly, however, this interaction would also lead to tighter latching between CH1 (to which the N-terminal is covalently attached) and CH2, thereby further stabilizing the closed ABD state and leading to the observed decrease in actin binding.

## 3.3 SYSTEMS MODELING OF THE TANDEM PHOSPHORYLATION MECHANISM

We next explored the constraints imposed by adding an extra phosphorylation switch to control the functional Y31 regulation of actin binding. To do this, we used a systems modeling approach to study phosphorylation levels of Y31 in models where two phosphorylation sites are linked in the same molecule. This approach is based on an earlier approach to quantify crosstalk and competition between phosphorylation sites on different molecules [104]. This work in turn builds off that of Goldbeter and Koshland [105], who first showed that a ultrasensitive response can occur for a single phosphorylation site on a substrate at saturating conditions (relative to the

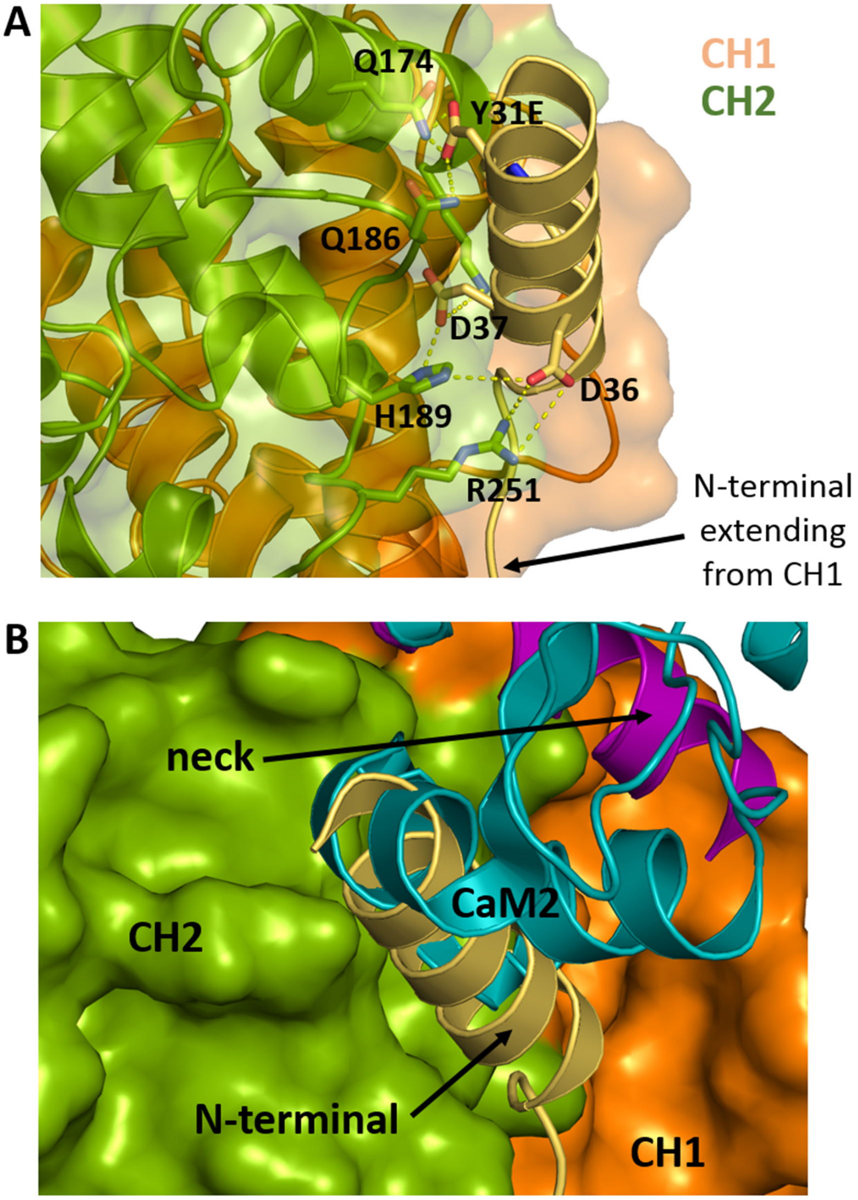


**Figure 3.7: Phosphorylation of Y31 latches the ABD in a closed configuration.**

The helical element of the N-terminal region comprising residues 31-41 and with the Y31E mutation was used as ligand for virtual docking to the crystal structure of a human ACTN4 ABD mutant (PDB ID 2R0O [57]), with K255E backmutated to Lys. Virtual docking was done using ClusPro [51, 106], with top-ranked docked models filtered based on balanced contributions between electrostatic and hydrophobic contacts. (**A**) Top-ranked docked model of the helical motif around Y31E and ACTN4 ABD. Several residues that make contacts, including the Y31E phosphomimic, are labeled. Q174, Q186, and H189 were shown earlier to be involved in interactions with CaM2

(Fig. 2.8). Black arrow points to where the N-terminal connects to CH1, indicating that this is a feasible docked model. (**B**) Same docked model superimposed with the core ternary complex (CH2-neck-CaM2). Note the clash between the N-terminal helix and CaM2. The helical element containing phosphorylated Y31 can thus detach the ABD from the ternary complex, while latching both CH domains together in the closed configuration.

concentrations of kinase and phosphatase in the system) such that the phosphorylation level of the site displays a switch-like behavior (either 0% or 100%) with respect to a parameter $r$ (defined as the ratio of the maximum kinase velocity over the maximum phosphatase velocity). Here, we use four different models: (A) in the presence of a single functional phosphorylation site; (B) a tandem model with both sites phosphorylated by a single kinase; (C) the same model but where the kinase has a weaker affinity towards the switch site; and, (D) a tandem model where the two sites are phosphorylated by different kinases. A description of these models is first given in the next few subsections, followed by the results obtained from the simulations of these models.

### 3.3.1 General notes on modeling approach

The single phosphorylation site, two-site 1-kinase/1-phosphatase (1K1P) tandem phosphorylation, and two-site 2-kinase/1-phosphatase (2K1P) tandem phosphorylation models were simulated using a rule-based approach with the BioNetGen software v2.2.2 [107, 108]. To simplify the models, we assumed for all the models that (1) a single phosphatase acts on the phosphorylation sites and (2) the catalytic rates of the kinase/s and phosphatase are the same. An additional assumption for the tandem phosphorylation models is that only one site out of the two can be bound at any instant, either by kinase or phosphatase. This is to account for steric hindrance that would prevent the simultaneous binding of different kinase/phosphatase molecules to two nearby sites. Although these assumptions would have allowed the models here to be solved without using a rule-based approach, we opted on using the latter in anticipation of future research where the restriction of simultaneous binding of enzymes to both sites is removed.

For the single-site and 1K1P models, the kinase is represented by the molecule type *kin(b)*, indicating that there is only one binding site on the kinase. For the 2K1P models, the two kinases are represented as *kin1(b)* and *kin2(b)*. For all the models, the single phosphatase is represented as *pho(b)*.

For all models, the binding, unbinding, and catalytic rates of kinase/s and phosphatase at either of the sites are set at 1.0 $\mu M^{-1} \cdot s^{-1}$, 0.001 $s^{-1}$, and 0.999 $s^{-1}$. This leads to a Michaelis constant of 1 μM for the hypothetical kinase/s and phosphatase in the simulations. The only exception is in model (C), where the binding rate of the single kinase to the Y4 site is set lower at 0.02 $\mu M^{-1} \cdot s^{-1}$ (while keeping the binding rate to the Y31 site at 1.0 $\mu M^{-1} \cdot s^{-1}$). This leads to a Michaelis constant in this model for the Y4 site of 50 μM.

### 3.3.2 Single phosphorylation site model

The N-terminal region was described in this model using the following molecule type:

*nter(b,s~U~P)*

where *s* refers to the single phosphorylation site at Y31 while *b* refers to the corresponding binding site. Site *s* can be either unphosphorylated (*U*) or phosphorylated (*P*) at any instant. The following reaction rules were then used for describing the single phosphorylation site model:

3.3.2A: *nter(b,s~U) + kin(b) <–> nter(b!1,s~U).kin(b!1)* *kin_f, kin_r*
3.3.2B: *nter(b!1,s~U).kin(b!1) –> nter(b,s~P) + kin(b)* *kin_cat*
3.3.2C: *nter(b,s~P) + pho(b) <–> nter(b!1,s~P).pho(b!1)* *pho_f, pho_r*
3.3.2D: *nter(b!1,s~P).pho(b!1) –> nter(b,s~U) + pho(b)* *pho_cat*

Rules 3.3.2A and 3.3.2B describe kinase binding to *b* and phosphorylation of *s*, respectively. Rules 3.3.2C and 3.3.2D give the corresponding rules for phosphatase binding to *b* and desphosphorylation of *s*, respectively.

In simulating this model, the phosphatase concentration was kept fixed at 1 nM, while the kinase and N-terminal region concentrations were varied between 0–2 nM and 0–100 μM, respectively. All simulations were run until the default convergence criterion for steady-state levels set by BioNetGen was reached. In plotting the results, the x-axis was made dimensionless by normalizing the N-terminal region concentration by the Michaelis constant of the kinase (1 μM), and similarly for the y-axis by normalizing the kinase activity (i.e. the maximum velocity of the kinase) by the phosphatase activity. Note that the enzyme activities are taken to be proportional to their concentration.

### 3.3.3 Two-site 1-kinase/1-phosphatase (1K1P) tandem model

The N-terminal region was described in this model using the following molecule type:

*nter(b1,b2,s1~U~P,s2~U~P)*

where *s1* and *s2* refer to the phosphorylation sites Y4 and Y31, respectively, while *b1* and *b2* refer to the corresponding binding sites. Both *s1* and *s2* can be either unphosphorylated (*U*) or phosphorylated (*P*) at any instant. The following reaction rules were then used for describing the 1K1P model with tandem phosphorylation:

3.3.3A: *nter(b1,b2,s1~U) + kin(b) <–> nter(b1!1,b2,s1~U).kin(b!1)* *kin_f, kin_r*

3.3.3B: *nter(b1!1,b2,s1~U).kin(b!1) –> nter(b1,b2,s1~P) + kin(b)* *kin_cat*

3.3.3C: *nter(b1,b2,s1~P,s2~U) + kin(b) <–> nter(b1,b2!1,s1~P,s2~U).kin(b!1)*
*kin_f, kin_r*

3.3.3D: *nter(b1,b2!1, s1~P,s2~U).kin(b!1) –> nter(b1,b2,s1~P,s2~P) + kin(b)*
*kin_cat*

3.3.3E: *nter(b1,b2,s1~P) + pho(b) <–> nter(b1!1,b2,s1~P).pho(b!1)* *pho_f, pho_r*

3.3.3F: *nter(b1!1,b2,s1~P).pho(b!1) –> nter(b1,b2,s1~U) + pho(b)* *pho_cat*

3.3.3G: *nter(b1,b2,s2~P) + pho(b) <–> nter(b1,b2!1,s2~P).pho(b!1)* *pho_f, pho_r*

3.3.3H: *nter(b1,b2!1,s2~P).pho(b!1) –> nter(b1,b2,s2~P) + pho(b)* *pho_cat*

Rules 3.3.3A and 3.3.3B describe kinase binding to *b1* and phosphorylation of *s1*, while rules 3.3.3E and 3.3.3F describe phosphatase binding to *b1* and dephosphorylation of *s1*. Rules 3.3.3C and 3.3.3D, and 3.3.3G and 3.3.3H, give the corresponding rules for sites *b2* and *s2*. In this tandem model, kinase binding and phosphorylation of the first site does not depend on the phosphorylation state of the second site (Rules 3.3.3A and 3.3.3B). However, kinase binding and phosphorylation of the second site does require prior phosphorylation of the first site (Rules 3.3.3C and 3.3.3D). Note that phosphatase binding and dephosphorylation at either site does not depend on the phosphorylation state of the other site (Rules 3.3.3E to 3.3.3H).

In simulating this model, the kinase, phosphatase, and ACTN4 N-terminal concentrations were the same as those set in the single phosphorylation site model. As mentioned earlier, the only difference between models (B) and (C) is that the latter has a 50-fold weaker binding affinity to the Y4 site. All simulations were run until the default convergence criterion for steady-state levels set by BioNetGen was reached. Plotting of results here also used the same approach as mentioned for the single phosphorylation site model.

### 3.3.4 Two-site 2-kinase/1-phosphatase (2K1P) tandem model

The N-terminal region was described in this model using the following molecule type:

*nter(b1,b2,s1~U~P,s2~U~P)*

where *s1* and *s2* refer to the phosphorylation sites Y4 and Y31, respectively, while *b1* and *b2* refer to the corresponding binding sites. Both *s1* and *s2* can be either unphosphorylated (*U*) or phosphorylated (*P*) at any instant. The following reaction rules were used for describing the 2K1P model with tandem phosphorylation:

3.3.4A: *nter(b1,b2,s1~U) + kin1(b) <–> nter(b1!1,b2,s1~U).kin1(b!1)* *kin1_f, kin1_r*

3.3.4B: *nter(b1!1,b2,s1~U).kin1(b!1) –> nter(b1,b2,s1~P) + kin1(b)* *kin1_cat*

3.3.4C: *nter(b1,b2,s1~P,s2~U) + kin2(b) <–> nter(b1,b2!1,s1~P,s2~U).kin2(b!1)* *kin2_f, kin2_r*

3.3.4D: *nter(b1,b2!1,s1~P,s2~U).kin2(b!1) –> nter(b1,b2,s1~P,s2~P) + kin2(b)* *kin2_cat*

3.3.4E: *nter(b1,b2,s1~P) + pho(b) <–> nter(b1!1,b2,s1~P).pho(b!1)* *pho_f, pho_r*

3.3.4F: *nter(b1!1,b2,s1~P).pho(b!1) –> nter(b1,b2,s1~U) + pho(b)* *pho_cat*

3.3.4G: *nter(b1,b2,s2~P) + pho(b) <–> nter(b1,b2!1,s2~P).pho(b!1)* *pho_f, pho_r*

3.3.4H: *nter(b1,b2!1,s2~P).pho(b!1) –> nter(b1,b2,s2~P) + pho(b)* *pho_cat*

The description of rules 3.3.4A to 3.3.4H essentially follow that of rules 3.3.3A to 3.3.3H in the earlier 1K1P tandem model, with the only difference being that sites *b1* and *s1* are recognized by one kinase type (*kin1*) while sites *b2* and *s2* are recognized by another kinase type (*kin2*).

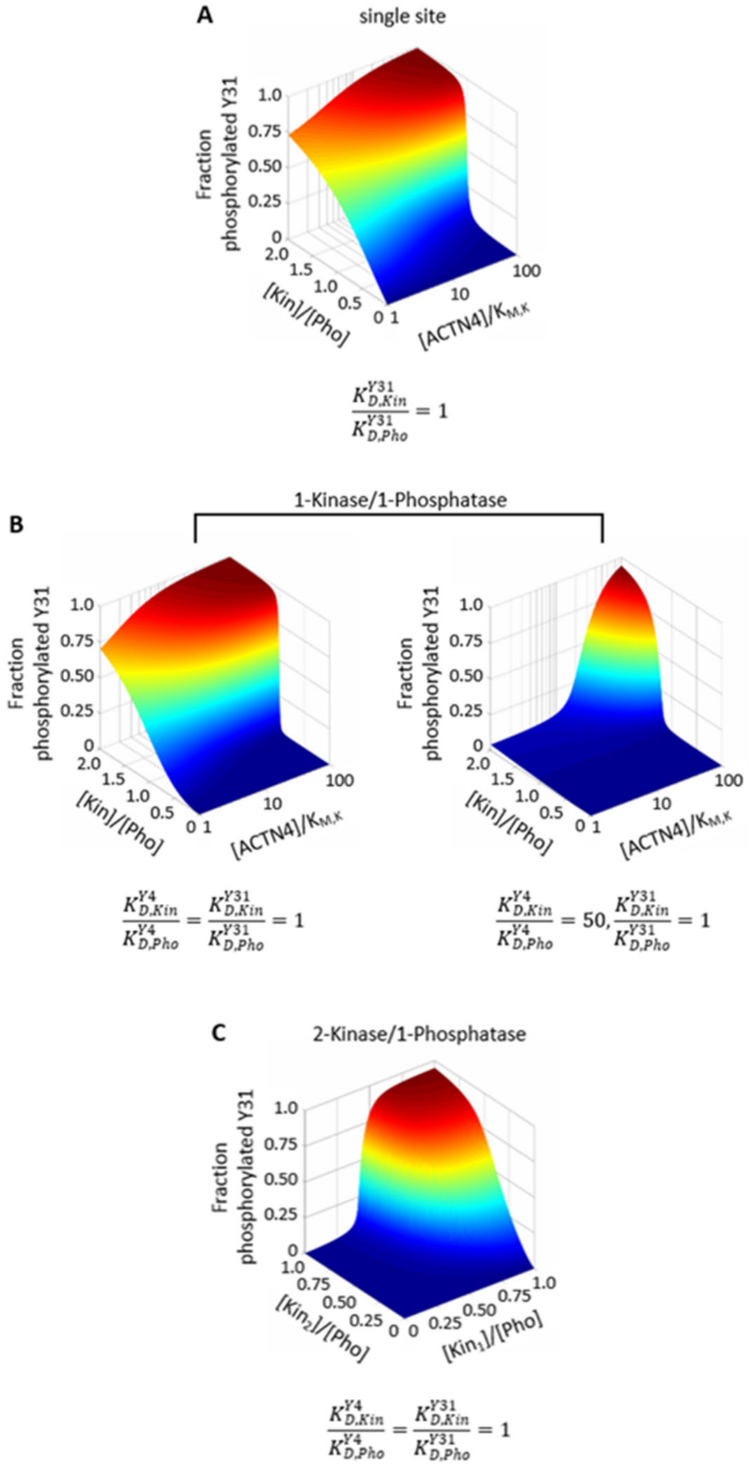
In simulating this model, the phosphatase concentration was kept fixed at 1 nM, while the concentrations of kinases 1 and 2 were varied both between 0–2 nM. The N-terminal region concentration was kept fixed at 20 μM simply to saturate the enzymes in order to see if any ultrasensitive (i.e., switch-like "on-and-off") behavior is exhibited. All simulations were run until the default convergence criterion for steady-state levels set by BioNetGen was reached. In plotting the results, the x- and y-axes were made dimensionless by normalizing the activity of both kinases by the phosphatase activity (where the enzyme activities are taken to be proportional to their concentration).

### 3.3.5 Results

We found that the phosphorylation dynamics of Y31 would not differ significantly between the models with a single phosphorylation site (Fig. 3.8A) and with two phosphorylation sites that are recognized by the same kinase (Fig. 3.8B, left plot). However, lowering the affinity of the kinase to the Y4 switch site (Fig. 3.8B, right plot) does drastically lower the overall Y31 phosphorylation. In other words, the 1K1P tandem model predicts that in order to preserve the functionality of Y31, the kinase affinity towards the Y4 switch site should be stronger or, at least, equal to that towards the Y31 site. The latter is consistent with the observation that the Y4 switch is the major site of phosphorylation upon growth factor stimulation, while the critical Y31 is substoichiometrically phosphorylated [21]. Interestingly, Fig. 3.8C shows that in a tandem model where a different kinase phosphorylates the Y4 site, it is changes in the activity of this kinase (and not the Y31 kinase) that show an ultrasensitive behavior for Y31 phosphorylation. Note that the "on-and-off" switch takes place when the Y4 kinase is at least half as active as phosphatase, and the full response (either 0% or close to 100%) occurs only when the Y31 kinase is also at least half as active as the phosphatase. This shows that the combined activities of both kinases must exceed that of the single phosphatase in order for the latter to be saturated with both phosphorylated Y4 and Y31 substrate and thus allow Y31 to become fully phosphorylated. On the other hand, changes in the activity of the Y31 kinase lead to a hyperbolic response for Y31 phosphorylation, which is especially noticeable at higher activities of the Y4 kinase. This may be because the phosphatase is already saturated with phosphorylated Y4 substrate before the Y31 kinase is turned on, and so the phosphorylated Y31 levels will rise slowly instead of abruptly as the Y31 kinase activity is increased.

## 3.4 SEARCHING FOR THE KINASES THAT ACT ON Y4 AND Y31

Identifying the putative kinase(s) that are responsible for phosphorylating Y4 and Y31 of the ACTN4 N-terminal region is a challenging problem. Initial approaches to identify candidate kinase(s) have not yielded unambiguous results (data not shown). However, both Y4 and Y31 are found within the same motif in mammalian sequences, MX**DY**XA, which suggests that a single



single site

$$\frac{K_{D,Kin}^{Y31}}{K_{D,Pho}^{Y31}} = 1$$

1-Kinase/1-Phosphatase

$$\frac{K_{D,Kin}^{Y4}}{K_{D,Pho}^{Y4}} = \frac{K_{D,Kin}^{Y31}}{K_{D,Pho}^{Y31}} = 1$$

$$\frac{K_{D,Kin}^{Y4}}{K_{D,Pho}^{Y4}} = 50, \frac{K_{D,Kin}^{Y31}}{K_{D,Pho}^{Y31}} = 1$$

2-Kinase/1-Phosphatase

$$\frac{K_{D,Kin}^{Y4}}{K_{D,Pho}^{Y4}} = \frac{K_{D,Kin}^{Y31}}{K_{D,Pho}^{Y31}} = 1$$

**Figure 3.8: Systems modeling of tandem phosphorylation for linked sites.**

Modeling results for ACTN4 N-terminus phosphorylation using (**A**) a single phosphorylation site system; (**B**) a 1-kinase/1-phosphatase (1K1P) tandem phosphorylation system where the kinase has either similar binding affinity to both sites (right) or stronger binding affinity to Y31 (left); and (**C**) a 2-kinase/1-phosphatase (2K1P) tandem phosphorylation system. In (**A**) and (**B**), the fraction of phosphorylated Y31 is plotted as a function of the kinase concentration (normalized by the phosphatase concentration) and the total N-terminal concentration (normalized by the kinase/phosphatase $K_M$ and plotted on a log scale). In (**C**), the fraction of phosphorylated Y31 is plotted as a function of the concentrations of kinases 1 and 2 ($Kin_1$ and $Kin_2$; both normalized by the phosphatase concentration) that act on Y4 and Y31, respectively. For the 2K1P model, the N-terminal region concentration was kept fixed at 20 μM to saturate the enzymes. All enzyme concentrations are assumed to be directly proportional to their activities.

class of kinase (or even the same kinase) likely recognizes and phosphorylates both sites. To test the significance of this motif in EGF-induced phosphorylation of the ACTN4 N-terminal, we mutated the aspartic acid residue before each phosphorylatable tyrosine to see if this would impair kinase activity. MD simulations indicate that Y31 gets exposed for both the D3A and D30A mutants confirm this prediction (Fig. 3.9A). With both Y4 and Y31 exposed in these mutants, any changes in phosphorylation levels can then be attributed to the mutations directly affecting kinase activity, rather than to phosphorylatable tyrosines being buried and kinase inaccessible. Phosphorylation detection assays (Fig. 3.9B) show that D3A has a reduced phosphorylation level compared to WT, but still higher than Y4E indicating that D3A impairs (and does not completely abolish) Y4 phosphorylation. The similar applies to D30A in terms of its effects on Y31 phosphorylation (D30A has higher phosphorylation than Y31E). The amount of phosphorylation observed for D3/30A is lower than both D3A and D30A, but slightly higher than Y4/31E, which is consistent with the above.

We performed a scan of potential kinases for the Y4 and Y31 sites using the PhosphoNET online resource (http://www.phosphonet.ca, [109]), and found that the most likely candidate kinases are all membrane-bound receptor tyrosine kinases (RTKs), with several that are either downstream of or exhibit crosstalk with EGFR (Table 3.1). This would be consistent with ACTN4 being predominantly localized at the membrane to participate in the regulation of actin cytoskeleton remodeling [110, 111].

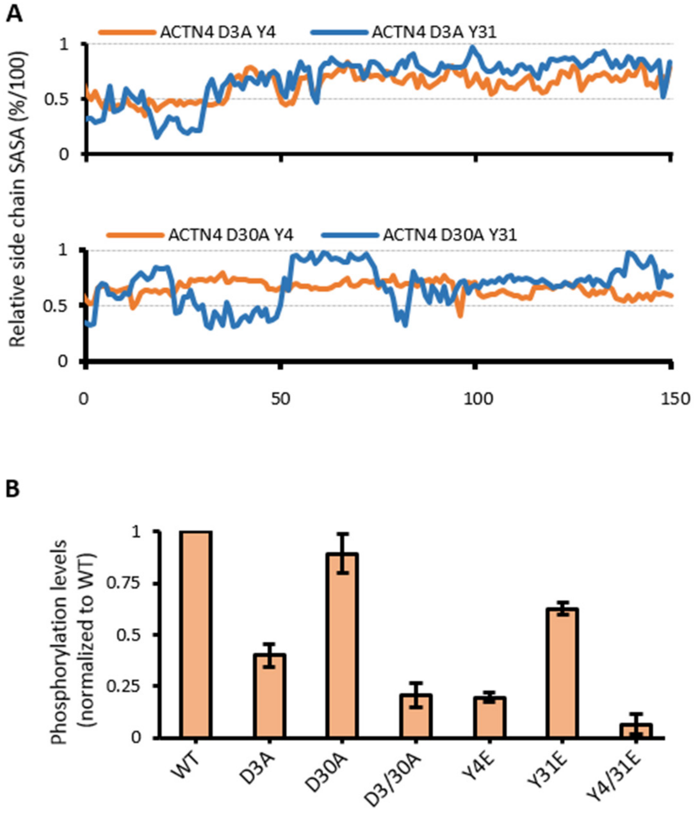


**Figure 3.9: Mutations in the MXDY$_{4,31}$XA motif decrease tyrosine phosphorylation.**

(**A**) Plots of side chain solvent-accessible surface area (SASA) over 150-ns MDs for Y4 and Y31 from D3A (top) and D30A (bottom) ACTN4 mutants. In addition to keeping the Y4 side chain exposed, the Y31 side chain also becomes exposed in both mutants. (**B**) Phosphorylation detection assays were performed for WT, D3A, D30A, D3/30A, Y4E, Y31E and Y4/31E constructs of ACTN4. The plot shows the quantitation of two experiments for amount of EGF-induced phosphorylated tyrosine normalized to WT levels. Error bars indicate s.e.m. The experiments were performed by Dr. Hanshuang Shao, under the supervision of Dr. Alan Wells, both from the Department of Pathology, University of Pittsburgh.

**Table 3.1: Top three RTKs predicted by PhosphoNET for ACTN4 Y4 and Y31.**

| Y4 | Y31 |
|---|---|
| AXL[a] (P30530) [112] | MERTK[a] (Q12866) |
| RET (P07949) [113] | AXL[a] (P30530) [112] |
| Tyro3[a] (Q06418) | EGFR (P00533) |

In parentheses are the corresponding UniProt IDs for each RTK.
In square brackets are references which show that the RTK is either downstream of or exhibits crosstalk with EGFR.
[a] Tyro3, AXL, and MERTK are all part of the TAM family of RTKS [114].

## 3.5 LOCALE-SPECIFIC REGULATION OF ACTN4 N-TERMINAL PHOSPHORYLATION VIA M-CALPAIN CLEAVAGE

To understand the evolutionary pressure to design the alternate switch site at Y4 in the ACTN4 N-terminal region, we first note that we have previously determined that ACTN4 contains a primary m-calpain cleavage site between residues 13 and 14 in the N-terminal [36]. As cleavage at this site removes Y4 permanently, we confirmed via MD simulations that the truncated protein (removing residues 1 to 13) keeps Y31 buried and inaccessible to its kinase (Fig. 3.10A). The latter was also experimentally validated by a phosphorylation assay of the truncated mutant that showed significantly reduced Y31 phosphorylation (Fig. 3.10B) and by the observation that cleavage was not found to directly affect actin binding [36]. Note that this cleavage site is not conserved in the other three ACTN isoforms (Fig. 3.1). Collectively, these results indicate that by conserving both the m-calpain cleavage site and the tandem phosphorylation sites in the ACTN4 sequence, mammalian cells have incorporated a mechanism by which the Y4 switch can be permanently turned off, thereby disabling the main regulatory pathway to inhibit the binding of ACTN4 to actin filaments.

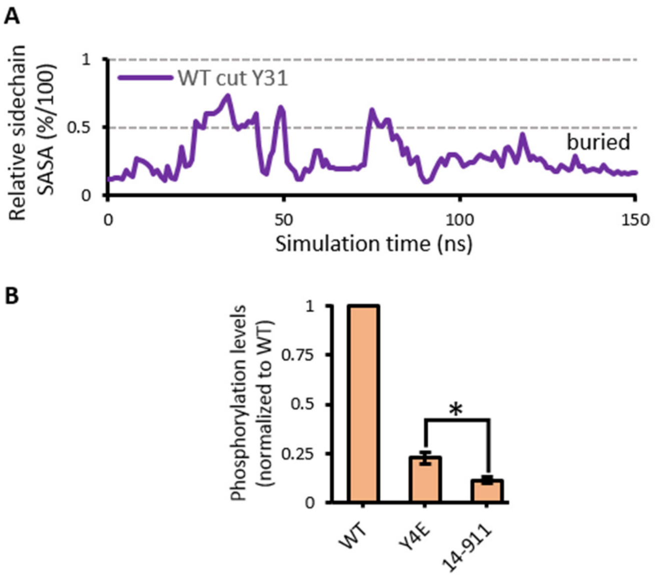


**Figure 3.10: m-Calpain cleavage of the ACTN4 N-terminal disables Y31 phosphorylation.**

(A) Relative side chain solvent-accessible surface area (SASA) for the Y31 side chain over a 150-ns MD simulation from unphosphorylated Δ1-13 (CAPN2-cleaved) ACTN4. (B) Phosphorylation detection assays were performed for WT, Y4E, and 14-911 constructs of ACTN4. The plot shows the quantitation of three experiments for amount of EGF-induced phosphorylated tyrosine normalized to WT levels. For WT, phosphorylation of Y4 and Y31 is being detected. Both Y4E and 14-911 measure Y31 phosphorylation levels only; 14-911 shows decreased Y31 phosphorylation compared to Y4E. Error bars indicate s.e.m. [* $P < 0.05$ based on student's t-test.] The experiments were performed by Dr. Hanshuang Shao, under the supervision of Dr. Alan Wells, both from the Department of Pathology, University of Pittsburgh.

From a cellular perspective, the aforementioned mechanism provides a rationale for the assembly and disassembly of locale-specific adhesion contacts during cell motility, triggered by the fact that activated m-calpain is localized only at the rear of motile cells [115, 116]. At the same time that the leading edge of the cell protrudes and makes new adhesion contacts with the outside extracellular matrix (ECM), the rear of the cell has to remain stably attached to the matrix through its own adhesion contacts. Hence, by cleaving the ACTN4 N-terminal region such that Y31 is no longer phosphorylatable even in the presence of EGF, m-calpain stabilizes the crosslinks made by ACTN4 between actin filaments and the adhesion plaques at the rear contacts. The m-calpain cleavage of talin later loosens these linkages at the rear when the rest of the cell body needs to move forward [117]. On the other hand, at the leading edge where m-calpain is not present, EGF-induced tandem phosphorylation of the ACTN4 N-terminus leads to reduced actin binding, which would then allow for more dynamic restructuring of the actin cytoskeletal network as the leading edge extends forward and forms new adhesion contacts. In other words, tandem phosphorylation of ACTN4 in conjunction with m-calpain leads to locale-specific function within the cell such that EGF induces reduction of actin binding by ACTN4 at the leading edge, while m-calpain cleavage acts as the toggle to block this reduction at the rear of the cell (Fig. 3.11). This locale-specific function is not possible using ACTN1, with its shorter and more hydrophilic N-terminal that has Y12 (homologue of ACTN4 Y31) always exposed. This means that with ACTN1 alone, adhesion contacts at the rear end of the cell would also be less stable (Fig. 3.12), and so the lack of a base to coordinate cell spreading would make it difficult for the cell to move forward in a directed fashion.

## 3.6 SUMMARY

In this chapter, we have shown that a novel tandem phosphorylation mechanism is involved in the EGF-induced reduction of actin binding by ACTN4 upon phosphorylation of Y4 and Y31 in its disordered N-terminal region. Only phosphorylated Y31 is necessary and sufficient to decrease actin binding, however Y31 is buried within the rest of the N-terminal while Y4 is solvent exposed in the unphosphorylated protein. Y4 thus gets phosphorylated first, and the

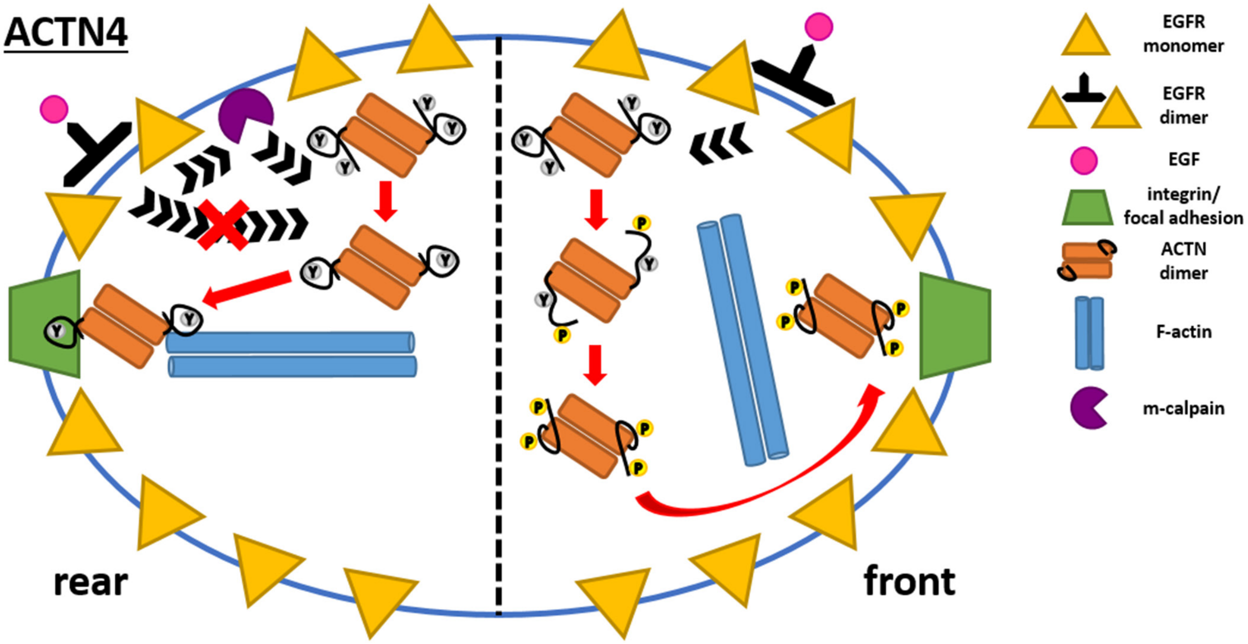


**Figure 3.11: Tandem phosphorylation and m-calpain cleavage of ACTN4 allow for locale-specific functionality in cell motility.**

At the leading edge (front) of the cell, activated EGFR leads to dual phosphorylation of Y4 and Y31 in the ACTN4 N-terminal, which causes reduced actin binding. This leads to reduced crosslinking of actin filaments with other filaments and integrins at focal adhesion sites, thereby allowing for more dynamic restructuring of the actin cytoskeleton to allow the leading edge to move forward and make new adhesion contacts. While this is happening, the rear of the cell must initially remain in place and keep stable its own adhesion contacts with the ECM. Since EGFR is distributed equally over the entire surface of the cell [118], then ACTN4 at the rear could potentially be doubly phosphorylated at the N-terminal and lead to the same phenotype as found at the leading edge. However, m-calpain is localized only at the rear of motile cells and also gets activated upon activation of EGFR. This allows m-calpain to cleave the ACTN4 N-terminal between Y13-G14, keeping Y31 inaccessible to its kinase. The crosslinks between actin filaments and rear adhesion contacts can then be kept stable while the leading edge is inching forward.

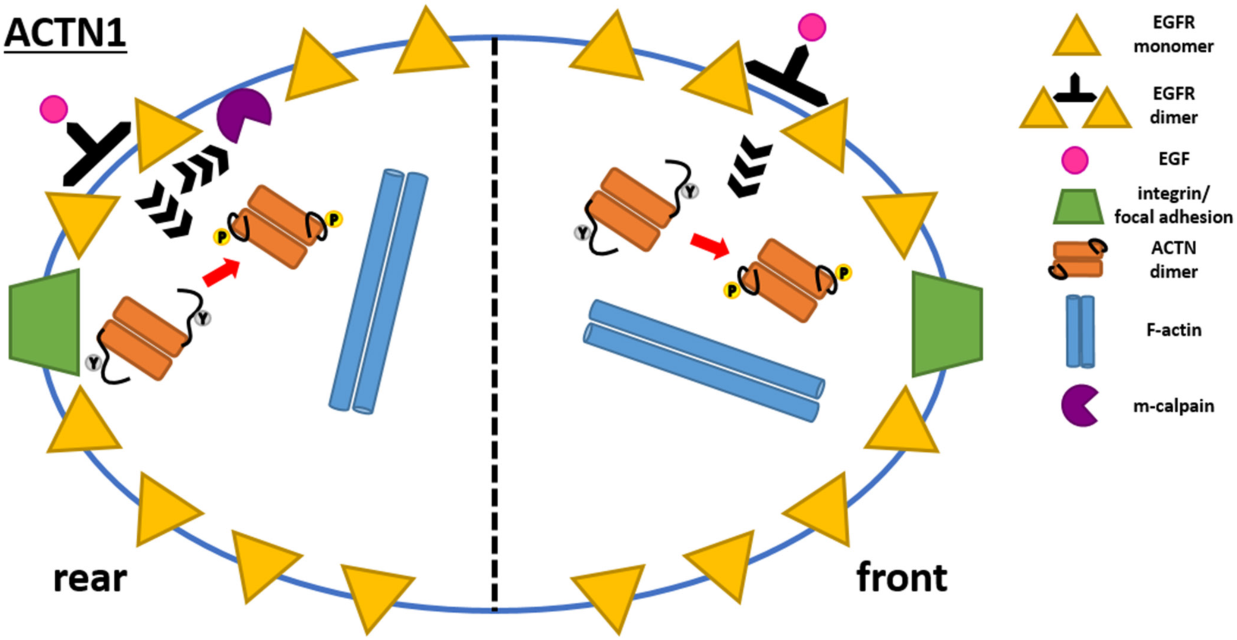


**Figure 3.12: ACTN1 alone would weaken crosslinks between actin filaments and focal adhesions at both the front and rear of the cell.**

ACTN1 has only one N-terminal phosphorylation site at Y12, which is homologous to Y31 of ACTN4. At the front of the cell, ACTN1 would lead to the same phenotype of more dynamic actin cytoskeleton remodeling upon EGFR activation. This phenotype would also be seen at the rear end of the cell, however, if only ACTN1 were present as activated m-calpain would not act on the ACTN1 N-terminal such that Y12 becomes kinase-inaccessible. Locale-specific functionality during cell motility would thus not be possible with only the always exposed Y12 phosphorylation site of ACTN1.

effect of this is a change in the molten globule-like state of the N-terminal such that Y31 becomes exposed and kinase accessible. This leads to phosphorylation of Y31, which can then latch both CH domains of the ABD in the closed conformation such that actin binding is reduced. Using a systems modeling approach, we find that a single kinase or separate kinases acting on both sites will allow for a switch-like behavior to turn on or off the phosphorylation of the functional Y31. The likely candidate kinases for phosphorylation of these two sites are the membrane-bound receptor tyrosine kinases (RTKs). Finally, we show that the function of m-calpain cleavage of the N-terminal region is to remove the Y4 switch site such that Y31 phosphorylation is permanently disabled. We also provide a model which shows that tandem phosphorylation of the N-terminal region, in conjunction with m-calpain cleavage, can allow for locale-specific functionality between the leading edge and rear end of the cell during motility. In the next chapter, we will look at the other ways by which the cell harnesses external cues to regulate the actin-binding function of ACTN4, and try to gain insight into the molecular mechanisms behind these using the atomic model of the full homodimer developed in Chapter 2.

# CHAPTER 4

# MODELING THE MOLECULAR MECHANISMS BY WHICH EXTERNAL CUES REGULATE ACTN4 FUNCTION

## 4.1 BACKGROUND

In addition to the EGF-induced phosphorylation of Y4 and Y31 in the ACTN4 N-terminal region that was discussed in the previous chapter, there are other mechanisms by which binding of ACTNs to actin filaments is regulated: (i) phosphorylation of Y265 in the CH2 domain [21], (ii), binding of calcium ions to the CaM-like domain in the calcium-sensitive non-muscle isoforms [19, 20], (iii) limited proteolysis by the ubiquitous intracellular enzymes calpain-1 and calpain-2 [36-38], and (iv) binding of phosphatidylinositol moieties (e.g. PIP2 and PIP3) to a putative binding site on the ABD [39]. In this chapter, we will use the atomic model for the full WT ACTN4 homodimer that was developed in Chapter 2 to study the molecular mechanisms behind these external cues that lead to regulation of actin binding.

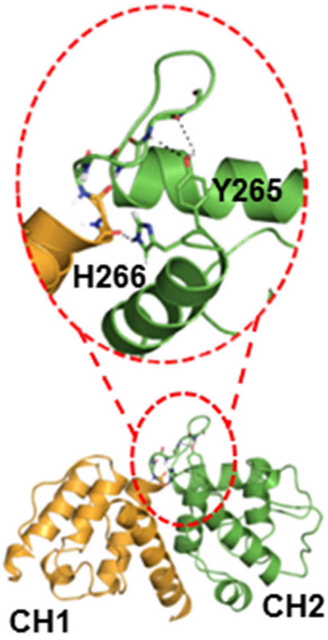


**Figure 4.1: Y265 and the H266 make contacts that stabilize the ABD closed conformation.**

Crystal structure of CH1 (orange) and CH2 (green) from ACTN4 ABD (PDB ID 2R0O [57]). The zoomed area highlights polar contacts of Y265 and H266 with the backbone of the loop that connects the two CH domains. Adapted from Figure 9A in [67].

## 4.2 PHOSPHORYLATION OF Y265 EASES ABD OPENING

Residue Y265 is located on the last helix of the ordered CH2 domain of ACTN4, and high-throughput phosphoproteomics studies have identified this as a phosphorylatable tyrosine [23, 24]. Compared to the decreased actin binding that occurs upon dual phosphorylation of the ACTN4 N-terminal region, phosphorylation of Y265 instead leads to increased actin binding [21]. Y265 is conserved in all four human ACTN isoforms, and crystal structures of the various ACTN ABDs [55-57] reveal that the hydroxyl group of Y265 makes hydrogen bonds with the

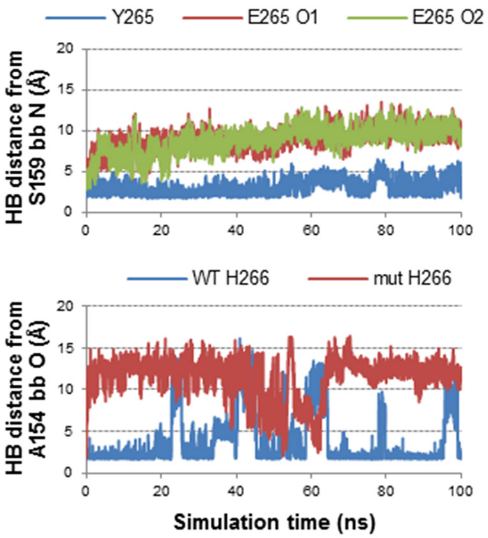


**Figure 4.2: Phosphomimic of Y265 breaks contacts that help tether CH1 and CH2 together.** Plots show the evolution of the hydrogen bond distance between the S159 backbone and the side chain at residue 265 (top) and between the A154 backbone and H266 side chain (bottom) from 100-ns explicit solvent MD simulations. For the WT simulation (blue curves), these contacts are kept mostly stable throughout the trajectory. For the Y265E simulation (magenta and green curves), the contact with the S159 backbone is not formed with either carboxylate oxygen of the glutamic acid mutation, while the contact involving H266 is broken for most of the trajectory. Adapted from Figure 9B in [67].

backbone N-H and C-O groups of S159 while the adjacent H266 interacts with the backbone oxygen atom of A154 (Fig. 4.1). These bonds help to stabilize the closed ABD form, and constrain the loop between CH1 and CH2 that contacts the neck. An MD simulation of our full WT homodimer model shows that these noncovalent bonds are highly stable, while the negatively-charged moiety of the phosphomimic mutation Y265E immediately disrupts them (Fig. 4.2). We note that the opening of the ABD structure (Fig. 2.11) does not necessarily require breaking the loop-neck interactions involving residues E162 and S164, but only the contact between E161-R280 (Fig. 2.9). These results indicate that phosphorylation at Y265 can regulate

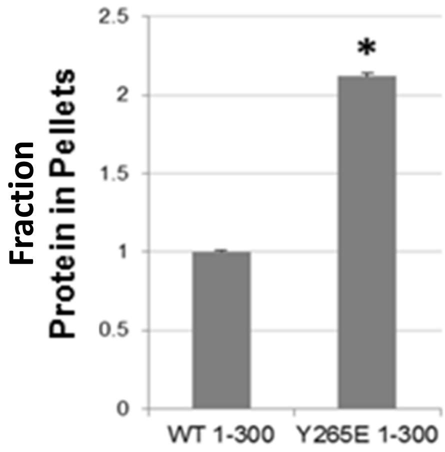


**Figure 4.3: Phosphomimic at Y265 increases actin binding of the ABD.**

Actin binding assays were performed for WT and Y265E constructs of ACTN4 residues 1-300, with ultracentrifugation separating unbound ACTN4 (supernatant or S fraction) from actin-bound ACTN4 (pellet or P fraction). The plot shows the quantitation of three experiments for amounts of ACTN4 in the P fraction. By destabilizing the closed conformation of the ABD, the Y265 phosphomimic shows highly increased binding to actin filaments compared to the WT construct. Error bars indicate s.e.m. [* $P < 0.05$ based on student's t-test.] The experiments were performed by Dr. Hanshuang Shao, under the supervision of Dr. Alan Wells, both from the Department of Pathology, University of Pittsburgh. Adapted from Figure 9C in [67].

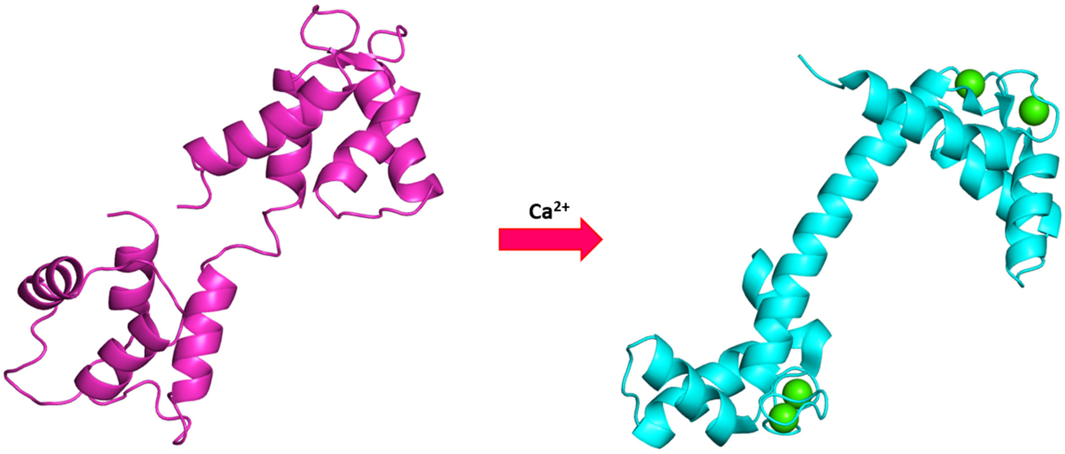
actin binding by breaking contacts with the loop connecting CH1 and CH2, thereby disengaging the CH domains and allowing the ABD to adopt an open conformation. To confirm this prediction, we performed actin-binding assays on the ACTN4 1-300 fragment, which removes any influence of other ACTN4 domains, for both WT and Y265E. The phosphomimic mutant was almost all bound to F-actin compared to the WT construct (Fig. 4.3), demonstrating that Y265E (and by extension, phosphorylated Y265) allows the ABD to be adopt a predominantly open state.

## 4.3 BINDING OF CALCIUM IONS MAY BREAK THE CAM2/NECK COMPLEX

Calcium ion concentrations of greater than $10^{-7}$ M have been shown to reduce actin binding of the calcium-sensitive non-muscle ACTN isoforms [19]. These ions bind to the calmodulin (CaM)-like domain located towards the C-terminus of the monomer; whereas calmodulin itself is capable of binding up to four calcium ions (through its four EF hand motifs), the CaM-like domain of non-muscle ACTNs can only bind a maximum of 2 ions (through the two EF hand motifs in CaM1) as mutations in CaM2 have rendered it incapable of binding calcium. Since the function of CaM2 is to bind the helical neck region in order to form a stable CH2/neck/CaM2 ternary complex in the homodimer, then it makes sense that CaM2 has lost its ability to bind calcium ions. This leaves the question, though, as to how reduced actin binding results from the binding of calcium to CaM1.

To investigate this, we first tried to draw clues from the structural changes that occur in the calmodulin protein once it binds calcium. Aside from opening the hydrophobic binding pockets of both of the globular heads, the most striking difference between calcium-free and calcium-bound calmodulin is the helical folding of the linker that connects the two heads (Fig. 4.4). As CaM2 is part of a ternary complex in the ACTN4 homodimer, a structural change such as that seen for calmodulin may likely put strain on the multi-domain assembly that would alter its structure and thereby explain the observed changes in actin binding.

We employed a targeted molecular dynamics (TMD) approach to force the helical folding of the linker between CaM1 and CaM2 in our structural model of the ACTN4 homodimer end region, and see if this folding is capable of altering the multi-domain assembly. In TMD simulations, a subset of atoms is guided towards a target structure via steering forces that are



**calcium-free calmodulin**

**calcium-bound calmodulin**

**Figure 4.4: Calcium ions trigger the helical folding of the calmodulin linker region.**

Coordinates for the calcium-free (left) and calcium-bound (right) structures of calmodulin are from PDB IDs 1CFD [82] and 1EXR [119], respectively. Calcium ions bound to the EF hand motifs are shown as green spheres.

based on the root-mean-square (RMS) distance between the current and target coordinates of the selected atoms [120, 121]. At each time step $t$, the current RMS distance for these atoms is computed, followed by the additional force on these atoms as given by the gradient of the following potential:

$$U_{TMD} = \frac{k}{2}[RMS(t) - RMS_0(t)]^2$$

where $RMS(t)$ is the instantaneous best-fit RMS distance between the current and target coordinates, $RMS_0(t)$ is a linear evolution from the initial RMS distance at the first time step to the target RMS distance at the last time step, and $k$ is the force constant for this harmonic potential. The target coordinates of the selected atoms is provided as a reference file at the start of the simulation. While TMD has been shown to generate plausible transition pathways from the initial to target structure, the main limitation of this approach is that the holonomic constraint may force the system to cross very high energy barriers such that pathways which are otherwise inaccessible at physiological temperatures are taken during the simulation [122]. This can lead to the following problems: (i) TMD does not always yield reversible pathways [123], (ii) there is a tendency for large-scale motions in the protein to precede small-scale ones [124], and (iii) the simulation does not sample an equilibrium ensemble and thus cannot be used to calculate equilibrium properties such as free energy profiles [125].

We used our model of the ACTN4 homodimer end region (Fig. 2.10), excluding the disordered N-terminal region, as the starting coordinates of the TMD simulation. For the reference structure, we built a model of a helix encompassing residues A830 to L855 and used the backbone atoms of these residues as the target subset. Solvation, charge neutralization, minimization, and position-restrained equilibration were then performed as described in the MD protocol in Appendix A, followed by a short 10 ns TMD simulation with a target RMSD of 0 Å. We found that during the course of the simulation, the helical folding of the linker between CaM1 and CaM2 was accompanied by a 90° bending of about half of the helical neck region away from the hydrophobic binding pocket of CaM2 (Fig. 4.5). The relative binding ΔG for the CaM2/neck complex in this pose was estimated at -9.5 kcal/mol using FastContact [79], around half of that for the starting structure (Table 2.1). This suggests that the helical folding of the CaM1-CaM2 linker upon binding of calcium ions to CaM1 causes the CaM2/neck complex to break. How this could lead to reduced actin binding is one topic for future research. In particular, by extending this TMD simulation beyond 10 ns, it would be interesting to see if the CaM2/neck does completely break apart and how this would affect the positions of the other domains in the assembly.

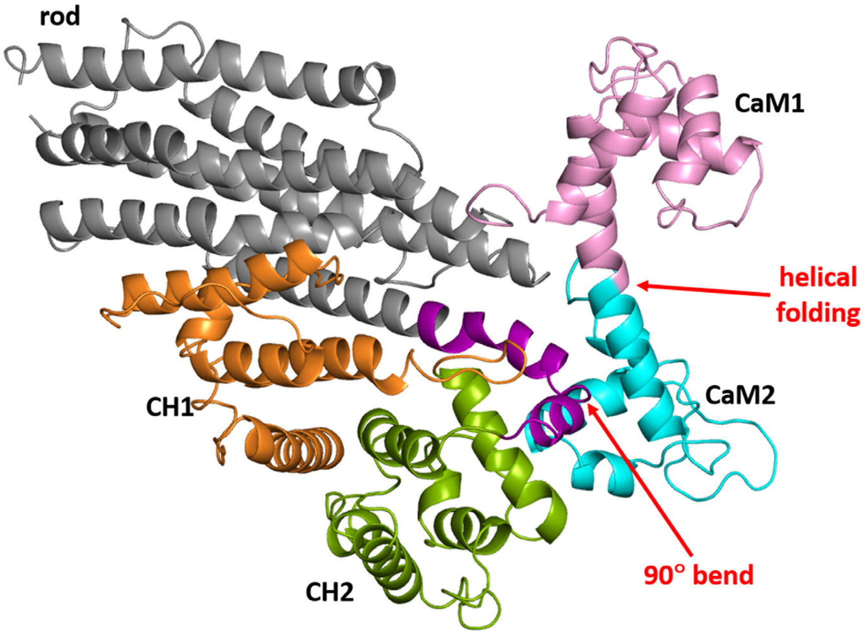


**Figure 4.5: Helical folding of CaM1-CaM2 linker may break the CaM2/neck complex.**

Final structure of the ACTN4 homodimer end region after a 10-ns TMD simulation that forces the folding of the CaM1-CaM2 linker into a helix. This structure is positioned for easy comparison with the starting structure in Fig. 2.10. Arrows point to the linker that undergoes helical folding and to the resulting 90° bending of the helical neck region.

## 4.4 CAM2 PROTECTS THE NECK FROM M-CALPAIN CLEAVAGE

The calpain enzymes area family of intracellular cysteine proteases, with the ubiquitous members being μ-calpain (calpain 1) and m-calpain (calpain 2) [126]. The names of these two members come from their *in vitro* calcium concentration requirements in order to be activated, with μ-calpain requiring micromolar concentrations and m-calpain requiring millimolar concentrations. Calpains are heterodimers; the large subunit contains the catalytic domain, a C2-like domain, and a calcium-binding domain with five EF hand motifs, and the small subunit is also a calcium-binding domain with another five EF hand motifs. The calcium requirement comes from the fact that the catalytic triad of residues (C105, H262, and N286) is not assembled in the calcium-free protein, and binding of calcium ions to the EF hand motifs presumably leads to allosteric changes that allow the formation of a functional active site [127-129]. Both μ-calpain and m-calpain are important regulators of cell migration [130]; m-calpain, in particular, is localized at the inner surface of the cell membrane towards the rear of motile cells (through its phosphoinositide-binding C2-like domain), where its activation is regulated by the phosphoinositide PIP2 [115, 116].

Two m-calpain cleavage sites were previously identified for human ACTN4: a primary site in the N-terminal between Y13-G14 and a secondary site in the neck region between K283-V284 (Fig. 4.6) [36]. The implications of cleavage at the N-terminal site, in conjunction with tandem phosphorylation of Y4 and Y31, on locale-specific regulation of actin cytoskeletal remodeling were already discussed earlier (see Chapter 3.5). For the neck site, the main implication is that cleavage at the neck would detach the ABD from the rest of the homodimer. Thus, crosslinking of actin filaments to one another or proteins within adhesion contacts is lost, even if the ABD is still able to bind to actin filaments.  In this section, m-calpain cleavage of the neck region site will be the focus.

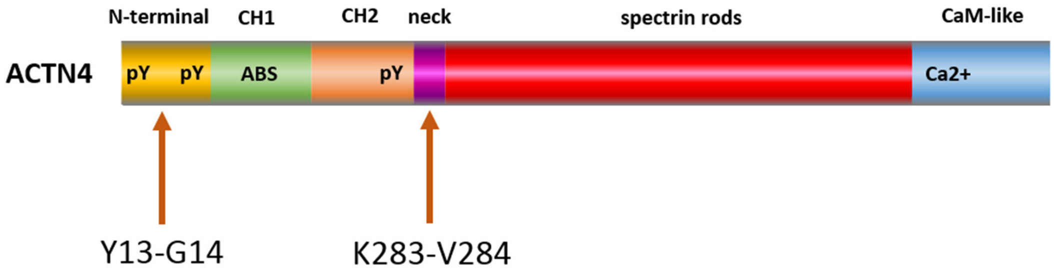


**Figure 4.6: m-Calpain cleaves human ACTN4 at two sites.**

Y13-G14 in the disordered N-terminal region is the primary site for m-calpain cleavage, while K283-V284 in the neck region is a secondary site.

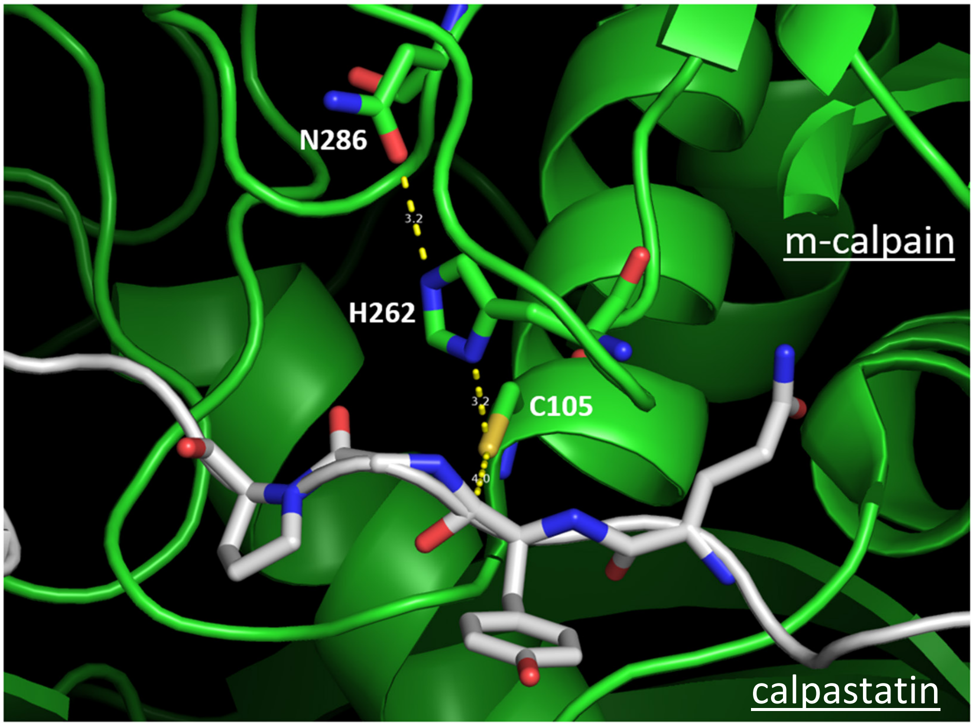


**Figure 4.7: Target sequences adopt an extended conformation when binding to m-calpain.** A crystal structure for the complex between m-calpain (colored green) and its inhibitor calpastatin (colored white) shows that the binding site near the catalytic triad (C105, H262, N286; shown in sticks) of m-calpain has only enough space to bind target peptides that adopt extended conformations (PDB ID 3BOW [131]).

We first note that based on crystal structures of m-calpain bound to its inhibitor protein calpastatin [131, 132], there is only enough room for a peptide with an extended conformation in the binding site near the catalytic triad (Fig. 4.7). This observation allows for recognition of the N-terminal cleavage site, which tends to adopt a coil structure in our MD simulations of WT N-terminal region (Fig. 3.4). Our full model of the ACTN4 homodimer region, however, has the neck folded into a helix when part of the CH2-neck-CaM2 ternary complex (Fig. 2.10). This presents a problem for binding to m-calpain, as there is not enough space to fit a helical structure in the binding site of the protease. One possibility is that the helical conformation of the neck region is only stable when in complex with at least CaM2; consistent with this, a low-resolution cryo-EM of chicken smooth muscle ACTN (PDB ID 1SJJ [66]) does not show the CaM2-neck complex as the neck region was resolved as taking an extended conformation. Assays for m-calpain cleavage indeed show that CaM2 acts as an inhibitory domain for cleavage of the neck region site [36], which supports the notion that by folding into a helix and binding to CaM2, the neck region site is protected from cleavage. Interestingly, it was also found that in truncation mutants where more than 11 residues were removed from the C-terminus of ACTN4, resistance to m-calpain cleavage at the neck site is lost and both cell migration and spreading become significantly impaired. We performed MD simulations of full and several truncation mutants of CaM2, and found that removal of more than 11 residues from the C-terminus leads to a collapse of the hydrophobic binding pocket of CaM2 (Fig. 4.8). This means that by destabilizing the structure of CaM2, the helical structure of the neck region becomes less stable such that access to K283-V284 for m-calpain cleavage is increased.

## 4.5 BINDING OF PHOSPHOINOSITIDES ALLOWS ACTN4 TO CROSSLINK ACTIN FILAMENTS NEAR THE MEMBRANE

Among the regulators of the α-actinins are several phosphoinositides, including phosphatidylinositol 4,5-biphosphate (PIP2) and phosphatidylinositol 3,4,5-triphosphate (PIP3) [133, 134]. A putative binding site for phosphoinositides was first found using binding assays with ACTN peptides generated by proteolysis, and the strongest binding peptide was found to be part of the CH2 domain of the ABD, containing the basic residues K181, H189, and K193

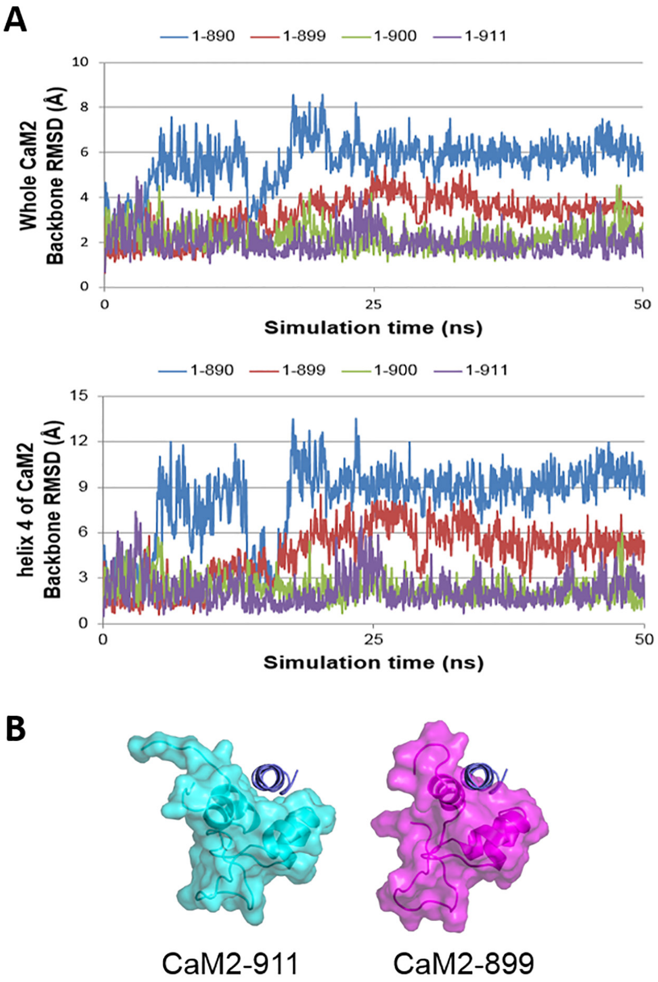


**Figure 4.8: Truncation of more than the last 11 residues of CaM2 hampers neck binding.**

(**A**) Top plot gives the backbone RMSD for MD simulations of full CaM2 (840-911) and several CaM2 truncation mutants (840-900, 840-899, and 840-890). Truncating more than 11 residues from the C-terminus shows increased backbone RMSD as the simulation progresses (840-890 and 840-899), while truncating less than or equal to 11 residues shows stable backbone RMSD over the trajectories (840-900 and 840-911). Bottom plot gives the backbone RMSD for the last helix (helix 4) of CaM2 alone, and shows that much of the increased backbone RMSD seen in the top plot for 840-890 and 840-899 is due to movements of this particular helix. (**B**) Snapshots from the full (840-911; left) and truncated (840-899; right) CaM2 simulations with CaM2 shown in surface for both. The helical neck region (not part of the MD simulations) is drawn in cartoons and along its helical axis to show that truncation of more than 11 residues destabilizes helix 4 such that it collapses the hydrophobic binding pocket, thereby preventing formation of the CaM2-neck complex.

(numbering based on human ACTN4 sequence, [39]). Among these three residues, K181 has found to be the most important for binding of phosphoinositides [10]. Since this candidate binding site was inferred using peptide fragments, however, there is a possibility that there may be other relevant residues that are spatially close to this peptide in the tertiary structure but faraway sequence-wise. Indeed, analysis of the electrostatics for the surface of the human ACTN3 crystal structure suggests a different set of three basic residues on the CH2 domain to be important for phosphoinositide binding: R175, K181, and R204 [56]. Residue K181 is the only residue common to both candidate sites, consistent with the experimental evidence regarding its significance to binding of phosphoinositides.

As there are currently no experimentally-solved structures the complex of a phosphoinositide with the ABD of any ACTN, we used a virtual docking approach to gain insight into how phosphoinositides can bind to the ABD. Here, we used AutoDock Vina [135] to dock the negatively-charged inositol head group of PIP3 on the ABD of human ACTN4. The force field topology and parameters for the PIP3 ligand were generated using the SwissParam web interface [136], which uses (i) the Merck Molecular Force Field (MMFF) [137-141] to derive the bond, angle, dihedral, and improper dihedral energy terms as well as the atomic partial charges, and (ii) the CHARMM22 force field [142] for the atomic van der Waals parameters. We first did a docking screen using rigid side chains on the ABD to speed up the search. Interestingly, the resulting top-ranked poses were clustered near K181, H189, and K193, the set of basic residues identified in the peptide fragment analysis. We then performed a docking screen with flexible side chains for these three basic residues as well as other nearby polar residues (N182, N187, H189, K217, and R251) to possibly find a more energetically favorable

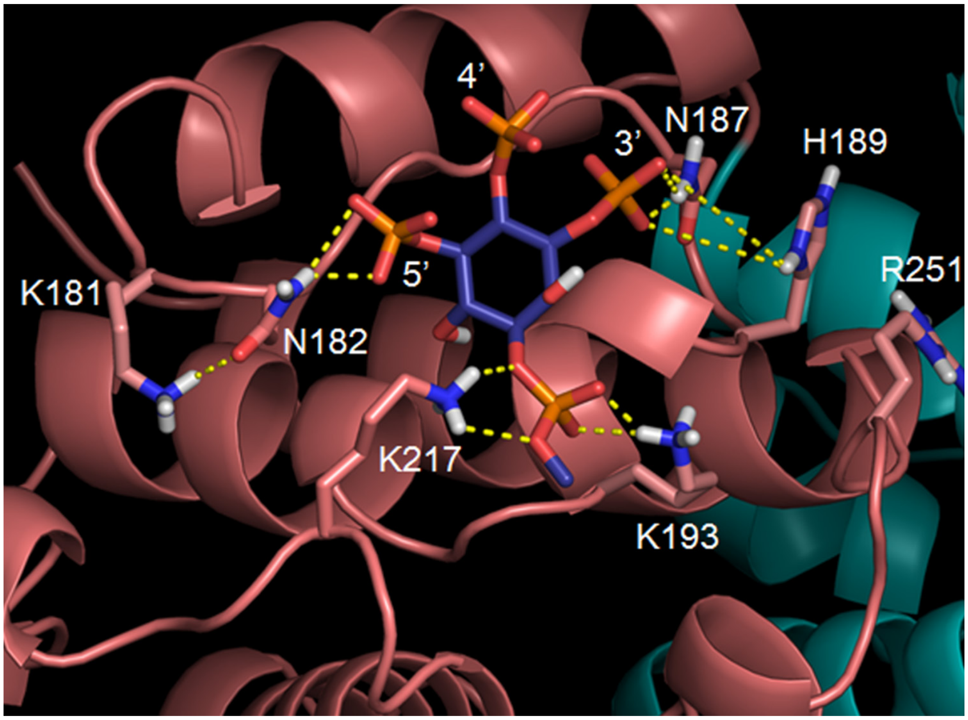


**Figure 4.9: Virtual docking predicts that the PIP3 head group binds to the CH2 domain.**

CH1 and CH2 of the ABD are shown in cartoons and colored pink and cyan, respectively. Shown in sticks are the inositol head group of PIP3 in the middle, along with several polar residues (including K181, H189, and K193) that stabilize the binding of the PIP3 head group to CH2. The 1’-phosphate of the ligand is shown at the bottom as interacting with K193 and K217; the fatty acid tails of PIP3 would extend downward from this phosphate.

positioning of these residues with respect to the ligand. The top-ranked pose from this second screen is shown in Fig. 4.9; this pose shows the 1’-phosphate interacting with K193 and additionally K217, and the 3’-phosphate interacting with H189 and additionally N187. K181 does not interact directly with the ligand, but is actually bridged by the N182 side chain with the 5’-phosphate of the ligand. Binding assays suggest that the ABD binds both PIP2 and PIP3 with same affinity [41], suggesting that the presence of the 5’-phosphate is not crucial for binding and can be replaced with a hydroxyl group (which is found in the 5’ position of PIP2). Further research that we plan on doing here will involve additional computational validation of this docked pose, for instance by performing MD simulations to see if the inositol head group remains stably bound to CH2 in the presence of solvent and ions.

Since phosphoinositides are concentrated on the cytosolic face of plasma membranes [143] and can thus target ACTN4 molecules to the inner membrane, we wondered if our atomic model of the full ACTN4 homodimer (Fig. 2.10) would allow for this using the top-ranked binding pose between the PIP3 head group and CH2 domain. As shown in Fig. 4.10, our full model allows for the full homodimer to localize near the membrane, without any protein domains making clashes with it, using the binding pose in Fig. 4.9. More importantly, using our model with open ABD shows that the homodimer can crosslink actin filaments in such a way that these are parallel to the plane of the membrane. Since the actin cytoskeleton plays an important role near the membrane in maintaining cell shape, the model shown in Fig. 4.10 offers the possibility for the involvement of ACTN4 as a key player in this process.

## 4.6 SUMMARY

In this chapter, we used the atomic model for the full WT ACTN4 homodimer that was developed in Chapter 2 to study the molecular mechanisms that allow various external cues to regulate the actin-binding function of ACTN4. Phosphorylation of Y265 breaks contacts involving Y265 and H266 to ease the opening of the ABD. Binding of calcium ions to CaM1 may break the CaM2/neck complex, and is still a topic for future research. m-Calpain cleavage of a site located in the neck region will destroy the ability of the homodimer to crosslink actin filaments, however this site is protected from the protease in the CaM2/neck complex. Finally, a

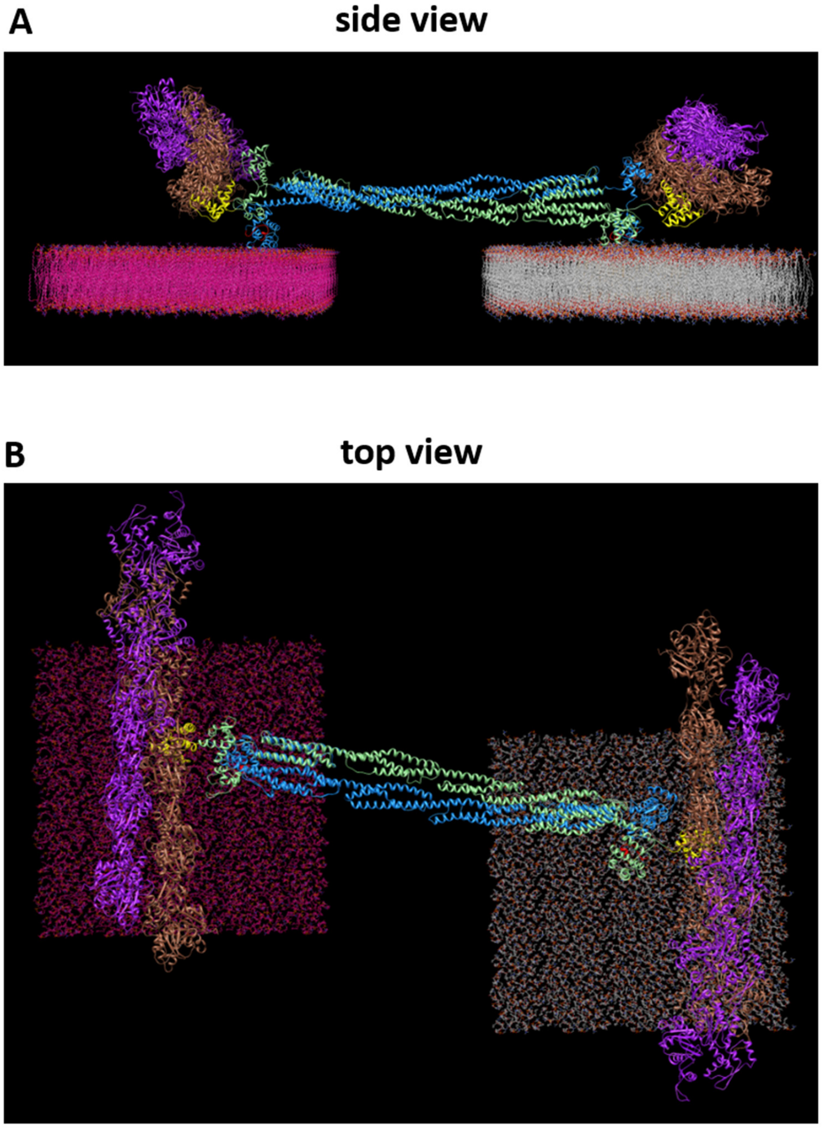


**Figure 4.10: Actin crosslinking at the membrane using our full ACTN4 homodimer model.**

(**A**) Side view and (**B**) top view of a model for crosslinking of actin filaments by the ACTN4 homodimer at the membrane surface. The two antiparallel monomers of ACTN4 are shown in the middle in blue and green. Double helical actin filaments (brown and purple) are shown on both ends of the ACTN4 homodimer. Membranes (shown with pink or white hydrophobic tails) were generated using the Membrane Plugin of VMD [81]. One of the head groups on the left side membrane was replaced with the PIP3 inositol head group, and the ACTN4 homodimer was added to the model using the PIP3/CH2 binding pose in Fig. 4.9. Structure of the actin filaments and their binding to the CH1 domain of ACTN 4 were taken from PDB ID 3LUE [18]. Model was built and rendered using Chimera [144].

binding site for phosphoinositides in the CH2 domain can localize ACTN4 to the surface of the inner membrane, and our full model of the homodimer shows that ACTN4 can crosslink actin filaments such that these are parallel to the surface of the membrane. In the next chapter, we will take our first steps in developing a network-level model that can tie together most of these mechanisms into a computational tool for predicting the degree of actin binding by ACTN4 in the presence of multiple signals.

# CHAPTER 5

# A NETWORK-LEVEL MODEL FOR PREDICTING THE ACTIN-BINDING RESPONSE OF ACTN4 TO MULTIPLE SIGNALS

## 5.1 BACKGROUND

In the previous chapters, the molecular mechanisms by which external cues lead to regulation of the actin-binding function of ACTN4 were investigated primarily using a host of computational structural methods. These methods led to insights that allowed us to: (i) build an all-atom structural model for the full ACTN4 homodimer; (ii) discover that a novel tandem phosphorylation mechanism that acts in protein IDRs is involved in the inhibition of actin binding seen upon EGF-induced of two tyrosine residues (Y4 and Y31) in the unstructured N-terminal region of ACTN4; and (iii) study the mechanisms that lead to ACTN4 regulation via Y265 phosphorylation, binding of calcium ions, m-calpain cleavage of site in the N-terminal and neck regions, and binding of phosphoinositides. There is likely a dynamic interplay among these different cues such that actin crosslinking by ACTN4 is under the influence of multiple simultaneous cues at a particular time, however experimental data on upregulation or downregulation of ACTN4 function is available only for individual cues that are acting alone. An interesting and useful question is if we can use the structural insights gained in the previous chapters to develop a quantitative model that can accurately predict the actin-binding response of ACTN4 when multiple external cues come into play.

## 5.2 A TWO-STEP BINDING MODEL FOR WT AND PY265 ACTN4

As a step in this direction, we first use a two-step binding model to describe the binding of WT ACTN4 homodimer to actin filaments. The two steps correspond to (i) opening of the ABD that frees up the actin-binding site (ABS) on CH1 to bind to actin, followed by (ii) actual binding of the opened ABD to actin filaments (Fig. 5.1). From the ordinary differential equations (ODEs) that describe the reactions rates for these two steps, we can derive the following equation for the overall $K_D$ of binding one end of the ACTN4 homodimer to actin:

$$K_D = K_{bind} \times [(K_{open})^{-1} + 1]$$

where $K_{open}$ gives the rate of ABD opening and $K_{bind}$ gives the binding rate of the ABS on CH1 (with the ABD already in an open conformation) to actin.

This two-step binding model should also be applicable to ACTN4 with phosphorylated Y265, since this modification will enhance the ABD opening rate but without disturbing the CH2/neck/CaM2 ternary complex or any other domains (Chapter 4.2). Experimental binding curves from our collaborators (Dr. Alan Wells and Dr. Hanshuang Shao, Department of Pathology, University of Pittsburgh) allowed us to estimate via nonlinear regression the

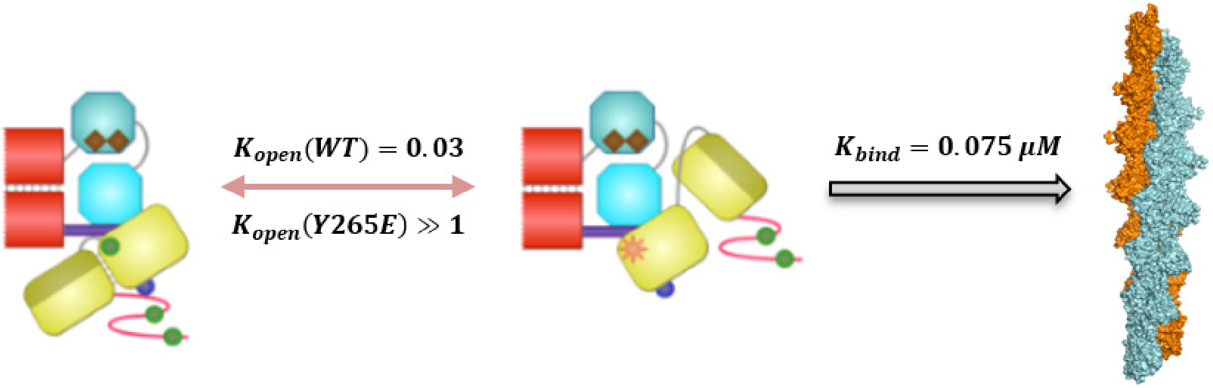


**Figure 5.1: Two-step binding model for WT ACTN4 homodimer.**

The following are shown for the ACTN4 homodimer end region (left and middle): N-terminal region (pink), ABD (yellow), neck (purple), spectrin rods (red), CaM1 (dark blue), and CaM2 (light blue). Also shown are the tyrosine phosphorylation sites (green circles), calcium binding sites (brown diamonds), and ABS2 on CH1 (hashed area). The ABD opening rates for WT and pY265 ACTN4, and the binding rate of exposed CH1 to actin, were obtained by solving our two-step binding $K_D$ equation, using $K_D$ values for WT and pY265 ACTN4 that were estimated from binding curves from our experimental collaborators.

overall $K_D$ for WT and pY265 (using the Y265E phosphomimic) ACTN4 as 2.6 ± 0.4 μM and 0.075 ± 0.02 μM, respectively. Note that $K_{bind}$ gives the binding rate of the exposed ABS on CH1 (implying open ABD) to actin, so this parameter will have the same value for both WT and pY265 ACTN4. Since Y265E shifts towards predominantly open ABD (i.e., $K_{open}(pY265) >> 1$) without affecting either $K_{bind}$ or the integrity of the open CH2/neck/CaM2 ternary complex, the above equation implies that $K_{bind}$ = 0.075 μM and $K_{open}(WT)$ = 0.03. These values predict that around 43% (WT) and 91% (pY265) of ACTN4 molecules will be found in complex with actin at equilibrium, in agreement with typical values seen in actin-binding assays. More importantly, these values reveal aspects of the molecular basis of ACTN4 regulation that are very difficult to probe experimentally. For instance, the parameter for $K_{open}(WT)$ suggests that the ABD opens on average only once for every 34 WT ACTN4 molecules.

## 5.3 A NETWORK-LEVEL MODEL RECAPITULATES SEMI-QUANTITATIVE ACTIN BINDING DATA

We were then able to build upon the two-step binding model to incorporate transitions based on our structural models for N-terminal phosphorylation at Y4/Y31 and binding of calcium ions. This network-level model was implemented in the SimBiology Toolbox of MATLAB, using the following 14 reversible reactions:

5.3A: $C_{tri} <-> O_{tri}$

5.3B: $A + O_{tri} <-> A \cdot O_{tri}$

5.3C: $C_{Pnter} <-> O_{Pnter}$

5.4D: $A + O_{Pnter} <-> A \cdot O_{Pnter}$

5.4E: $C_{tri} <-> C_{cal}$

5.4F: $O_{tri} <-> O_{cal}$

5.4G: $C_{cal} <-> O_{cal}$

5.4H: $A + O_{cal} <-> A \cdot O_{cal}$

5.4I: $C_{Pnter} <-> C_{Pnter,cal}$

5.4J: $O_{Pnter} <-> O_{Pnter,cal}$

5.4K: $C_{Pnter,cal} <-> O_{Pnter,cal}$

5.4L: $A + O_{Pnter,cal} <-> A \cdot O_{Pnter,cal}$

5.4M: $O_{tri}$ <-> $O_{Pnter}$

5.4N: $O_{cal}$ <-> $O_{Pnter,cal}$

where:

$C_{tri}$, $O_{tri}$ = closed and open ABD for ACTN4 with intact ternary complex (row II)

$C_{Pnter}$, $O_{Pnter}$ = closed and open ABD for ACTN4 with phosphorylated N-terminal (row III)

$C_{cal}$, $O_{cal}$ = closed and open ABD for calcium-bound ACTN4 (row I)

$C_{latch,cal}$, $O_{latch,cal}$ = closed and open ABD for calcium-bound ACTN4 with phosphorylated N-terminal (row IV)

$A$ = unbound actin filaments

$A \cdot O_{tri}$, $A \cdot O_{latch}$, $A \cdot O_{cal}$, $A \cdot O_{latch,cal}$ = complexes of actin filaments bound with a particular open ABD ACTN4 form

where the row numbers (I to IV) indicate where the species is located in the diagram of the network-level model shown in Fig. 5.2. The above reversible reactions were solved until steady-state levels were reached using the ode45 solver of MATLAB, with the ordinary differential equations (ODEs) governing the rates of change in concentration for each chemical species generated automatically by MATLAB. Each reversible reaction requires user-specified forward and reverse rate parameters. Units of micromolarity (μM) were used for all species concentrations. Most of the reactions are unimolecular transitions except for those that involve the binding of an open ABD ACTN4 form to actin. For the latter, the forward rate parameters have units of 1/(μM · s), while the unimolecular transitions (and the reactions involving the unbinding of open ABD from its complex with actin) have units of 1/s.

In all the simulations the starting concentration of actin is set at 2.99 μM while all ACTN4 are initialized in the $C_{tri}$ form with a concentration of 2.38 μM. These concentrations match those used by our experimental collaborators for *in vitro* actin-binding assays under a variety of conditions (ACTN4 with either WT, Y265E phosphomimic, or Y4/Y31 double phosphomimic, and either with or without calcium ions). The network-level model shown in Fig. 5.2 consists of four interconnected binding pathways with eight predicted states for different multi-domain complex structures. The four layers in this model denote: actin binding of WT and pY265 (or Y265E) ACTN in the presence (I) or absence (II) of calcium ions, and actin binding of pY4/pY31 (or Y4/31E) in the presence (III) or absence (IV) of calcium ions. The rate parameters are depicted in Fig. 5.2 as the ratio of the reverse rate over the forward rate parameters; values in black are based on the earlier two-step binding model while values in red were manually fit. The network-level model accounts for increased actin binding upon Y265 phosphorylation, decreased actin binding upon Y4/Y31 phosphorylation, decreased actin binding for WT ACTN4 in the presence of calcium, and observations that pY265 and pY4/pY31 are calcium-insensitive in terms of actin binding. Table 5.1 shows that we were able to manually fit the parameters that did not come from the earlier two-step binding model, such that there is good agreement between the values predicted by the network model and semi-quantitative experimental values for the amount of bound ACTN4 under the conditions listed above.

The next step for future research would be to test the robustness of the fitted parameters in this network model under varying initial conditions (e.g., varying concentrations of actin and ACTN4). In addition, a complete network-level model would include m-calpain cleavage and binding of phosphoinositides; the particular challenge with these is the compartmental nature of their function (e.g., m-calpain cleavage is localized at the rear end of motile cells, phosphoinositide binding takes place the inner membrane surface). Finally, this model simplifies the representation of actin as a mixture of monomers instead of as a "rigid" filamentous structure. The latter in fact imposes a structural constraint on how ACTN4 molecules crosslink actin; experimental binding curves for WT ACTN4 have shown that at most 1 out of every 4 actin monomers are bound to ACTN4 even under saturating conditions [145]. This observation was explicitly factored into the model by dividing the total actin concentration, instead of being a result that naturally emerges from the model.

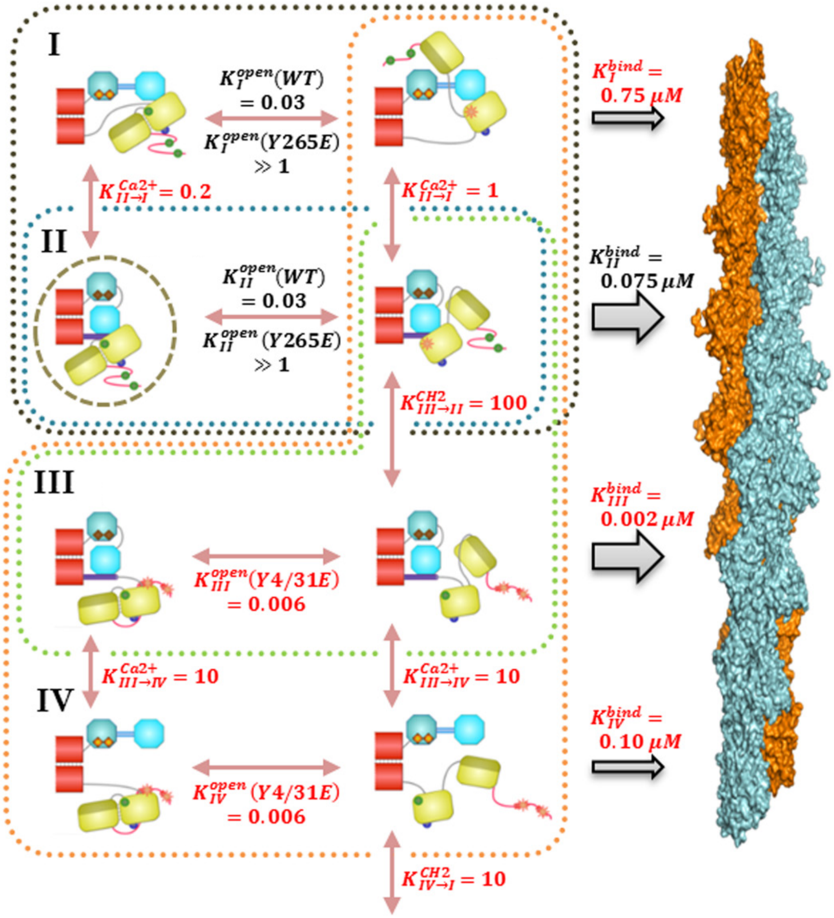


**Figure 5.2: Network-level model of actin binding by ACTN4 under various external cues.**

The description of ACTN4 in each of the eight predicted states follows that given in Fig. 5.1. Key rates are shown for the different transitions, with model-determined and fitted parameter values shown in black and red, respectively. Coordinates for F-actin structure were taken from PDB ID 3MFP [146].

**Table 5.1: Comparison of experimental and model-predicted amounts of bound ACTN4.**

| Construct | Calcium ions | Experimental[a] | Predicted[a] |
|---|---|---|---|
| WT | – | 100[b] | 100 |
| WT | + | 79 | 74 |
| Y265E | – | 208 | 210 |
| Y265E | + | 208 | 209 |
| Y4/31E | – | 57 | 53 |
| Y4/31E | + | 55 | 54 |

All values are normalized relative to actin binding for WT ACTN4 in the absence of calcium.

[a] Expressed as percentage after normalization; results from single experiment.

[b] Actin-binding assays show approximately 40-50% bound ACTN4 before normalization.

## 5.4 SUMMARY

In this chapter, we have developed a preliminary network-level model, using parameters derived from a two-step binding model of WT and pY265 ACTN4 as well as manually fitted parameters, that is able to recapitulate semi-quantitative experimental actin-binding data for WT, Y265 phosphomimic, and Y4/Y31 double phosphomimic conditions, either in the presence or absence of calcium ions. Pending further robustness testing of our fitted parameters, this model can serve as the foundation for building a computational tool for predicting the actin-binding response under multiple external cues. In addition to incorporating other conditions, such as m-calpain cleavage and phosphoinositide binding, that require adding the notion of compartmentalization to the model, a further improvement would be to be able to translate this model such that it is geared more towards the modeling of a cell-based environment than an *in vitro* assay environment (e.g., phosphorylation is expressed in terms of kinase and phosphatase reactions, instead of phosphomimic protein constructs).

## CHAPTER 6

# CONCLUSIONS AND FUTURE RESEARCH

The research presented in this dissertation focused on investigating the molecular mechanisms by which the binding of α-actinin-4 (ACTN4) to actin filaments is differentially modulated by the following regulatory events: (i) tyrosine phosphorylation; (ii) binding of calcium ions; (iii) limited proteolysis by the m-calpain enzyme; and (iv) binding of phosphoinositide moieties. Using a combination of computational approaches from structural and systems modeling in conjunction with experimental validation, we were able to gain novel insights in important questions regarding cellular regulation actin cytoskeletal crosslinking by ACTN4, especially in response to the above external cues. Below, we briefly recap our main contributions as presented in the earlier chapters, along with possible directions for future studies.

In Chapter 2, we developed an all-atom structural model for the full wild-type ACTN4 homodimer, in which the multiple domains of the protein were brought together into a single structure using information from experimentally-solved structures of pairwise complexes along with homology models to several of the domains. We showed that a novel ternary complex involving the CH2/neck/CaM2 domains forms the core structure of the end region of the homodimer, and three sets of interfaces that contribute to stabilizing this ternary complex were experimentally validated. This full atomic model serves as a starting point for studying how various external cues affect the interactions between domains that lead to the regulation of the actin-binding function of ACTN4.

In Chapter 3, we showed that a novel tandem phosphorylation mechanism takes place during epidermal growth factor (EGF)-induced dual phosphorylation of Y4 and Y31 in the disordered N-terminal region of ACTN4, which leads to inhibition of actin binding. In particular, we found that only phosphorylated Y31 is necessary and sufficient to cause a reduction in actin binding, but the Y31 side chain is actually buried within the rest of the N-terminal while the Y4 side chain is solvent exposed, and thus accessible to kinase, in the unphosphorylated protein. It is only after prior phosphorylation of Y4 that the Y31 side chain gets exposed and becomes available for phosphorylation. A rule-based systems modeling approach showed that having a single kinase type or different kinase types act on the two sites both allow for an ultrasensitive response such that the phosphorylation of Y31 can be switched on-or-off by making changes in the relative activity of the kinase/s. A scan for potential kinases of the two sites using the PhosphoNET online suggests that membrane-bound receptor tyrosine kinases (RTKs) are the likely candidates. Assays by our collaborators are currently underway to test if the top-predicted kinase, AXL from the TAM family of RTKs, is responsible for Y4/Y31 phosphorylation. Finally, we show that the function of m-calpain cleavage of the N-terminal region is to remove the Y4 switch site such that Y31 phosphorylation is permanently disabled. This allows tandem phosphorylation to work in conjunction with m-calpain cleavage (which is localized to the rear end of motile cells) to allow for locale-specific functionality between the leading and rear ends of the cell during motility. Since phosphorylation occurs with greater frequency in the intrinsically disordered regions (IDRs) of proteins, an interesting topic for future research would be to search for examples of other proteins where two (or more) phosphorylation sites within and IDR involve tandem phosphorylation.

In Chapter 4, we used the all-atom model of full WT ACTN4 homodimer developed in Chapter 2 to study the molecular mechanisms that allow various external cues to regulate the actin-binding function of ACTN4. First, we found via MD simulations that phosphorylation of Y265 breaks contacts involving Y265 and H266 that should allow the ABD to adopt an open conformation (i.e., the complex between CH1 and CH2 domains is broken) such that CH1 cis free to bind actin. These simulations did not show the actual opening of the ABD, and a possible further extension here could be to use MD algorithms that enhance the sampling of configuration space, such as accelerated MD [147] or scaled MD [148], to see if CH1 and CH2 are indeed more easily separated upon phosphorylation of Y265. We next found, using targeted MD, that binding of calcium ions to CaM1 may break the CaM2/neck complex; the short simulations done here will need to be extended to gain further insight on how this leads to reduced actin binding. Next, we found that complexation with CaM2 protects a site in the neck region from m-calpain cleavage. Last, we used a virtual docking approach to dock the hydrophilic phosphoinositide head group to the CH2 domain; this docked model shows how ACTN4 can get localized to the surface of the inner membrane, and taken in combination with our full homodimer model shows how actin filaments can be crosslinked by ACTN4 into networks that are parallel to the surface of the membrane.

In Chapter 5, we have developed a preliminary network-level model that can serve as the starting point for building a computational tool that predicts the actin-binding response under a combination of multiple external cues. This model uses parameters derived from a two-step binding model of WT and pY265 ACTN4 as well as manually fitted parameters, and can recapitulate semi-quantitative experimental actin-binding data for WT, Y265 phosphomimic, and Y4/Y31 double phosphomimic conditions, either in the presence or absence of calcium ions. Directions for future research here include: (i) robustness testing and sensitivity analysis of the fitted parameters in the model; (ii) incorporation of additional conditions into the model, such as m-calpain cleavage and phosphoinositide binding, that require adding the notion of compartmentalization to the model; and (iii) translation of this model into a form that it is geared more towards the modeling of a cell-based environment than *in vitro* assays (e.g., model phosphorylation as a set of kinase and phosphatase reactions, rather than through the use of phosphomimic protein constructs).

# APPENDIX A

# PROTOCOL FOR MOLECULAR DYNAMICS SIMULATIONS

All molecular dynamics (MD) simulations were performed using GPU-accelerated Amber version 12.3.1 [149, 150] with the Amber ff12SB force field. All simulations were done with TIP3P solvation [151] in an octahedral water box. Counterions were added to neutralize the system charge. A cutoff of 8.0 Å was used for calculating short-range Lennard-Jones and electrostatic interactions, and long-range electrostatic contributions were computed using the particle-mesh Ewald approach [152]. A time step of 2 fs was used and all bond lengths involving hydrogen were kept constant using SHAKE (or SETTLE for water) [153, 154]. Initial energy minimization was performed using steepest descent for 1000 steps then conjugate gradient minimization for another 1000 steps. This was followed by NVT equilibration at 300 K for 50 ps and NPT equilibration at 1 atm and 300 K for another 50 ps. All heavy atoms had position restraints up to this point. Unrestrained NVT equilibration at 300K was then done for 50 ns, followed by the production runs under NVT at 300 K.

# BIBLIOGRAPHY

[1] Pollard TD and Cooper JA. (2009) Actin, a Central Player in Cell Shape and Movement. *Science* **326**:1208-1212.

[2] Schmidt A and Hall MN. (1998) Signaling to the actin cytoskeleton. *Annual Review of Cell and Developmental Biology* **14**:305-338.

[3] Geeves MA and Holmes KC. (1999) Structural mechanism of muscle contraction. *Annual Review of Biochemistry* **68**:687-728.

[4] Gardel ML, Shin JH, MacKintosh FC, Mahadevan L, Matsudaira P, and Weitz DA. (2004) Elastic behavior of cross-linked and bundled actin networks. *Science* **304**:1301-1305.

[5] Huttenlocher A and Horwitz AR. (2011) Integrins in cell migration. *Cold Spring Harbor Perspectives in Biology* **3**:a005074.

[6] Ciobanasu C, Faivre B, and Le Clainche C. (2012) Actin dynamics associated with focal adhesions. *International Journal of Cell Biology* **2012**:941292.

[7] Otto JJ. (1994) Actin-bundling proteins. *Current Opinion in Cell Biology* **6**:105-109.

[8] Otey CA and Carpen O. (2004) Alpha-actinin revisited: a fresh look at an old player. *Cell Motility and the Cytoskeleton* **58**:104-111.

[9] Sjöblom B, Salmazo A, and Djinović-Carugo K. (2008) Alpha-actinin structure and regulation. *Cellular and Molecular Life Sciences* **65**:2688-2701.

[10] Davison MD and Critchley DR. (1988) alpha-Actinins and the DMD protein contain spectrin-like repeats. *Cell* **52**:159-160.

[11] Blanchard A, Ohanian V, and Critchley D. (1989) The structure and function of alpha-actinin. *Journal of Muscle Research & Cell Motility* **10**:280-289.

[12] Ylänne J, Scheffzek K, Young P, and Saraste M. (2001) Crystal structure of the alpha-actinin rod reveals an extensive torsional twist. *Structure* **9**:597-604.

[13] Flood G, Kahana E, Gilmore AP, Rowe AJ, Gratzer WB, and Critchley DR. (1995) Association of structural repeats in the alpha-actinin rod domain. Alignment of inter-subunit interactions. *Journal of Molecular Biology* **252**:227-234.

[14] Flood G, Rowe AJ, Critchley DR, and Gratzer WB. (1997) Further analysis of the role of spectrin repeat motifs in alpha-actinin dimer formation. *European Biophysics Journal* **25**:431-435.

[15] The UniProt Consortium. (2014) Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Research* **42**:D191-D198.

[16] Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, and Higgins DG. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7**:539.

[17] Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, and Lopez R. (2010) A new bioinformatics analysis tools framework at EMBL–EBI. *Nucleic Acids Research* **38**:W695-W699.

[18] Galkin VE, Orlova A, Salmazo A, Djinovic-Carugo K, and Egelman EH. (2010) Opening of tandem calponin homology domains regulates their affinity for F-actin. *Nature Structural & Molecular Biology* **17**:614-616.

[19] Witke W, Hofmann A, Köppel B, Schleicher M, and Noegel AA. (1993) The Ca(2+)-binding domains in non-muscle type alpha-actinin: biochemical and genetic analysis. *The Journal of Cell Biology* **121**:599-606.

[20] Imamura M, Sakurai T, Ogawa Y, Ishikawa T, Goto K, and Masaki T. (1994) Molecular cloning of low-Ca(2+)-sensitive-type non-muscle alpha-actinin. *European Journal of Biochemistry* **223**:395-401.

[21] Shao H, Wu C, and Wells A. (2010) Phosphorylation of alpha-actinin 4 upon epidermal growth factor exposure regulates its interaction with actin. *The Journal of Biological Chemistry* **285**:2591-2600.

[22] Izaguirre G, Aguirre L, Hu YP, Lee HY, Schlaepfer DD, Aneskievich BJ, and Haimovich B. (2001) The cytoskeletal/non-muscle isoform of alpha-actinin is phosphorylated on its actin-binding domain by the focal adhesion kinase. *The Journal of Biological Chemistry* **276**:28676-28685.

[23] Rush J, Moritz A, Lee KA, Guo A, Goss VL, Spek EJ, Zhang H, Zha XM, Polakiewicz RD, and Comb MJ. (2005) Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nature Biotechnology* **23**:94-101.

[24] Rikova K, Guo A, Zeng Q, Possemato A, Yu J, Haack H, Nardone J, Lee K, Reeves C, Li Y, Hu Y, Tan Z, Stokes M, Sullivan L, Mitchell J, Wetzel R, Macneill J, Ren JM, Yuan J, Bakalarski CE, Villen J, Kornhauser JM, Smith B, Li D, Zhou X, Gygi SP, Gu TL, Polakiewicz RD, Rush J, and Comb MJ. (2007) Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell* **131**:1190-1203.

[25] Shao H, Wang JH, Pollak MR, and Wells A. (2010) Alpha-actinin-4 is essential for maintaining the spreading, motility and contractility of fibroblasts. *PLoS One* **5**:e13921.

[26] Fraley TS, Pereira CB, Tran TC, Singleton C, and Greenwood JA. (2005) Phosphoinositide binding regulates alpha-actinin dynamics: mechanism for modulating cytoskeletal remodeling. *The Journal of Biological Chemistry* **280**:15479-15482.

[27] Dandapani SV, Sugimoto H, Matthews BD, Kolb RJ, Sinha S, Gerszten RE, Zhou J, Ingber DE, Kalluri R, and Pollak MR. (2007) Alpha-actinin-4 is required for normal podocyte adhesion. *The Journal of Biological Chemistry* **282**:467-477.

[28] Kanchanawong P, Shtengel G, Pasapera AM, Ramko EB, Davidson MW, Hess HF, and Waterman CM. (2010) Nanoscale architecture of integrin-based cell adhesions. *Nature* **468**:580-584.

[29] Roca-Cusachs P, del Rio A, Puklin-Faucher E, Gauthier NC, Biais N, and Sheetz MP. (2013) Integrin-dependent force transmission to the extracellular matrix by alpha-actinin triggers adhesion maturation. *Proceedings of the National Academy of Sciences* **110**:E1361-E1370.

[30] Honda K, Yamada T, Endo R, Ino Y, Gotoh M, Tsuda H, Yamada Y, Chiba H, and Hirohashi S. (1998) Actinin-4, a novel actin-bundling protein associated with cell motility and cancer invasion. *The Journal of Cell Biology* **140**:1383-1393.

[31] Honda K, Yamada T, Hayashida Y, Idogawa M, Sato S, Hasegawa F, Ino Y, Ono M, and Hirohashi S. (2005) Actinin-4 increases cell motility and promotes lymph node metastasis of colorectal cancer. *Gastroenterology* **128**:51-62.

[32] Kikuchi S, Honda K, Tsuda H, Hiraoka N, Imoto I, Kosuge T, Umaki T, Onozato K, Shitashige M, Yamaguchi U, Ono M, Tsuchida A, Aoki T, Inazawa J, Hirohashi S, and Yamada T. (2008) Expression and gene amplification of actinin-4 in invasive ductal carcinoma of the pancreas. *Clinical Cancer Research* **14**:5348-5356.

[33] Yamamoto S, Tsuda H, Honda K, Onozato K, Takano M, Tamai S, Imoto I, Inazawa J, Yamada T, and Matsubara O. (2009) Actinin-4 gene amplification in ovarian cancer: a candidate oncogene associated with poor patient prognosis and tumor chemoresistance. *Modern Pathology* **22**:499-507.

[34] Sen S, Dong M, and Kumar S. (2009) Isoform-specific contributions of alpha-actinin to glioma cell mechanobiology. *PLoS One* **4**:e8427.

[35] Quick Q and Skalli O. (2010) Alpha-actinin 1 and alpha-actinin 4: contrasting roles in the survival, motility, and RhoA signaling of astrocytoma cells. *Experimental Cell Research* **316**:1137-1147.

[36] Shao H, Travers T, Camacho CJ, and Wells A. (2013) The carboxyl tail of alpha-actinin-4 regulates its susceptibility to m-calpain and thus functions in cell migration and spreading. *The International Journal of Biochemistry & Cell Biology* **45**:1051-1063.

[37] Selliah N, Brooks WH, and Roszman TL. (1996) Proteolytic cleavage of alpha-actinin by calpain in T cells stimulated with anti-CD3 monoclonal antibody. *The Journal of Immunology* **156**:3215-3221.

[38] Sprague CR, Fraley TS, Jang HS, Lal S, and Greenwood JA. (2008) Phosphoinositide binding to the substrate regulates susceptibility to proteolysis by calpain. *The Journal of Biological Chemistry* **283**:9217-9223.

[39] Fukami K, Sawada N, Endo T, and Takenawa T. (1996) Identification of a phosphatidylinositol 4,5-bisphosphate-binding site in chicken skeletal muscle alpha-actinin. *The Journal of Biological Chemistry* **271**:2646-2650.

[40] Fukami K, Furuhashi K, Inagaki M, Endo T, Hatano S, and Takenawa T. (1992) Requirement of phosphatidylinositol 4,5-bisphosphate for alpha-actinin function. *Nature* **359**:150-152.

[41] Fraley TS, Tran TC, Corgan AM, Nash CA, Hao J, Critchley DR, and Greenwood JA. (2003) Phosphoinositide binding inhibits alpha-actinin bundling activity. *The Journal of Biological Chemistry* **278**:24039-24045.

[42] Dror RO, Dirks RM, Grossman JP, Xu H, and Shaw DE. (2012) Biomolecular simulation: a computational microscope for molecular biology. *Annual Review of Biophysics* **41**:429-452.

[43] Friedrichs MS, Eastman P, Vaidyanathan V, Houston M, Legrand S, Beberg AL, Ensign DL, Bruns CM, and Pande VS. (2009) Accelerating molecular dynamic simulation on graphics processing units. *Journal of Computational Chemistry* **30**:864-872.

[44] Le Grand S, Götz AW, and Walker RC. (2013) SPFP: Speed without compromise—A mixed precision model for GPU accelerated molecular dynamics simulations. *Computer Physics Communications* **184**:374-380.

[45] Zuckerman DM. (2011) Equilibrium sampling in biomolecular simulations. *Annual Review of Biophysics* **40**:41-62.

[46] Ponder JW, Wu C, Ren P, Pande VS, Chodera JD, Schnieders MJ, Haque I, Mobley DL, Lambrecht DS, DiStasio RA, Jr., Head-Gordon M, Clark GN, Johnson ME, and Head-

Gordon T. (2010) Current status of the AMOEBA polarizable force field. *The Journal of Physical Chemistry B* **114**:2549-2564.

[47] Jiang W, Hardy DJ, Phillips JC, Mackerell AD, Jr., Schulten K, and Roux B. (2011) High-performance scalable molecular dynamics simulations of a polarizable force field based on classical Drude oscillators in NAMD. *The Journal of Physical Chemistry Letters* **2**:87-92.

[48] Sousa SF, Fernandes PA, and Ramos MJ. (2006) Protein-ligand docking: current status and future challenges. *Proteins* **65**:15-26.

[49] Pons C, Grosdidier S, Solernou A, Perez-Cano L, and Fernandez-Recio J. (2010) Present and future challenges and limitations in protein-protein docking. *Proteins* **78**:95-108.

[50] Gabb HA, Jackson RM, and Sternberg MJE. (1997) Modelling protein docking using shape complementarity, electrostatics and biochemical information. *Journal of Molecular Biology* **272**:106-120.

[51] Comeau SR, Gatchell DW, Vajda S, and Camacho CJ. (2004) ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics* **20**:45-50.

[52] Koes DR, Baumgartner MP, and Camacho CJ. (2013) Lessons learned in empirical scoring with smina from the CSAR 2011 benchmarking exercise. *Journal of Chemical Information and Modeling* **53**:1893-1904.

[53] Jain AN. (2006) Scoring functions for protein-ligand docking. *Current Protein & Peptide Science* **7**:407-420.

[54] Ekman D, Björklund ÅK, Frey-Skött J, and Elofsson A. (2005) Multi-domain proteins in the three kingdoms of life: orphan domains and other unassigned regions. *Journal of Molecular Biology* **348**:231-243.

[55] Borrego-Diaz E, Kerff F, Lee SH, Ferron F, Li Y, and Dominguez R. (2006) Crystal structure of the actin-binding domain of alpha-actinin 1: evaluating two competing actin-binding models. *Journal of Structural Biology* **155**:230-238.

[56] Franzot G, Sjöblom B, Gautel M, and Djinović Carugo K. (2005) The crystal structure of the actin binding domain from alpha-actinin in its closed conformation: structural insight into phospholipid regulation of alpha-actinin. *Journal of Molecular Biology* **348**:151-165.

[57] Lee SH, Weins A, Hayes DB, Pollak MR, and Dominguez R. (2008) Crystal structure of the actin-binding domain of alpha-actinin-4 Lys255Glu mutant implicated in focal segmental glomerulosclerosis. *Journal of Molecular Biology* **376**:317-324.

[58] Korenbaum E and Rivero F. (2002) Calponin homology domains at a glance. *Journal of Cell Science* **115**:3543-3545.

[59] Gimona M, Djinovic-Carugo K, Kranewitter WJ, and Winder SJ. (2002) Functional plasticity of CH domains. *FEBS Letters* **513**:98-106.

[60] Bresnick AR, Janmey PA, and Condeelis J. (1991) Evidence that a 27-residue sequence is the actin-binding site of ABP-120. *The Journal of Biological Chemistry* **266**:12989-12993.

[61] Kuhlman PA, Hemmings L, and Critchley DR. (1992) The identification and characterisation of an actin-binding site in alpha-actinin by mutagenesis. *FEBS Letters* **304**:201-206.

[62] Levine BA, Moir AJ, Patchell VB, and Perry SV. (1990) The interaction of actin with dystrophin. *FEBS Letters* **263**:159-162.

[63] Levine BA, Moir AJ, Patchell VB, and Perry SV. (1992) Binding sites involved in the interaction of actin with the N-terminal region of dystrophin. *FEBS Letters* **298**:44-48.

[64] Dominguez R. (2004) Actin-binding proteins--a unifying hypothesis. *Trends in Biochemical Sciences* **29**:572-578.

[65] Young P and Gautel M. (2000) The interaction of titin and alpha-actinin is controlled by a phospholipid-regulated intramolecular pseudoligand mechanism. *The EMBO Journal* **19**:6331-6340.

[66] Liu J, Taylor DW, and Taylor KA. (2004) A 3-D reconstruction of smooth muscle alpha-actinin by CryoEm reveals two different conformations at the actin-binding region. *Journal of Molecular Biology* **338**:115-125.

[67] Travers T, Shao H, Wells A, and Camacho CJ. (2013) Modeling the assembly of the multiple domains of alpha-actinin-4 and its role in actin cross-linking. *Biophysical Journal* **104**:705-715.

[68] Ohtsuka H, Yajima H, Maruyama K, and Kimura S. (1997) Binding of the N-terminal 63 kDa portion of connectin/titin to alpha-actinin as revealed by the yeast two-hybrid system. *FEBS Letters* **401**:656-657.

[69] Ohtsuka H, Yajima H, Maruyama K, and Kimura S. (1997) The N-terminal Z repeat 5 of connectin/titin binds to the C-terminal region of alpha-actinin. *Biochemical and Biophysical Research Communications* **235**:1-3.

[70] Sorimachi H, Freiburg A, Kolmerer B, Ishiura S, Stier G, Gregorio CC, Labeit D, Linke WA, Suzuki K, and Labeit S. (1997) Tissue-specific expression and alpha-actinin binding

properties of the Z-disc titin: implications for the nature of vertebrate Z-discs. *Journal of Molecular Biology* **270**:688-695.

[71] Young P, Ferguson C, Banuelos S, and Gautel M. (1998) Molecular structure of the sarcomeric Z-disk: two types of titin interactions lead to an asymmetrical sorting of alpha-actinin. *The EMBO Journal* **17**:1614-1624.

[72] Trinick J. (1996) Titin as a scaffold and spring. Cytoskeleton. *Current Biology* **6**:258-260.

[73] Atkinson RA, Joseph C, Dal Piaz F, Birolo L, Stier G, Pucci P, and Pastore A. (2000) Binding of alpha-actinin to titin: implications for Z-disk assembly. *Biochemistry* **39**:5255-5264.

[74] Joseph C, Stier G, O'Brien R, Politou AS, Atkinson RA, Bianco A, Ladbury JE, Martin SR, and Pastore A. (2001) A structural characterization of the interactions between titin Z-repeats and the alpha-actinin C-terminal domain. *Biochemistry* **40**:4957-4965.

[75] Atkinson RA, Joseph C, Kelly G, Muskett FW, Frenkiel TA, Nietlispach D, and Pastore A. (2001) Ca2+-independent binding of an EF-hand domain to a novel motif in the alpha-actinin-titin complex. *Nature Structural Biology* **8**:853-857.

[76] Rice P, Longden I, and Bleasby A. (2000) EMBOSS: the European Molecular Biology Open Software Suite. *Trends in Genetics* **16**:276-277.

[77] Eisenberg D, Schwarz E, Komaromy M, and Wall R. (1984) Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *Journal of Molecular Biology* **179**:125-142.

[78] The PyMOL Molecular Graphics System, Version 1.6.9.0. Schrödinger, LLC.

[79] Champ PC and Camacho CJ. (2007) FastContact: a free energy scoring tool for protein-protein complex structures. *Nucleic Acids Research* **35**:W556-W560.

[80] Privalov PL. (1990) Cold denaturation of proteins. *Critical Reviews in Biochemistry and Molecular Biology* **25**:281-305.

[81] Humphrey W, Dalke A, and Schulten K. (1996) VMD: visual molecular dynamics. *Journal of Molecular Graphics* **14**:33-8, 27-28.

[82] Kuboniwa H, Tjandra N, Grzesiek S, Ren H, Klee CB, and Bax A. (1995) Solution structure of calcium-free calmodulin. *Nature Structural Biology* **2**:768-776.

[83] Zhou H and Murthy KS. (2003) Identification of the G protein-activating sequence of the single-transmembrane natriuretic peptide receptor C (NPR-C). *American Journal of Physiology. Cell Physiology* **284**:C1255-C1261.

[84] Goldsmith SC, Pokala N, Shen W, Fedorov AA, Matsudaira P, and Almo SC. (1997) The structure of an actin-crosslinking domain from human fimbrin. *Nature Structural Biology* **4**:708-712.

[85] Sevcik J, Urbanikova L, Kost'an J, Janda L, and Wiche G. (2004) Actin-binding domain of mouse plectin. Crystal structure and binding to vimentin. *European Journal of Biochemistry* **271**:1873-1884.

[86] Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, and Dunker AK. (2004) The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Research* **32**:1037-1049.

[87] Tyanova S, Cox J, Olsen J, Mann M, and Frishman D. (2013) Phosphorylation variation during the cell cycle scales with structural propensities of proteins. *PLoS Computational Biology* **9**:e1002842.

[88] Liu J, Faeder JR, and Camacho CJ. (2009) Toward a quantitative theory of intrinsically disordered proteins and their function. *Proceedings of the National Academy of Sciences* **106**:19819-19823.

[89] Iakoucheva LM, Brown CJ, Lawson JD, Obradovic Z, and Dunker AK. (2002) Intrinsic disorder in cell-signaling and cancer-associated proteins. *Journal of Molecular Biology* **323**:573-584.

[90] Espinoza-Fonseca LM, Kast D, and Thomas DD. (2007) Molecular dynamics simulations reveal a disorder-to-order transition on phosphorylation of smooth muscle myosin. *Biophysical Journal* **93**:2083-2090.

[91] Espinoza-Fonseca LM, Kast D, and Thomas DD. (2008) Thermodynamic and structural basis of phosphorylation-induced disorder-to-order transition in the regulatory light chain of smooth muscle myosin. *Journal of the American Chemical Society* **130**:12208-12209.

[92] Chen W, Dang T, Blind RD, Wang Z, Cavasotto CN, Hittelman AB, Rogatsky I, Logan SK, and Garabedian MJ. (2008) Glucocorticoid receptor phosphorylation differentially affects target gene expression. *Molecular Endocrinology* **22**:1754-1766.

[93] Metcalfe EE, Traaseth NJ, and Veglia G. (2005) Serine 16 phosphorylation induces an order-to-disorder transition in monomeric phospholamban. *Biochemistry* **44**:4386-4396.

[94] Sugita Y, Miyashita N, Yoda T, Ikeguchi M, and Toyoshima C. (2006) Structural changes in the cytoplasmic domain of phospholamban by phosphorylation at Ser16: a molecular dynamics study. *Biochemistry* **45**:11752-11761.

[95] Bertolucci CM, Guibao CD, and Zheng JJ. (2008) Phosphorylation of paxillin LD4 destabilizes helix formation and inhibits binding to focal adhesion kinase. *Biochemistry* **47**:548-554.

[96] Hegedus T, Serohijos AW, Dokholyan NV, He L, and Riordan JR. (2008) Computational studies reveal phosphorylation-dependent changes in the unstructured R domain of CFTR. *Journal of Molecular Biology* **378**:1052-1063.

[97] Lowry DF, Stancik A, Shrestha RM, and Daughdrill GW. (2008) Modeling the accessible conformations of the intrinsically unstructured transactivation domain of p53. *Proteins* **71**:587-598.

[98] Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, and Gascuel O. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Research* **36**:W465-W469.

[99] Edgar RC. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**:1792-1797.

[100] Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, and Gascuel O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* **59**:307-321.

[101] Chevenet F, Brun C, Banuls AL, Jacq B, and Christen R. (2006) TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics* **7**:439.

[102] Das RK and Pappu RV. (2013) Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proceedings of the National Academy of Sciences* **110**:13392-13397.

[103] Uversky VN, Santambrogio C, Brocca S, and Grandori R. (2012) Length-dependent compaction of intrinsically disordered proteins. *FEBS Letters* **586**:70-73.

[104] Rowland MA, Fontana W, and Deeds EJ. (2012) Crosstalk and competition in signaling networks. *Biophysical Journal* **103**:2389-2398.

[105] Goldbeter A and Koshland DE, Jr. (1981) An amplified sensitivity arising from covalent modification in biological systems. *Proceedings of the National Academy of Sciences* **78**:6840-6844.

[106] Comeau SR, Gatchell DW, Vajda S, and Camacho CJ. (2004) ClusPro: a fully automated algorithm for protein-protein docking. *Nucleic Acids Research* **32**:W96-W99.

[107] Hlavacek WS, Faeder JR, Blinov ML, Posner RG, Hucka M, and Fontana W. (2006) Rules for modeling signal-transduction systems. *Science STKE* **2006**:re6.

[108] Faeder JR, Blinov ML, and Hlavacek WS. (2009) Rule-based modeling of biochemical systems with BioNetGen. *Methods in Molecular Biology* **500**:113-167.

[109] Safaei J, Manuch J, Gupta A, Stacho L, and Pelech S. (2011) Prediction of 492 human protein kinase substrate specificities. *Proteome Science* **9 Suppl 1**:S6.

[110] Gonzalez AM, Otey C, Edlund M, and Jones JC. (2001) Interactions of a hemidesmosome component and actinin family members. *Journal of Cell Science* **114**:4197-4206.

[111] Michaud JL, Chaisson KM, Parks RJ, and Kennedy CR. (2006) FSGS-associated alpha-actinin-4 (K256E) impairs cytoskeletal dynamics in podocytes. *Kidney International* **70**:1054-1061.

[112] Gusenbauer S, Vlaicu P, and Ullrich A. (2013) HGF induces novel EGFR functions involved in resistance formation to tyrosine kinase inhibitors. *Oncogene* **32**:3846-3856.

[113] Croyle M, Akeno N, Knauf JA, Fabbro D, Chen X, Baumgartner JE, Lane HA, and Fagin JA. (2008) RET/PTC-induced cell growth is mediated in part by epidermal growth factor receptor (EGFR) activation: evidence for molecular and functional interactions between RET and EGFR. *Cancer Research* **68**:4183-4191.

[114] Linger RM, Keating AK, Earp HS, and Graham DK. (2008) TAM receptor tyrosine kinases: biologic functions, signaling, and potential therapeutic targeting in human cancer. *Advances in Cancer Research* **100**:35-83.

[115] Shao H, Chou J, Baty CJ, Burke NA, Watkins SC, Stolz DB, and Wells A. (2006) Spatial localization of m-calpain to the plasma membrane by phosphoinositide biphosphate binding during epidermal growth factor receptor-mediated activation. *Molecular and Cellular Biology* **26**:5481-5496.

[116] Leloup L, Shao H, Bae YH, Deasy B, Stolz D, Roy P, and Wells A. (2010) m-Calpain activation is regulated by its membrane localization and by its binding to phosphatidylinositol 4,5-bisphosphate. *The Journal of Biological Chemistry* **285**:33549-33566.

[117] Franco SJ, Rodgers MA, Perrin BJ, Han J, Bennin DA, Critchley DR, and Huttenlocher A. (2004) Calpain-mediated proteolysis of talin regulates adhesion dynamics. *Nature Cell Biology* **6**:977-983.

[118] Bailly M, Wyckoff J, Bouzahzah B, Hammerman R, Sylvestre V, Cammer M, Pestell R, and Segall JE. (2000) Epidermal growth factor receptor distribution during chemotactic responses. *Molecular Biology of the Cell* **11**:3873-3883.

[119] Wilson MA and Brunger AT. (2000) The 1.0 A crystal structure of Ca(2+)-bound calmodulin: an analysis of disorder and implications for functionally relevant plasticity. *Journal of Molecular Biology* **301**:1237-1256.

[120] Schlitter J, Engels M, Krüger P, Jacoby E, and Wollmer A. (1993) Targeted Molecular Dynamics Simulation of Conformational Change-Application to the T ↔ R Transition in Insulin. *Molecular Simulation* **10**:291-308.

[121] Schlitter J, Engels M, and Kruger P. (1994) Targeted molecular dynamics: a new approach for searching pathways of conformational transitions. *Journal of Molecular Graphics* **12**:84-89.

[122] van der Vaart A and Karplus M. (2005) Simulation of conformational transitions by the restricted perturbation-targeted molecular dynamics method. *The Journal of Chemical Physics* **122**:114903.

[123] Apostolakis J, Ferrara P, and Caflisch A. (1999) Calculation of conformational transitions and barriers in solvated systems: Application to the alanine dipeptide in water. *The Journal of Chemical Physics* **110**:2099-2108.

[124] Huang H, Ozkirimli E, and Post CB. (2009) A Comparison of Three Perturbation Molecular Dynamics Methods for Modeling Conformational Transitions. *Journal of Chemical Theory and Computation* **5**:1301-1314.

[125] Ovchinnikov V and Karplus M. (2012) Analysis and elimination of a bias in targeted molecular dynamics simulations of conformational transitions: application to calmodulin. *The Journal of Physical Chemistry B* **116**:8584-8603.

[126] Goll DE, Thompson VF, Li H, Wei W, and Cong J. (2003) The calpain system. *Physiological Reviews* **83**:731-801.

[127] Hosfield CM, Elce JS, Davies PL, and Jia Z. (1999) Crystal structure of calpain reveals the structural basis for Ca(2+)-dependent protease activity and a novel mode of enzyme activation. *The EMBO Journal* **18**:6880-6889.

[128] Strobl S, Fernandez-Catalan C, Braun M, Huber R, Masumoto H, Nakagawa K, Irie A, Sorimachi H, Bourenkow G, Bartunik H, Suzuki K, and Bode W. (2000) The crystal structure of calcium-free human m-calpain suggests an electrostatic switch mechanism for activation by calcium. *Proceedings of the National Academy of Sciences* **97**:588-592.

[129] Pal GP, De Veyra T, Elce JS, and Jia Z. (2003) Crystal structure of a micro-like calpain reveals a partially activated conformation with low Ca2+ requirement. *Structure* **11**:1521-1526.

[130] Franco SJ and Huttenlocher A. (2005) Regulating cell migration: calpains make the cut. *Journal of Cell Science* **118**:3829-3838.

[131] Hanna RA, Campbell RL, and Davies PL. (2008) Calcium-bound structure of calpain and its mechanism of inhibition by calpastatin. *Nature* **456**:409-412.

[132] Moldoveanu T, Gehring K, and Green DR. (2008) Concerted multi-pronged attack by calpastatin to occlude the catalytic cleft of heterodimeric calpains. *Nature* **456**:404-408.

[133] Fukami K, Endo T, Imamura M, and Takenawa T. (1994) alpha-Actinin and vinculin are PIP2-binding proteins involved in signaling by tyrosine kinase. *The Journal of Biological Chemistry* **269**:1518-1522.

[134] Yin HL and Janmey PA. (2003) Phosphoinositide regulation of the actin cytoskeleton. *Annual Review of Physiology* **65**:761-789.

[135] Trott O and Olson AJ. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry* **31**:455-461.

[136] Zoete V, Cuendet MA, Grosdidier A, and Michielin O. (2011) SwissParam: a fast force field generation tool for small organic molecules. *Journal of Computational Chemistry* **32**:2359-2568.

[137] Halgren TA. (1996) Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry* **17**:490-519.

[138] Halgren TA. (1996) Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. *Journal of Computational Chemistry* **17**:520-552.

[139] Halgren TA. (1996) Merck molecular force field. III. Molecular geometries and vibrational frequencies for MMFF94. *Journal of Computational Chemistry* **17**:553-586.

[140] Halgren TA and Nachbar RB. (1996) Merck molecular force field. IV. conformational energies and geometries for MMFF94. *Journal of Computational Chemistry* **17**:587-615.

[141] Halgren TA. (1996) Merck molecular force field. V. Extension of MMFF94 using experimental data, additional computational data, and empirical rules. *Journal of Computational Chemistry* **17**:616-641.

[142] MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kuczera J, Yin D, and Karplus M. (1998) All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins†. *The Journal of Physical Chemistry B* **102**:3586-3616.

[143] Shewan A, Eastburn DJ, and Mostov K. (2011) Phosphoinositides in cell architecture. *Cold Spring Harbor Perspectives in Biology* **3**:a004796.

[144] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, and Ferrin TE. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25**:1605-1612.

[145] Weins A, Schlondorff JS, Nakamura F, Denker BM, Hartwig JH, Stossel TP, and Pollak MR. (2007) Disease-associated mutant alpha-actinin-4 reveals a mechanism for regulating its F-actin-binding affinity. *Proceedings of the National Academy of Sciences* **104**:16080-16085.

[146] Fujii T, Iwane AH, Yanagida T, and Namba K. (2010) Direct visualization of secondary structures of F-actin by electron cryomicroscopy. *Nature* **467**:724-728.

[147] Hamelberg D, Mongan J, and McCammon JA. (2004) Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *The Journal of Chemical Physics* **120**:11919-11929.

[148] Sinko W, Miao Y, de Oliveira CA, and McCammon JA. (2013) Population based reweighting of scaled molecular dynamics. *The Journal of Physical Chemistry B* **117**:12759-12768.

[149] Case DA, Darden TA, Cheatham TEI, Simmerling CL, Wang J, Duke RE, Luo R, Walker RC, Zhang W, Merz KM, Roberts B, Hayik S, Roitberg A, Seabra G, Swails J, Götz AW, Kolossváry I, Wong KF, Paesani F, Vanicek J, Wolf RM, Liu J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Cai Q, Ye X, Wang J, Hsieh M-J, Cui G, Roe DR, Mathews DH, Seetin MG, Salomon-Ferrer R, Sagui C, Babin V, Luchko T, Gusarov S, Kovalenko A, and Kollman PA. (2012) AMBER 12. University of California, San Francisco.

[150] Salomon-Ferrer R, Götz AW, Poole D, Le Grand S, and Walker RC. (2013) Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *Journal of Chemical Theory and Computation* **9**:3878-3888.

[151] Jorgensen WL. (1981) Quantum and statistical mechanical studies of liquids. 10. Transferable intermolecular potential functions for water, alcohols, and ethers. Application to liquid water. *Journal of the American Chemical Society* **103**:335-340.

[152] Darden T, York D, and Pedersen L. (1993) Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **98**:10089-10092.

[153] Ryckaert J-P, Ciccotti G, and Berendsen H. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* **23**:327-341.

[154] Miyamoto S and Kollman PA. (1992) SETTLE: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry* **13**:952-962.