UCSF
UC San Francisco Electronic Theses and Dissertations

Title
Conformational Switches Regulate Clathrin Mediated Endocytosis

Permalink
https://escholarship.org/uc/item/0fd1v310

Author
Wilbur, Jeremy D.

Publication Date
2008-08-15

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Conformational Switches Regulate Clathrin Mediated Endocytosis

by

Jeremy D. Wilbur

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biophysics

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Acknowledgements

I would like to thank all those that have contributed to my scientific career. From my first days in the laboratory of Dr. Juliet Fuhrman I have been blessed with fun co-workers and an intellectually stimulating environments. During my time in the laboratory of Dr. Xandra Breakefield I was given tremendous freedoms to explore scientific questions that peaked my interest. Upon entering my dissertation laboratories under the guidance of Dr. Robert Fletterick and Dr. Frances Brodsky, the freedom to seek my own scientific questions was very important to me. Within these labs I found my own path with the support of my mentors. To them I am greatful for what they have taught me both scientifically and personally.

To my friends and colleagues within all these labs, I thank you as well. Without you the richness of my experience would have been greatly tempered.

Below is a list of specific people who have contributed to publications that will result from this dissertation work.

Involvement of Coauthors – Scientific Contributions

The work presented here is the result of efforts by many collaborating scientists. Though primarily my work many people contributed to collecting and interpreting data. Without these contributions this work would not have reached equivalent quality and impact. This

work has resulted in three publications which are all currently in the review process. Below the scientific contribution by others is defined. I thank those involved.

Chapter 1 – Conformation Switching of Clathrin Light Chain Regulates Clathrin lattice assembly. In reviewed favorably by Nature.

This long term project was initiated by Joel Ybe and Peter Hwang and Frances Brodsky. Each person contributed significant scientific guidance. Joel Ybe and Peter Hwang grew the crystals and collected the crystal data. Robert Fletterick provide significant guidance relating to X-ray crystallography and protein structure and function analysis. Development of the all atom model used in crystallography analysis was collaboration between Ben Sellers, Matt Jacobson and myself and proved critical as described in chapter 4. Michael Lane collected much of the biochemical data and tirelessly purified proteins.

Chapter 2 – Clathrin light chain regulates actin binding by hip1 and hip1R. Reviewed favorably by the Journal of Biological Chemistry.

This work was primarily of my own design with significant scientific guidance from Frances Brodsky and Peter Hwang. Robert Fletterick contributed additional scientific guidance. Venus Manalo and Jean Chen contributed important cellular data without which the biological context of this work would be unclear.

Appendix 1 - The Disordered C-terminus of RcdA Is Necessary for Efficient CtrA Proteolysis in Caulobacter crescentus. Manuscript in progress. To be submitted for publication.

This work was in collaboration with James Taylor and Kathleen Ryan of University California Berkeley. My contribution to this work was to determine the structure of RcdA. This resulted in the determination of a novel fold and the identification of disordered regions of the protein. James Taylor provided the cellular context for structural ideas. Kathleen Ryan provided scientific guidance in biochemical and biological functions of RcdA.

Abstract
Conformational Switches Regulate Clathrin Mediated Endocytosis

Jeremy D. Wilbur

Clathrin mediated endocytosis is a fundamental cellular process used to regulate the response of cells to their environment. It has been usurped as a means for pathogens to enter cells. Clathrin mediated endocytosis is driven by the polymerization of clathrin into a polyhedral lattice. Lattice formation is critical for stabilization of membranes and organization of the protein machinery used to form a budding vesicle. Clathrin has a unique three legged structure. The legs interweave to form the lattice. Regulation of the process of lattice assembly has been observationally described for over twenty years. The mechanisms of regulation are only now being defined.

The clathrin endocytic machinery is closely associated with the actin cytoskeleton. This may allow the force of actin polymerization to be harnessed for coated vesicle formation or actin may play an additional stabilizing role during endocytosis. The interaction between clathrin and actin machineries is becoming an increasingly important research area. The details of the link between clathrin and actin are not yet determined.

To help define the mechanisms involved in forming clathrin coated vesicles and linking clathrin to the actin cytoskeleton biochemical and structural studies were undertaken. The first structure of the regulatory clathrin light chain suggested a conformational switch regulated clathrin assembly. Structural and biochemical work on

the huntingtin interacting proteins again suggested a conformational switch regulates the link between clathrin and actin.

This work suggests a general mechanism of numerous conformational switches being flipped to regulated clathrin mediated endocytosis.

Table of Contents

List of Tables

List of Figures

Introduction

Chapter 1

Chapter 2

Conformational Switches Regulate Clathrin Mediated Endocytosis

Introduction:

Cellular sensing of external stimuli requires a vast array of signaling receptors at the cell surface (Brodsky, Chen et al. 2001). Cellular organelles maintain structure and function through the flux of vesicles merging and separating from the delimiting membranes. Viral and bacterial pathogens enter their cellular host to escape immune detection and to proliferate (Bonazzi and Cossart 2006; Marsh and Helenius 2006). All these processes occur through the action of the clathrin machinery.

The hallmark of eukaryotic cells is the compartmentalization of cellular functions achieved with membrane bilayers. Compartmentalization allows the occurrence of increased diversity of cellular functions by separation of normally contradictory functions. Unfortunately, while the benefit of functional diversity is achieved, the exchange of proteins and metabolites is hindered due to the impermeability of the membranes. This leads to distinct topologically similar compartments within a cell that provide varied functions (Fig. 1). One way to overcome compartment barriers and exchange cellular molecules is through the transport of a membrane bound vesicles that forms from one organelle and fuses with another (Fig. 1). While this does not allow proteins or metabolites to cross the membrane this does allow relatively free movement of materials in topologically similar compartments.

The necessity of forming a vesicle from a surrounding membrane is the common feature within endocytosis, viral and bacterial pathogenesis, and organelle biogenesis. The most well studied means of achieving vesicle formation is to surround membranes in protein coats or large lattices formed from polymerization of proteins (McMahon and Mills 2004; Stagg, LaPointe et al. 2007). Protein coats provide a scaffold for organizing the numerous coat associated factors into a functioning machinery and likely stabilize energetically unfavorable highly curved membrane conformations (Brodsky, Chen et al. 2001). Additionally, protein coats may help provide mechanical force necessary to bend membranes into small vesicles (Hinrichsen, Meyerholz et al. 2006).

Three major types of coat proteins are known. Clathrin was the first discovered and plays major roles at the plasma membrane and trans-golgi. The other two coat proteins, COPI and COPII, specifically form vesicles within ER and golgi. Clathrin mediated vesicle formation continues to be the prototype for studying coat assembly mechanisms and the dynamic interplay of protein coats with accessory proteins.

Clathrin mediated vesicle formation proceeds through a hierarchical set of processes starting with a signaling event triggering the initiation of coat formation(Le Roy and Wrana 2005). For the instance of the cell surface receptor, binding of its cognate ligand initiates signaling which both passes information into the cell but also initiates the process of clathrin mediated endocytosis to shut down the signaling(Le Roy and Wrana 2005). Endocytosis of a cell surface receptor prevents chronic signaling through the activated receptor. The receptor is internalized into the cell and transported to the

endosome. At the endosome the receptor separates from its cognate ligand and is either sent back to the cell surface or further transported to the lysosome for degradation. Clathrin mediated endocytosis is the means for down-regulating and halting signaling in many cell surface receptors including those important in cellular homeostasis and disease.

Initiation of the formation of a clathrin coat requires many complex steps. Many highly regulated protein interactions are also required for productive endocytosis. First adaptor proteins, major constituents of the protein coat, bind to the cytoplasmic region of the cell surface receptor. Clathrin subsequently binds to the adaptor proteins and polymerizes into a polyhedral lattice(Ungewickell and Hinrichsen 2007). It is thought that the initial clathrin lattices formed on the plasma membrane are relatively flat when compared to the final vesicle. As previously mentioned flat membranes are the most favorable state for the lipid bilayer. Inducing curvature is the role of additional accessory proteins during endocytosis (Hinrichsen, Meyerholz et al. 2006; Ungewickell and Hinrichsen 2007). Curvature of both the membrane and the clathrin lattice is necessary. The epsin family of proteins can insert a hydrophobic helix into the membrane to deform the lipid bilayer (Owen, Collins et al. 2004). This may provide the initial means for promoting curvature. Finally, additional adaptor and accessory proteins directly link the membrane to the clathrin coat, provide regulatory checkpoints and link the budding vesicle to the cellular cytoskeleton for force generation and transport (Engqvist-Goldstein and Drubin 2003).

Though many of the molecular details of clathrin lattice assembly have been determined, the fundamental driving forces controlling lattice assembly, rearrangements and curvature

are still unknown. For the clathrin lattice a geometric means of energetic restraint can be attributed to the separation of flat versus curved lattices. Flat lattices have primarily a hexagonal meshwork. To form a completely closed spherical lattice, like a coated vesicle, exactly twelve pentagons must be incorporated into the hexagonal meshwork (Lisanti, Flanagan et al. 1984). Accessory proteins that act in formation of a clathrin lattice may also play a role in defining the number of pentagons forming in the clathrin lattice meshwork. The geometric regulation of clathrin lattice formation can be thought as an energetic proofreading mechanism for a clathrin lattice. As a flat lattice, the hexagonal meshwork is the lowest energy state (Fig. 2). If the hexagonal meshwork is bent significantly it will either disassemble or spring back to the flat state. This prevents unintended vesicle formation directly from the flat lattice with no regulation. If some factor alters the hexagonal meshwork to incorporate a heptagon or polygon other than a pentagon there is high amounts of molecular strain and the meshwork will likely rearrange back to the low energy hexagonal meshwork to relieve the molecular strain of the triskelion (Fig. 2). If however a pentagon is induced the new low energy state will be curved because the triskelion is mildly strained when forming a pentagonal meshwork and to relieve the molecular strain the meshwork adopts a three dimensional curve (Fig. 2). This likely provides a simple and elegant means for regulating clathrin lattice assembly and provides a pathway for small scale changes within a molecular to lead to large scale manipulations of huge macromolecular complexes.

Regulation of clathrin mediated endocytosis occurs at many levels. Tight regulation is necessary to prevent unintended signaling from occurring and to maintain the integrity of the cellular endo-membrane system. Regulation of endocytosis can occur at the systems

level by modulation of number and types of protein interactions that occur within the cell. Additionally, the canonical mode of regulation, phosporylation, acts during clathrin mediated endocytosis at multiple steps. At a more fundamental level, regulation of clathrin mediated endocytosis can occur through modulating the activity of individual proteins. The most basic step in clathrin mediated endocytosis, clathrin lattice formation, appears to be regulated in all these ways. The attachment of clathrin coated vesicles to the actin cytoskeleton appears to provide yet another site of regulation for this important process.

To fully define the mechanisms involved in each step of clathrin mediated endocytosis knowledge of all proteins involved and how they act is required. Since protein structure and function are intimately linked, the most fundamental starting point for elucidating mechanisms of cellular processes is with the structure of the proteins involved. Determining protein structures provides information on potential functional roles for proteins, allows determination of stability of proteins, provides insights into protein interactions and often defines a means of regulation of function. Because of the wealth of knowledge protein structure provides toward cellular mechanisms I sought to determine the structures of two important proteins involved in clathrin mediated endocytosis. First, I sought to determine the structure of clathrin and in particular clathrin light chain, the regulatory subunit of the clathrin molecule. Next, I sought to understand how clathrin coat structures were linked to the actin cytoskeleton by determining the structure of the protein huntingtin interacting protein-1 (hip1). I used the method of X-ray crystallography to determine these structures due to its position as the "gold standard" for

protein structure determination and for hip1 I additionally used electron microscopy for its ability to define the structure of individual protein molecules.

To determine mechanistic biological roles for these proteins, I sought to augment the protein structure data with techniques that quantitatively describe protein interactions and conformations so that I could define the modes of function for each protein. These methods include numerous biochemical assays, surface plasmon resonance, fluorescence resonance energy transfer, analytical ultracentrifugation, atomic force microscopy, mass spectrometry and circular dichroism. The diversity of approaches allowed functional dissection in mutliple ways leading to findings that would not have otherwise been discovered.

References

Bonazzi, M. and P. Cossart (2006). "Bacterial entry into cells: a role for the endocytic machinery." FEBS Lett 580(12): 2962-7.

Brodsky, F. M., C. Y. Chen, et al. (2001). "Biological basket weaving: formation and function of clathrin-coated vesicles." Annu Rev Cell Dev Biol 17: 517-68.

Engqvist-Goldstein, A. E. and D. G. Drubin (2003). "Actin assembly and endocytosis: from yeast to mammals." Annu Rev Cell Dev Biol 19: 287-332.

Hinrichsen, L., A. Meyerholz, et al. (2006). "Bending a membrane: how clathrin affects budding." Proc Natl Acad Sci U S A 103(23): 8715-20.

Le Roy, C. and J. L. Wrana (2005). "Clathrin- and non-clathrin-mediated endocytic regulation of cell signalling." Nat Rev Mol Cell Biol 6(2): 112-26.

Lisanti, M. P., M. Flanagan, et al. (1984). "Clathrin lattice reorganization: theoretical considerations." J Theor Biol 108(1): 143-57.

Marsh, M. and A. Helenius (2006). "Virus entry: open sesame." Cell 124(4): 729-40.

McMahon, H. T. and I. G. Mills (2004). "COP and clathrin-coated vesicle budding: different pathways, common approaches." Curr Opin Cell Biol 16(4): 379-91.

Owen, D. J., B. M. Collins, et al. (2004). "Adaptors for clathrin coats: structure and function." Annu Rev Cell Dev Biol 20: 153-91.

Stagg, S. M., P. LaPointe, et al. (2007). "Structural design of cage and coat scaffolds that direct membrane traffic." Curr Opin Struct Biol 17(2): 221-8.

Ungewickell, E. J. and L. Hinrichsen (2007). "Endocytosis: clathrin-mediated membrane budding." Curr Opin Cell Biol 19(4): 417-25.

Figure 1. Schematic diagram of topologically equivalent compartments in a cell. The white lumens of vesicles and organelles is topologically the same as the outside of the cell. The grey cytosol is on the opposite side of the membrane.

Figure 2. Schematic diagrams of clathrin lattice geometries. The hexagonal lattice is the most frequent flat lattice. The pentagonal lattice incorporates curvature. The heptagonal lattice though occasionally seen, is unlikely to form.

Figure 1.

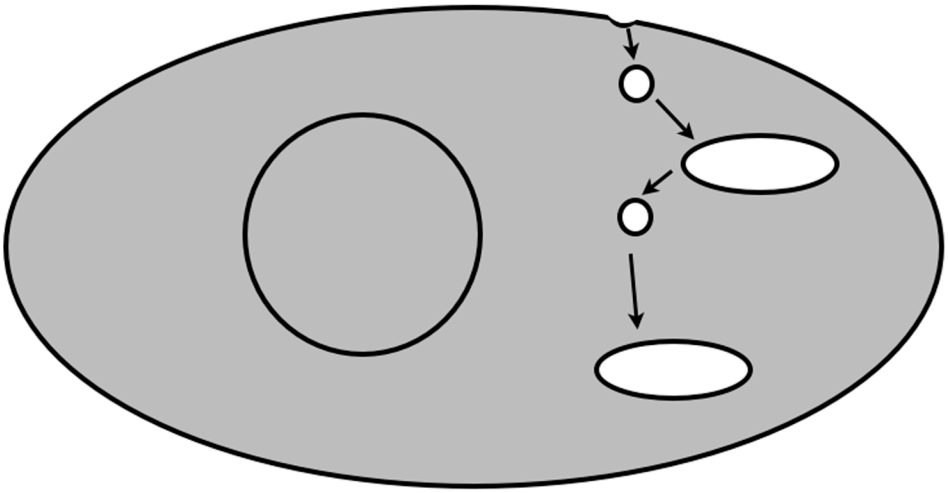
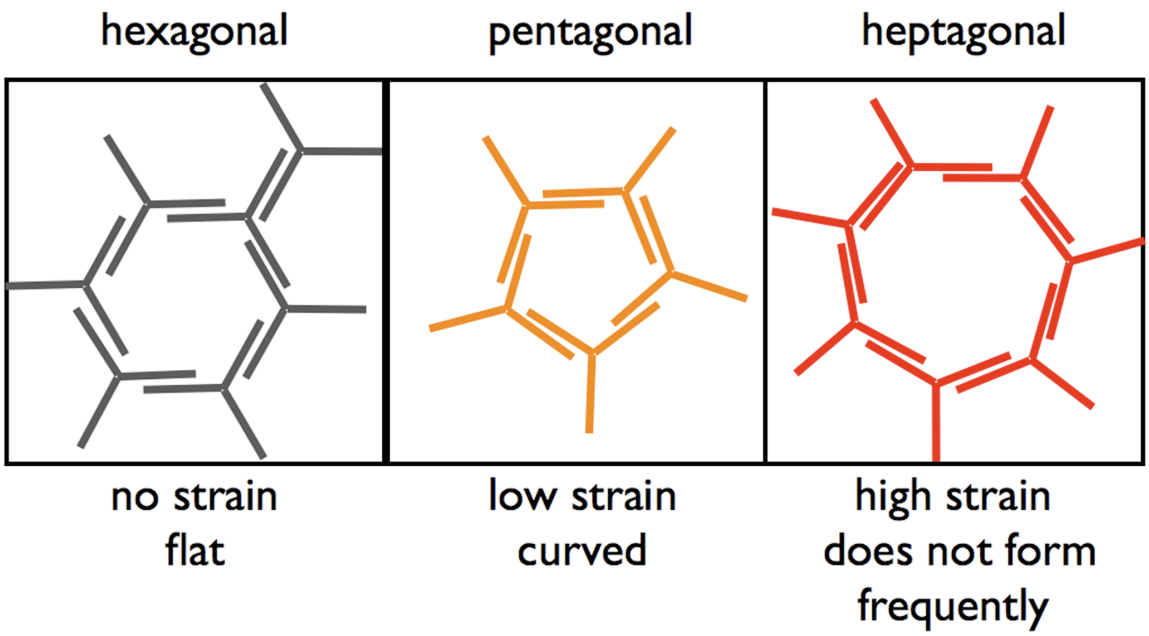


Figure 2.


hexagonal
pentagonal
heptagonal
no strain
flat
low strain
curved
high strain
does not form
frequently

Chapter 1 - Conformational regulation of clathrin assembly

Introduction

The unique chemical structure of proteins imbues them with unique biological functions. The precise three dimensional fold of a protein determines the surface properties and exposure of certain amino acids. Given that proteins function by interacting with other proteins or binding substrates, the distribution of surface amino acids and the exposure of internal amino acids greatly affect the activity of any protein. All means of regulating proteins essentially change either the surface properties or three dimensional structure or both. Recently, a more subtle means of regulation has been recognized. Modulating the inherent flexibility of proteins can have functional effects on the protein's function.

Conformational changes in the three dimensional structure of a protein provides a dramatic means of changing both the surface properties and exposure of amino acids. The amount of change proteins undergo varies dramatically, though often very modest changes can greatly affect protein function.

Of the proteins involved in clathrin mediated endocytosis few have had well characterized conformational changes associated with function or regulation. But given the size, number of the proteins necessary for endocytosis, and need for very tight regulation, it seems likely numerous conformational changes would be involved.

Clathrin is a uniquely shaped molecule forming a three legged pin wheel, or triskelion (Brodsky, Chen et al. 2001). Each clathrin triskelion consists of three clathrin heavy chain molecules and three light chain molecules (Brodsky, Chen et al. 2001; Kirchhausen 2000). Each clathrin heavy chain can bind one clathrin light chain. Each clathrin leg can be broken into six domains. Each domain plays a specific role in formation of clathrin lattices and clathrin function in general.

At the N-terminus is the globular terminal domain with a beta-propeller fold (ter Haar, Musacchio et al. 1998). This domain is responsible for interacting with adaptor proteins. Adaptor proteins bind in two separate surface grooves by burying large hydrophobic residues in these grooves (ter Haar, Harrison et al. 2000). Each groove binds a specific sequence motif though it is unclear how binding to these grooves is regulated. Adjacent to the terminal domain is the ankle region, which is followed by the distal leg. The ankle acts as a somewhat flexible linker between the terminal domain and the rest of the clathrin molecule. The ankle contributes to lattice assembly by interacting with ankle domains from other triskelia within the clathrin lattice. The distal leg has a large interaction surface involved with clathrin lattice formation and also binds domains from the GGA monomeric adaptor proteins (Knuehl, Chen et al. 2006). At the knee region the clathrin heavy chain has a distinct bend, though the bend is not uniform throughout all clathrin molecules or even within one triskelion when viewed by negative stain electron microscopy (Ungewickell and Branton 1981). The knee region appears to be a region of increased flexibility with the extended clathrin leg. Following the knee region is the clathrin proximal leg. The proximal leg is a rigid straight part of the clathrin molecule

and contains binding regions for the clathrin light chain; Chen, Reese et al. 2002). Finally following the proximal leg is the trimerization domain that interacts with other clathrin legs to for the triskelion shape. The major structural feature of the trimerization domain is the helical tripod of long alpha helices that provide most of the binding surface for trimerization (Fotin, Cheng et al. 2004). At the very C-terminus of clathrin heavy chain are 45 amino acids that form a random coil structure, likely due to numerous proline residues that prevent secondary structure elements from forming. Starting at the ankle region the clathrin molecule folds into a superhelix of alpha helices referred to the clathrin heavy chain repeat (Ybe, Brodsky et al. 1999). This repeat continues through the distal leg, knee and proximal leg.

Clathrin light chain is a small polypeptide chain that has little to no structure in solution on it own. Clathrin light chains however gain significant helical structure upon binding clathrin heavy chain. The binding affinity of clathrin light chain to clathrin heavy chain is approximately 20nM. This tight binding suggests that much of the clathrin light chain will be complexed with clathrin heavy chain if both are present in solution. Though the binding is not so tight as to render this an inseparable complex. The primary binding of clathrin light chain occurs through an extended helix that contains two tryptophans that contribute much of the binding energy (Chen, Reese et al. 2002). In the context of a helix these two tryptophans are aligned on one side of the extend clathrin light chain helix. The structures of regions of clathrin light chain outside of this central helix have been difficult to determine.

Below is the a set of work that focused on determining low resolution X-ray structures of clathrin and resulted in a manuscript which is currently in the review process.

Conformation switching of clathrin light chain regulates clathrin lattice assembly

Incollaborations with:

Peter K. Hwang[2], Joel A. Ybe[3,4,5], Michael Lane[3,4], Benjamin D. Sellers[1], Matthew P. Jacbson[3], Robert J. Fletterick[2], Frances M. Brodsky[3, 4]

[1]Graduate Program in Biophysics

[2]Department of Biochemistry and Biophysics

[3]Department of Pharmaceutical Chemistry

[4]The G.W. Hooper Foundation, Departments of Microbiology and Immunology and of Biopharmaceutical Sciences, University of California, San Francisco, California 94143

[5]Present address: Department of Biology, Indiana University, Bloomington, IN 47405

Clathrin-mediated membrane traffic influences cell signaling and metabolic homeostasis by regulating receptor targeting during endocytosis and organelle biogenesis. Receptor sorting in these pathways is a function of self-assembly of the triskelion-shaped clathrin protein into a polyhedral lattice that coats transport vesicles, leading to concentration of vesicle cargo (Brodsky, Chen et al. 2001; Ungewickell and Hinrichsen 2007). The triskelion is formed from three clathrin heavy chain (CHC) subunits, with associated clathrin light chain (CLC) subunits that modulate spontaneous CHC self-assembly (Ungewickell and Ungewickell 1991). Here we define a molecular mechanism for this regulation by determining the structure of CLC bound to the minimal self-assembling region of CHC. Solving this structure from the first co-crystal of clathrin subunits required a novel approach to analyzing low resolution (8.3 Å) crystal diffraction data. A single clathrin trimer was found in the asymmetric unit and two distinct conformations of CLC were resolved per trimer, each associated with conformational differences in CHC. Based on structure, mutagenesis and biochemical assays, we identify an electrostatic interaction between the subunits that allows a conformational change in CLC to alter the conformation of the CHC and influence assembly. This mechanism reveals that CLC negatively regulates clathrin lattice assembly by reducing CHC flexibility and explains why the influence of CLC on assembly has been difficult to detect at the cellular level. Clathrin self-assembly has inspired the design of small molecule self-assembly systems that form multidimensional polyhedral lattices (Matsuura, Murasato et al. 2005; Ghosh, Reches et al. 2007), so the mechanism we have defined

has significance for nano-system regulation, as well as for regulation of intracellular membrane traffic.

In a cell, clathrin lattice assembly is directed onto membrane surfaces and is the driving force for organizing the adaptor coat proteins, thus sorting their associated cargo into transport vesicles. Cellular clathrin assembly is easily regulated, as clathrin triskelia intertwine into lattices using broadly distributed, but weak interactions (Wakeham, Chen et al. 2003; Fotin, Cheng et al. 2004). Adaptor components of clathrin-coated vesicles (CCVs) regulate clathrin assembly in vivo, and in vitro studies indicate that this is achieved by manipulation of clathrin self-assembly (Greene, Liu et al. 2000). In vitro studies of purified clathrin and of the minimal self-assembling domain of clathrin, the trimerized C-terminal third (residues 1074-1675) forming the hub, have implicated the CLC subunit in negative regulation of the self-assembly reaction (Ungewickell and Ungewickell 1991; Liu, Wong et al. 1995; Ybe, Greene et al. 1998). CHC triskelia or hubs without CLC assemble readily at physiological pH, while those with bound CLC only assemble at lower pH. It is therefore puzzling that siRNA depletion of CLC from tissue culture cells or in vivo perturbation of CLC's ability to regulate clathrin assembly has little effect on endocytosis, one of the major pathways mediated by CCVs (Huang, Khvorova et al. 2004; Chen and Brodsky 2005; Girard, Allaire et al. 2005; Wang, Wang et al. 2006). The goal of this study was to determine by structural analysis how CLC regulates the basic clathrin assembly reaction in order to understand CLC activity in cells.

There are two isoforms of CLC in vertebrates (LCa and LCb) with 60% sequence identity and each has neuronal splicing variants (Holmes, Biermann et al. 1984; Wong, Ignatius et al. 1990). The differential and tissue specific functions of the CLCs have yet to be defined. Their longest shared sequence, a 22-residue stretch near their N-termini, is the binding site for huntingtin-interacting proteins (Hips) and comprises an acidic patch (EED) responsible for regulation of the pH sensitivity of clathrin self-assembly (Ybe, Greene et al. 1998; Chen and Brodsky 2005; Legendre-Guillemin, Metzler et al. 2005). CHCs are divided into domains starting at the N-terminal domain (residues 1 to 330 in mammalian CHC), which is followed by the ankle (~331 to 838), distal leg (~839 to 1073), knee (~1074 to 1197), proximal leg (~1198 to 1575) and trimerization domain (~1576 to 1675). A 2.9 Å resolution structure of the proximal leg revealed CHC repeats (CHCRs) comprising 10 helices of 10-12 residues each connected by loops, creating two helical and two loop faces to the triskelion leg (Ybe, Brodsky et al. 1999). Mapping CLC position along CHC by electron microscopy and immunolabeling has generated conflicting evidence for both a "U" shaped and an extended CLC (Ungewickell 1983; Nathke, Heuser et al. 1992). A reconstructed image of the assembled clathrin lattice between 7.9 and 12 Å resolution based on cryo-electron microscopy (cryoEM) showed density in which the backbone of the central third of the CLC could be resolved (Fotin, Cheng et al. 2004) as an alpha helix bound to the loop face of the CHC, as predicted by mutagenesis and modeling studies (Chen, Reese et al. 2002). How the N and C-terminal thirds of CLC interact with CHC has not yet been established.

To define further structural features of CLC-CHC interaction and establish the mechanism by which CLC influences clathrin self-assembly, we determined the structure of CLC (bovine LCb, neuronal isoform) bound to the CHC hub by X-ray crystallography. This structure represents the first X-ray crystal structure of a self-assembling unit of clathrin and the first structure of clathrin in an unassembled state. Purified hub-LCb complex was crystallized by hanging drop vapor diffusion and formed tetragonal crystals in the spacegroup $I4_122$. Despite significant efforts to improve the resolution, the best crystals diffracted to 8.3Å. Very few crystal structures have been determined at this low resolution. It was necessary to combine multiple sources of information to phase the crystal data by molecular replacement. Protein modeling in combination with the previously determined cryoEM structure was used to develop an all atom model of clathrin. Specifically, homology models of each CHCR and the trimerization domain were created based on revised sequence alignments for CHC and then structurally fit to the cryoEM structure. Loops between CHCRs were refined using a previously describe loop modeling protocol (Jacobson, Pincus et al. 2004). The all atom model was then energy minimized and truncated for molecular replacement, creating a final search probe that was a trimer of CHC residues 1280-1630 with all side chain atoms but no CLC information. Initial maps generated by molecular replacement clearly showed density that could be assigned to residues N-terminal to 1280 of CHC and to CLC. Additional modeled CHC regions were built onto the structure and refined against crystal data. Overall >90% of CLC was built as polyalanine helices and the entire structure was refined to Rwork and Rfree of 0.29 and 0.36 respectively (supplementary table S1). The previously determined 2.9 Å resolution structure of the proximal leg region (Ybe,

Brodsky et al. 1999), fit very well into the refined electron density of the clathrin hub from the 8.3 Å resolution hub-LCb co-crystal (Fig 1b, Fig1c).

There is one hub-LCb complex per asymmetric unit of the solved crystal structure and the three hub legs have unique structures. One crystal contact might reflect assembly like interactions but most contacts were compatible with hub in a disassembled state. Comparison to the cryoEM lattice structure revealed that the "pucker" of the unassembled CHC hub is similar to that of assembled triskelia and that the proximal leg regions are rigid and straight (Fig 1c, Fig 2a). In contrast, the knee regions in the crystal structure of disassembled hubs display a "bent" conformation for one leg and a "straight" conformation for two legs (Fig 2a, supplemental movie S1). The straight knee conformations are associated with extended forms of CLC (Fig 2a, supplemental movie S1), while the bent knee conformation displayed by the third leg is associated with a compact or "U"-shaped form of CLC (Fig 2a). The bent knee conformation overlayed very well with the structure of CHC from cryoEM reconstructions (supplementary Fig S1) and the associated CLC has contacts with the trimerization domain and about half of the proximal leg (Fig 2a, supplemental movie S1), with its N-terminus bent back towards the CHC trimerization domain. The CLCs associated with the straight knees extend from the trimerization domain to the knee region, covering the entire proximal leg, with their N-termini bound in the crease of the knee (Fig 2a, supplemental movie S1). For one of these extended CLCs, density assignable to nearly the entire CLC was visible in the structure. For the other two CLCs, the connection between the extreme N-terminus and the central domain was not resolved (Fig 2a, supplemental movie S1). In the compact

CLC, the apparent free segment was stabilized by a crystal contact and the region between the two arrowheads is presumably contiguous (Fig 2a). The C-terminus of each CLC, independent of leg-binding conformation, contacts helices derived from two different CHCs, explaining the biochemical and genetic evidence that CLCs contribute structural stability to the trimerization domain (Huang, Gullberg et al. 1997; Pishvaee, Munn et al. 1997; Chen, Reese et al. 2002; Wang, Wang et al. 2006; Ybe, Perez-Miller et al. 2007) (Fig 2a).

To test whether the two conformations of CLC in the crystal structure occur in solution, relative Förster resonance energy transfer (FRET) was measured between fluorescent probes attached near the N and C-termini of CLC. CLC (human LCb, neuronal isoform) that was engineered to have one cysteine at either terminus (mutations S9C and C190S) was labeled with a mixture of Alexa 555 and Alexa647. This procedure yielded CLCs with a different fluorophore at either terminus or labeled with the same fluorophore (Fig 2b). Labeled CLC were bound to CHC hubs and their FRET measured under disassembled conditions. The extended conformation of CLC would be too distant for significant FRET between the coupled fluorphores, but a background level of FRET (efficiency of 49% ± 8%) results from different fluorophores at the C-termini of CLCs juxtaposed at the hub trimerization domains (Fig 2b, Fig 2c). Hubs with labeled CLCs were then mixed with an excess of hubs with unlabeled CLCs, and FRET was measured following induction of assembly. The FRET signal was significantly increased (efficiency of 78% ± 10%) when hubs were assembled (Fig 2c). This change in FRET, reflecting increased proximity of the CLC termini, can be ascribed to a change in CLC

conformation upon assembly. The excess of hubs with unlabeled CLC eliminates the possibility that increased FRET upon assembly results from interactions between triskelia. These FRET data confirm that CLC undergoes a conformational change to a more compact state when clathrin assembles.

In the extended CLC conformation, the N-terminal domain interacts with CHC residues in the crease of the knee, revealing one small CLC helix packed into a loop of CHC comprised of basic residues (KR loop) (Fig 3a). This localization suggested the possibility that the KR loop might be the complementary half of an interaction with the N-terminal EED sequence of CLC that controls pH sensitivity of clathrin assembly. The KR loop is also the hinge point for the straight and bent CHC knee (Fig 3a), situating that loop at a site where CHC knee conformation would be affected by CLC binding. To test the role of the KR loop in electrostatic assembly regulation, K1163 was mutated to glutamate and R1165 was mutated to aspartate in recombinant CHC hub fragments. When CLC was bound to the KR-mutant hub, it was not able to inhibit assembly at pH 6.7, in contrast to CLC effects on wild-type hub assembly (Fig 3c). Both wild type and KR loop mutant CHC hub bound CLC with similar affinity revealing a slight difference in on rate, with no difference in off rate, and both hub forms were fully saturated with CLC (Fig 3b, data not shown). The effect of KR loop mutation on CLC regulation of hub assembly was identical to the effect of mutating the EED residues in CLC(Ybe, Greene et al. 1998), establishing the KR loop as the complementary half of the electrostatic switch by which CLC affects clathrin assembly.

In the extended CLC conformation, the N-terminus is bound to a face of the CHC that is not involved in direct leg-leg contacts in the assembled lattice. Thus negative regulation of CHC assembly by CLC is not simply due to steric blocking of lattice formation. Rather, the location of the CLC binding site in the crease of the knee suggests its influence on knee flexibility affects assembly. Flexibility at the triskelion knee has been proposed as a key contributor to the hexagon-pentagon switch that imparts curvature to the clathrin lattice (Musacchio, Smith et al. 1999). Additionally, the consistent angle of orientation of triskelion legs is lost when clathrin light chains are removed, supporting a role for CLC in controlling knee bending (Schmid, Matsumoto et al. 1982; Ungewickell 1983). To visualize CLC-induced changes in CHC conformation and how they relate to the mechanism of clathrin assembly, complete clathrin triskelia with straight or bent knees were modeled by structural alignment (Fig 4a). The hubs of these triskelia were then aligned according to hub orientation in an assembled lattice (to both a pentagon and a hexagon edge). This alignment revealed that the straight knee conformation would clash with adjacent triskelia, and be refractory to assembly (Fig 4b). Critical assembly interactions involving the distal leg and the ankle domains have been predicted (Wakeham, Chen et al. 2003; Fotin, Cheng et al. 2004) and both would be misaligned in the extended CLC-straight CHC knee conformation (Fig 4). The compact CLC conformation is permissive for assembly by allowing flexibility of the CHC knee and bending to accommodate to either hexagonal or pentagonal angles.

The precise geometry of clathrin triskelia plays a role in generating productive lattices (Greene, Liu et al. 2000; Fotin, Cheng et al. 2004). The structure that we have

determined reveals conformational changes of CLC that change triskelion geometry into a conformation incompatible with closed lattice assembly. The structure also identifies an electrostatic interaction between CLC and CHC that regulates CLC's ability to change CHC conformation and we demonstrate that this electrostatic interaction accounts for CLC effects on clathrin assembly in vitro. The molecular mechanism defined here reveals that CLC's negative regulation of clathrin assembly is due to its ability to decrease knee flexibility. Conversely, absence of CLC creates triskelia with greater knee flexibility. Therefore clathrin is able to assemble without CLC in vitro and to assemble from pieces when the distal leg is completely severed from the knee (Ungewickell, Unanue et al. 1982; Greene, Liu et al. 2000; Poupon, Girard et al. 2008), explaining its ability to assemble and function in tissue culture cells depleted of CLC (Huang, Khvorova et al. 2004). Proteomic analysis of CCVs from different tissues suggests that CLC expression levels are stoichiometric with CHC only in brain (Blondeau, Ritter et al. 2004). Fewer CLCs per triskelion in non-neuronal tissue would decrease the influence of CLCs on assembly and render various assembly-promoting adaptors more influential. Thus, we propose that the ability of CLC to control clathrin assembly is more effective in neurons where clathrin assembly has to be finely coordinated for synaptic vesicle protein recapture. This is consistent with the observation that in non-neuronal tissue the main detectable phenotype of CLC disruption is related to the Hip-binding function of CLCs and consequent regulation of the actin cytoskeleton, rather than an effect on CHC assembly(Chen and Brodsky 2005; Poupon, Girard et al. 2008). CLCs are also likely to have additional functions mediated by the various protein-binding modules that are present in the CLC sequences (Brodsky, Chen et al. 2001). It is interesting to note,

however, that phosphorylation sites for LCb and a binding site for the hsc70 uncoating protein on LCa are situated in the N-terminus where these properties could affect CLC conformation (Hill, Drickamer et al. 1988; DeLuca-Flaherty, McKay et al. 1990). The approximate position of the hsc70-binding site would be where the compact form of CLC bends back towards the trimerization domain, so hsc70 binding might cause CLC extension and thereby stimulate disassembly. The predicted phosphorylation sites for LCb are located in the knee-binding region, so phosphorylation might inhibit the extended conformation and promote assembly. These potential regulatory features also suggest that knee binding could be gradually manipulated such that CLC conformation would influence the hexagon-pentagon switch during lattice curvature. In conclusion we have determined the first complete structure of CLC and defined the mechanism for CLC regulation of clathrin lattice assembly. This should provide a framework for establishing how other modes of regulation occur in clathrin mediated endocytosis, particularly in neurons, and generate new concepts for understanding regulation of self-assembling systems in general.

Methods Summary

Bovine clathrin heavy chain residues 1074-1675 (hub) were expressed and purified as previously described (Liu, Wong et al. 1995). Bovine neuronal clathrin light chain b (LCb) used in crystallography was expressed and purified as previously described (Ybe, Greene et al. 1998). Human neuronal LCb was expressed and purified as a his-tagged fusion protein for assembly assays, FRET assays and SPR. His-tagged LCb behaved identically to untagged bovine LCb (data not shown). Crystals were grown in 200mM

Citrate, 16-22% glycerol, 2% trifluoroethanol. Data was collected at the Advanced Light Source beamline 8.3.1. Structure was phased by molecular replacement as described in the text and refined with CNS. FRET assays were performed by labeling a double mutant of LCb (S9C and C190S) with a mixture of Alexa 555 and Alexa 647 fluorescent dyes and measuring steady state energy transfer under assembly or disassembly conditions. Assembly assays were performed as previously described (Ybe, Greene et al. 1998). For SPR analysis LCb was covalently coupled to the surface of a CM5 flow cell (Biacore) and clathrin hub was flowed over the surface. Data was collected on a Biacore T100 and processed with Scrubber (University of Utah, Center for Biomolecular Interactions).

References

Blondeau, F., B. Ritter, et al. (2004). "Tandem MS analysis of brain clathrin-coated vesicles reveals their critical involvement in synaptic vesicle recycling." Proc Natl Acad Sci U S A **101**(11): 3833-8.

Brodsky, F. M., C. Y. Chen, et al. (2001). "Biological basket weaving: formation and function of clathrin-coated vesicles." Annu Rev Cell Dev Biol **17**: 517-68.

Chen, C. Y. and F. M. Brodsky (2005). "Huntingtin-interacting protein 1 (Hip1) and Hip1-related protein (Hip1R) bind the conserved sequence of clathrin light chains and thereby influence clathrin assembly in vitro and actin distribution in vivo." J Biol Chem **280**(7): 6109-17.

Chen, C. Y., M. L. Reese, et al. (2002). "Clathrin light and heavy chain interface: alpha-helix binding superhelix loops via critical tryptophans." Embo J **21**(22): 6072-82.

DeLuca-Flaherty, C., D. B. McKay, et al. (1990). "Uncoating protein (hsc70) binds a conformationally labile domain of clathrin light chain LCa to stimulate ATP hydrolysis." Cell **62**(5): 875-87.

Fotin, A., Y. Cheng, et al. (2004). "Molecular model for a complete clathrin lattice from electron cryomicroscopy." Nature **432**(7017): 573-9.

Ghosh, S., M. Reches, et al. (2007). "Bioinspired design of nanocages by self-assembling triskelion peptide elements." Angew Chem Int Ed Engl **46**(12): 2002-4.

Girard, M., P. D. Allaire, et al. (2005). "Non-stoichiometric relationship between clathrin heavy and light chains revealed by quantitative comparative proteomics of clathrin-coated vesicles from brain and liver." Mol Cell Proteomics **4**(8): 1145-54.

Greene, B., S. H. Liu, et al. (2000). "Complete reconstitution of clathrin basket formation with recombinant protein fragments: adaptor control of clathrin self-assembly." Traffic **1**(1): 69-75.

Hill, B. L., K. Drickamer, et al. (1988). "Identification of the phosphorylation sites of clathrin light chain LCb." J Biol Chem **263**(12): 5499-501.

Holmes, N., J. S. Biermann, et al. (1984). "Comparison of the primary structures of clathrin light chains from bovine brain and adrenal gland by peptide mapping." Embo J **3**(7): 1621-7.

Huang, F., A. Khvorova, et al. (2004). "Analysis of clathrin-mediated endocytosis of epidermal growth factor receptor by RNA interference." J Biol Chem **279**(16): 16657-61.

Huang, K. M., L. Gullberg, et al. (1997). "Novel functions of clathrin light chains: clathrin heavy chain trimerization is defective in light chain-deficient yeast." J Cell Sci **110 (Pt 7)**: 899-910.

Jacobson, M. P., D. L. Pincus, et al. (2004). "A hierarchical approach to all-atom protein loop prediction." Proteins **55**(2): 351-67.

Kirchhausen, T. (2000). "Clathrin." Annu Rev Biochem **69**: 699-727.

Knuehl, C., C. Y. Chen, et al. (2006). "Novel binding sites on clathrin and adaptors regulate distinct aspects of coat assembly." Traffic **7**(12): 1688-700.

Legendre-Guillemin, V., M. Metzler, et al. (2005). "Huntingtin interacting protein 1 (HIP1) regulates clathrin assembly through direct binding to the regulatory region of the clathrin light chain." J Biol Chem **280**(7): 6101-8.

Liu, S. H., M. L. Wong, et al. (1995). "Regulation of clathrin assembly and trimerization defined using recombinant triskelion hubs." Cell **83**(2): 257-67.

Matsuura, K., K. Murasato, et al. (2005). "Artificial peptide-nanospheres self-assembled from three-way junctions of beta-sheet-forming peptides." J Am Chem Soc **127**(29): 10148-9.

Musacchio, A., C. J. Smith, et al. (1999). "Functional organization of clathrin in coats: combining electron cryomicroscopy and X-ray crystallography." Mol Cell **3**(6): 761-70.

Nathke, I. S., J. Heuser, et al. (1992). "Folding and trimerization of clathrin subunits at the triskelion hub." Cell **68**(5): 899-910.

Pishvaee, B., A. Munn, et al. (1997). "A novel structural model for regulation of clathrin function." Embo J **16**(9): 2227-39.

Poupon, V., M. Girard, et al. (2008). "Clathrin light chains function in mannose phosphate receptor trafficking via regulation of actin assembly." Proc Natl Acad Sci U S A **105**(1): 168-73.

Schmid, S. L., A. K. Matsumoto, et al. (1982). "A domain of clathrin that forms coats." Proc Natl Acad Sci U S A **79**(1): 91-5.

ter Haar, E., S. C. Harrison, et al. (2000). "Peptide-in-groove interactions link target proteins to the beta-propeller of clathrin." Proc Natl Acad Sci U S A **97**(3): 1096-100.

ter Haar, E., A. Musacchio, et al. (1998). "Atomic structure of clathrin: a beta propeller terminal domain joins an alpha zigzag linker." Cell **95**(4): 563-73.

Ungewickell, E. (1983). "Biochemical and immunological studies on clathrin light chains and their binding sites on clathrin triskelions." Embo J **2**(8): 1401-8.

Ungewickell, E. and D. Branton (1981). "Assembly units of clathrin coats." Nature **289**(5796): 420-2.

Ungewickell, E., E. R. Unanue, et al. (1982). "Functional and structural studies on clathrin triskelions and baskets." Cold Spring Harb Symp Quant Biol **46 Pt 2**: 723-31.

Ungewickell, E. and H. Ungewickell (1991). "Bovine brain clathrin light chains impede heavy chain assembly in vitro." J Biol Chem **266**(19): 12710-4.

Ungewickell, E. J. and L. Hinrichsen (2007). "Endocytosis: clathrin-mediated membrane budding." Curr Opin Cell Biol **19**(4): 417-25.

Wakeham, D. E., C. Y. Chen, et al. (2003). "Clathrin self-assembly involves coordinated weak interactions favorable for cellular regulation." Embo J **22**(19): 4980-90.

Wang, J., Y. Wang, et al. (2006). "Clathrin light chain: importance of the conserved carboxy terminal domain to function in living cells." Traffic **7**(7): 824-32.

Wong, D. H., M. J. Ignatius, et al. (1990). "Neuron-specific expression of high-molecular-weight clathrin light chain." J Neurosci **10**(9): 3025-31.

Ybe, J. A., F. M. Brodsky, et al. (1999). "Clathrin self-assembly is mediated by a tandemly repeated superhelix." Nature **399**(6734): 371-5.

Ybe, J. A., B. Greene, et al. (1998). "Clathrin self-assembly is regulated by three light-chain residues controlling the formation of critical salt bridges." Embo J **17**(5): 1297-303.

Ybe, J. A., S. Perez-Miller, et al. (2007). "Light chain C-terminal region reinforces the stability of clathrin heavy chain trimers." Traffic **8**(8): 1101-10.

Acknowledgements

We thank Dr. Luke Rice and Dr. Pascal Egea for expert technical assistance with X-ray crystallography. This work was supported by NIH grants GM038093 to F.M.B and S10 RR023443 to R.J.F.

Figure 1. **8.3 Å resolution electron density of clathrin hub resolves both clathrin heavy chain helices and clathrin light chain structures. a.** Structure of the trimerization region of clathrin hub-LCb complex determined by X-ray crystallography. Clathrin heavy chain is in brown and clathrin light chain residues are in green. $2F_o$-F_c electron density (1.5σ) is shown as a blue mesh. **b.** and **c.** show the high resolution structure of the proximal leg (1B89) fitted into the $2F_o$-F_c electron density (mesh, 1.5σ) scaled to 1.5σ from clathrin hub. Arrows point to density corresponding to clathrin light chain. **b** is a cross section of **c** at the red line and red arrows correspond to same clathrin light chain density.

Figure 2. **Clathrin light chain adopts multiple conformations. a.** The overall structure of clathrin hub-LCb complex. Clathrin heavy chain residues are in brown and clathrin light chain is in green. Clathrin light chain extends from the knee region to the trimerization domain on two legs (arrows) while on the other leg, clathrin light chain forms a more compact structure (arrowheads). See movie S1 for 3-dimensional movies of the structure. Grey box corresponds to zoomed inset of the trimerization domain. Inset: At least one of three helices (indicated by asterisks) from the C-terminus of each clathrin light chain makes contact with two different clathrin heavy chains at the trimerization domain (the clathrin heavy chains are colored differently for clarity). **b.** Purified clathrin light chain was randomly labeled with Alexa 555 and Alexa 647 dyes at two positions, one near the N-terminus of clathrin light chain and one near the C-terminus. Labeled clathrin light chain retained its ability to prevent spontaneous assembly of clathrin hubs at pH 6.7. The cartoon shows the two dyes in blue and yellow

and indicates where FRET can occur. **c.** Labeled clathrin light chain was bound to clathrin hub and these complexes were mixed with complexes of hub with unlabeled clathrin light chain to prevent intermolecular energy transfer. Forster resonance energy transfer (FRET) was measured under conditions that favor either the unassembled or assembled state of the mixed labeled and unlabeled complexes. Each bar represents the average of three independent experiments plotted with standard error.

Figure 3. **The KR loop in the clathrin heavy chain knee participates in electrostatic regulation of assembly by clathrin light chain. a.** Expanded boxed regions show the KR loop (blue spheres, residues 1161-1165 in clathrin heavy chain) interacts with a helix from the extended clathrin light chain (green). The straight knee conformation (brown) associated with the extended form of clathrin light chain (green) and the bent knee conformation (cyan) associated with the compact form of clathrin light chain (not shown) have been overlaid in the boxes with the hinge point (red arrowheads) at the KR loop. For both knee conformations, helices on the proximal side of the arrowhead overlay well while helices distal to the arrowhead are shifted relative to each other. **b.** Surface plasmon resonance (SPR) analysis of clathrin light chain interactions with wild type clathrin hub or KR loop mutant hub. Human neuronal LCb was immobilized on the SPR chip surface and wild type or KR loop mutant hub flowed over. The fitted binding affinity for wild type was 25 nM while KR loop mutant was 59.1 nM. The difference in affinity was primarily due to a slowed on rate for the KR loop mutant likely due to decreased electrostatic attraction. **c.** Purified hubs (wild type or with mutations in the KR loop) were saturated with clathrin light chain (LCb) or left unoccupied and assayed

for assembly at pH 6.7 by light scattering. The ratio of assembly signal for hub with light chain to assembly signal for hub without light chain for either the wild type or KR loop mutant hubs is plotted (average of four independent experiments with standard error shown). Wild type and KR loop mutant hubs displayed comparable assembly signals without clathrin light chain bound.

Figure 4. **Altered clathrin knee flexibility associated with conformation switching of clathrin light chain explains light chain regulation of assembly. a.** Clathrin heavy chain triskelia with all straight knees or all bent knees were modeled by structure alignment from single legs. A distal leg was built from the C-alpha backbone of residues 1-1130 of the cryo-EM lattice and overlayed onto a straight or bent hub leg from the crystal structure, by alignment of residues 1074-1130, which are N-terminal to the hinge point (red arrowheads, Fig 3). Three straight legs or three bent legs were trimerized by alignment in the C-terminal part of the hub, which is invariant in the assembled or disassembled state. Straight conformation (brown) and bent conformation (cyan) are shifted 30° starting from the KR loop hinge point. **b.** Alignment of bent knee legs to adjacent triskelia (one full leg and two hub legs) from the cryoEM structure shows remarkable similarity to the assembled state. Black bars indicate regions aligned and the compact conformation of clathrin light chains is shown in green. **c.** Aligning straight knee legs to adjacent triskelia (one full leg and two hub legs) shows complete misalignment of the distal region and a significant steric clash of the knee region with the adjacent proximal region. Black bar indicate regions aligned and extended clathrin light chains are shown in green.

Figure S1. **Comparison of the crystallographically-determined structures of clathrin heavy chain legs from the hub-LCb complex with the structure of assembled clathrin determined by cryoelectron microscopy**.

Crystal structures in blue or brown were aligned to the cryo-EM structure (1XI4 – grey tubes) using the rigid part of the proximal leg (residues 1430-1560). The crystallographically determined leg structure with the bent knee (blue) aligns well while the crystallographically determined leg structure with the straight knee (brown) aligns well within the proximal leg region of the cryo-EM structure, but then deviates at the knee.

Figure S2. **Intentionally misplaced helices do not contribute positively to phase calculations.** Clathrin heavy chain in brown and clathrin light chain in green is shown surrounded by calculated 2fo-fc density (blue) and fo-fc density (red). Arrow points to a helix intentionally misplaced to test for model bias during refinement of the built structure. After one round of refinement this helix was clearly in negative density.

Figure S3. **New alignment of CHCRs from bovine clathrin heavy chain.** CHCR alignment to determine loop positions based on clathrin heavy chain sequences from nine species. Only bovine CHCR sequences are shown.

Table S1: Crystallographic data collection and refinement for clathrin hub-LCb complex

	Native data (high resolution shell)
Space group	$I4_122$
Cell dimensions (Å)	a=b=228.5 c=710.3
Wavelength (Å)	1.1
Resolution (Å)	250-8.3 (8.65-8.3)
R_{sym} (%)	7.5 (31.4)
I/σI	20 (4.2)
Completeness (%)	99 (97.2)
Redundancy	4.7 (3.6)
Total # Obs. Unique Reflections	9,144 (863)
Refinement	
Resolution (Å)	250-8.3
# of Reflections used in refinement	9,134 working set 8,238, test set 896
R_{work}	29.8
R_{free} (9.5%)	36.8

Figure. I

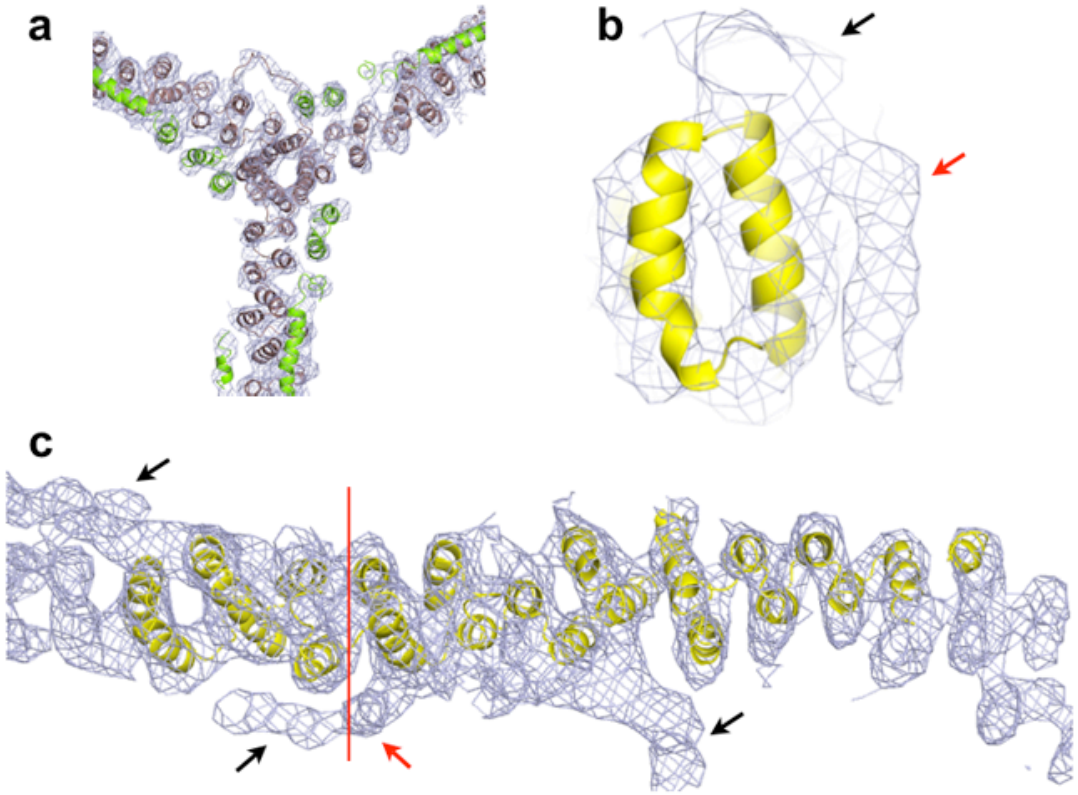


Figure. 2

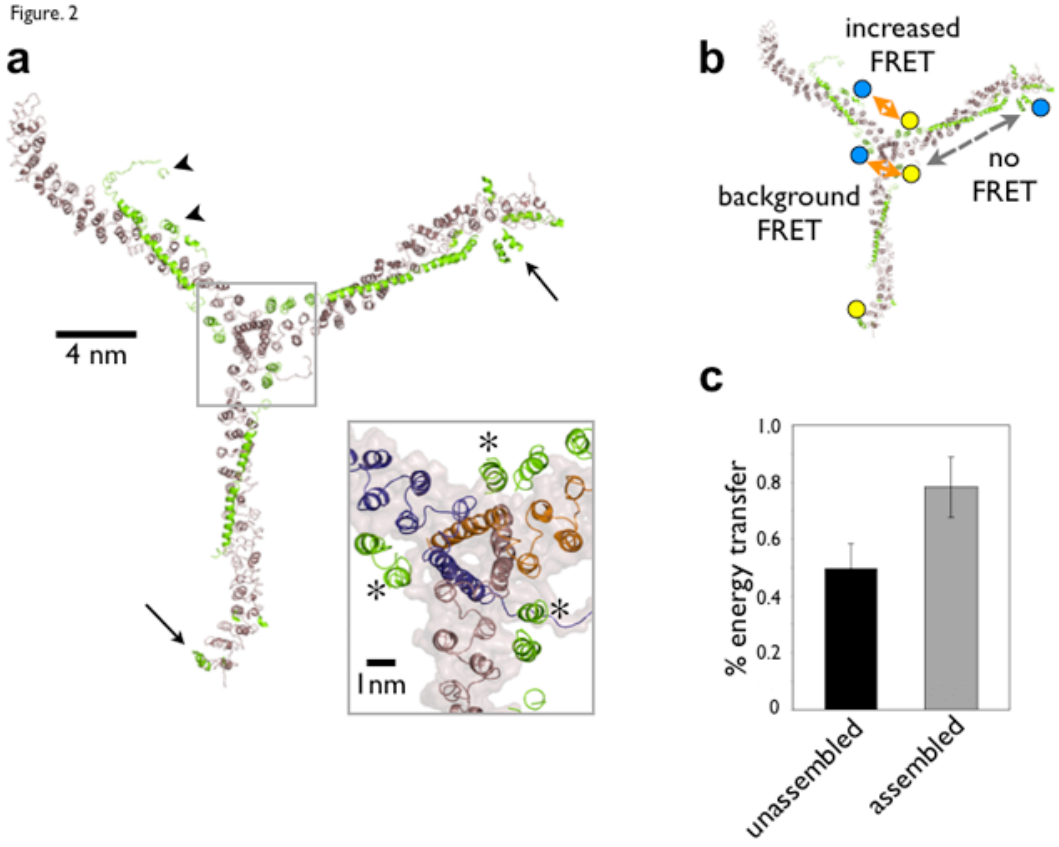

a
4 nm
1nm
b
increased
FRET
no
FRET
background
FRET
c
% energy transfer
1.0
0.8
0.6
0.4
0.2
0
unassembled
assembled

Figure. 3

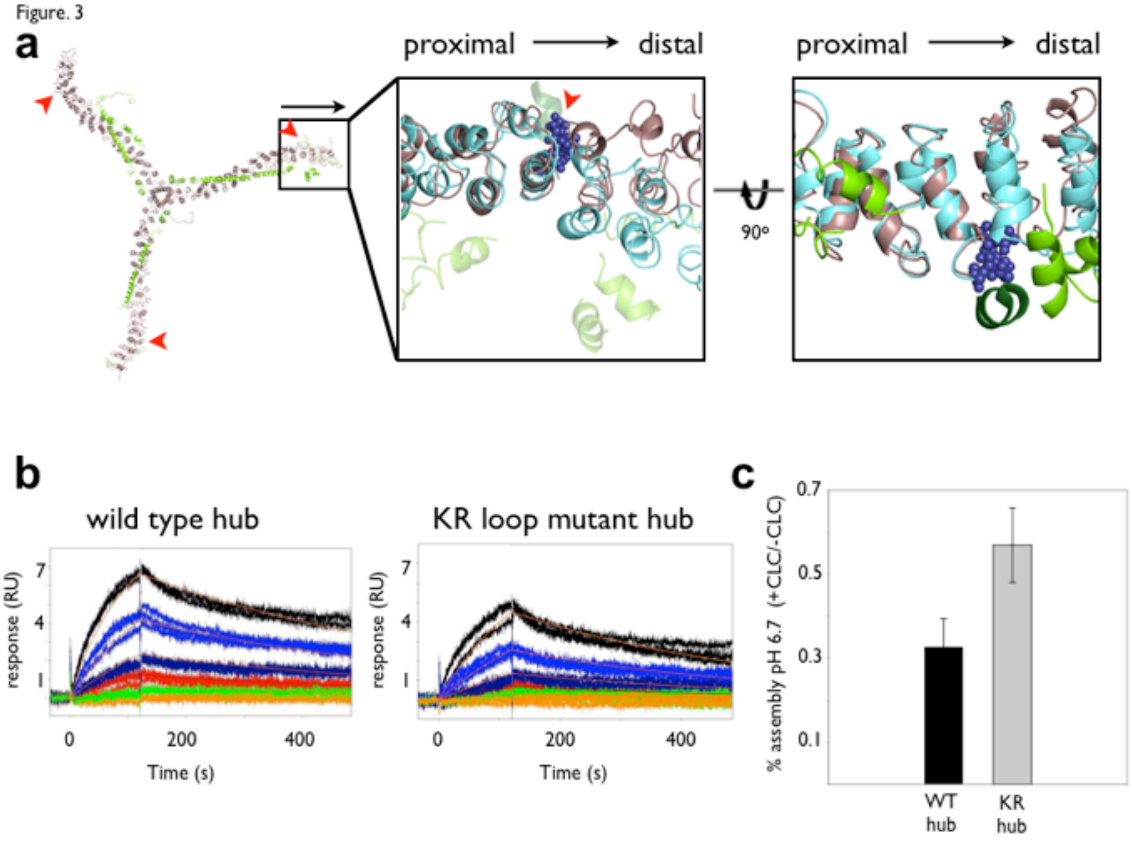
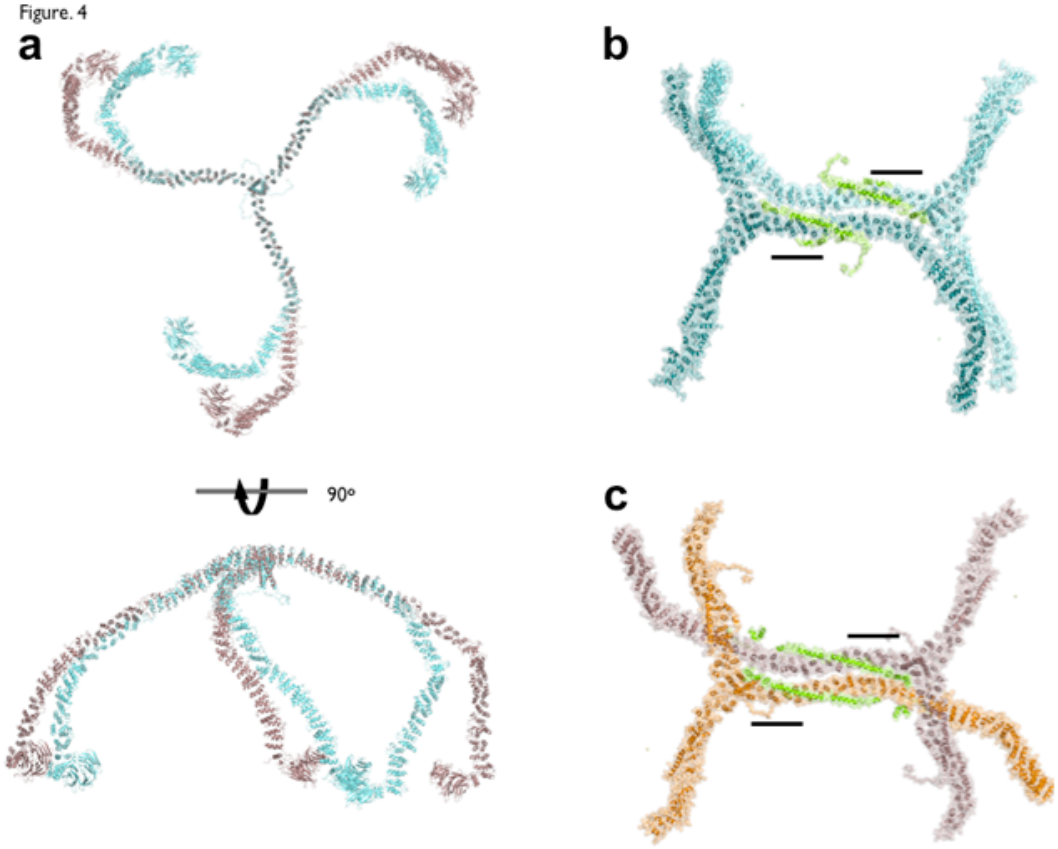
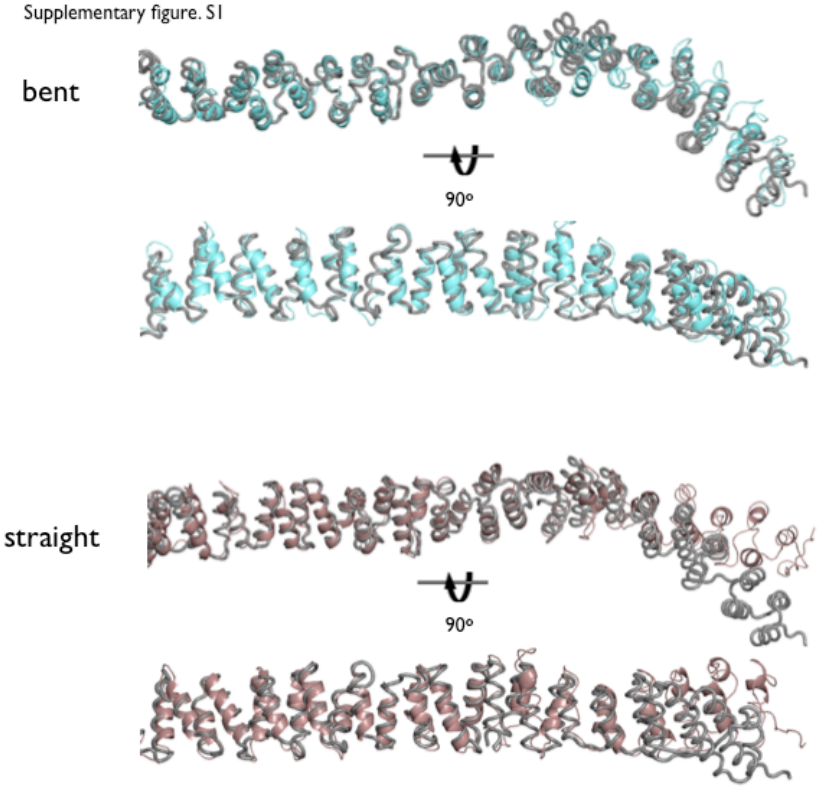

a
proximal
distal
proximal
distal
90°
b
wild type hub
KR loop mutant hub
response (RU)
7
4
1
0
200
400
Time (s)
c
% assembly pH 6.7 (+CLC/-CLC)
0.7
0.5
0.3
0.1
WT hub
KR hub

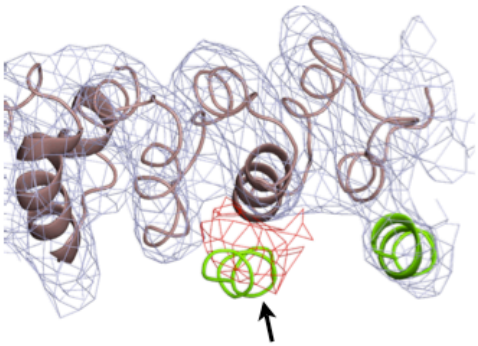
Figure. 4



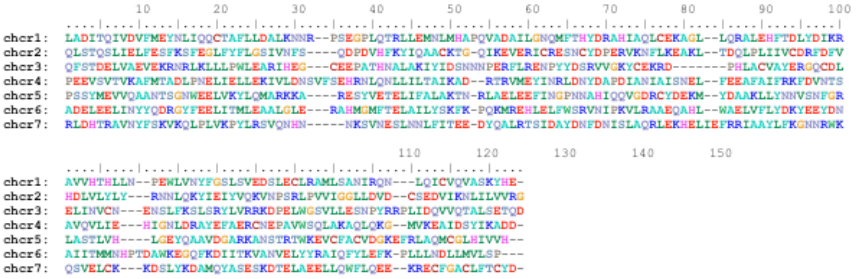
Supplementary figure. S1



Supplementary figure. S2



Supplementary figure. S3



Chapter 2 - Linking clathrin to the actin cytoskeleton

In addition to coat formation, increasing evidence is pointing toward actin polymerization as being a fundamental aspect to endocytosis (Engqvist-Goldstein and Drubin 2003). In yeast actin plays an indispensable role in clathrin mediated endocytosis (Engqvist-Goldstein and Drubin 2003). In higher eukaryotes it endocytosis can occur without actin, though it remains unclear how productive endocytosis is without actin (Pauly and Drubin 2007). It remains possible that there are many sub-pathways or cargo specific differences in clathrin mediated endocytosis some of which require actin and other's which can proceed without it. The possibility of actin defined cargo specific pathways of clathrin mediated endocytosis is seems less likely with the identification of multiple actin binding proteins that function as fundmental components of endocytosis. For example, dynamin, a protein required for scission of the budding vesicle is an absolute requirement for productive endocytosis and directly binds the actin nucleating protein Arp2/3 (Schafer, Weed et al. 2002). The HIP proteins hip1 and hip1R directly bind clathrin and actin (Engqvist-Goldstein, Warren et al. 2001). Mechanistically, the role for direct binding of clathrin and actin by the HIP proteins has been thought of as a clamp, latching the clathrin coat to polymerizing actin. In yeast the HIP protein homolog Sla2p may in fact work as a clathrin/actin clamp (Ybe, Mishra et al. 2007). In higher eukaryotes a diversification of roles for actin is indicated by the strict necessity for clathrin but not actin in endocytosis. It may be that actin's role in higher eukaryotes has been modified to add another layer of regulation. To accommodate this hypothesized role in endocytosis

the link from clathrin to actin may have changed from a static clamp to a more dynamic regulatory system.

In yeast actin appears to surround a budding vesicle acting much like a clathrin coat to help arrange proteins and membrane into a productive endocytic event (Engqvist-Goldstein and Drubin 2003). The relatively rigid actin scaffold is pushed into the cytoplasm from the membrane by the combined work of actin nucleating proteins and myosin proteins localized to the plasma membrane (Toret and Drubin 2006).

In mammalian cells dynamin interactions with Arp2/3 most likely occur at the neck of budding vesicle given that dynamin has very strong localization to this region. This would indicate that actin polymerization occurs at the neck of budding vesicle rather than at the plasma membrane. Additionally, hip1R interacts with cortactin to negatively regulate actin assembly (Le Clainche, Pauly et al. 2007). Hip1R is distributed throughout a budding vesicle coat (Engqvist-Goldstein, Warren et al. 2001). This suggests that hip1R and cortactin contrary to the yeast mechanism may disassemble or cap actin polymerization around the vesicle.

To investigate role of actin in mammalian clathrin mediated endocytosis, I sought to determine the biochemical and biophysical properties of hip1 and hip1R some of the primary clathrin and actin linking proteins. These studies determine a new molecular mechanism of regulating the link between clathrin and actin. More generally, these studies investigate how dynamic properties of proteins can contribute to their function.

The result of this work has been submitted and reviewed at the Journal of Biological Chemistry. Revisions based on reviewers' comments have been made and this manuscript will be resubmitted. The rest of this chapter is the manuscript that is prepared for submission.

Actin binding by hip1 and hip1R proteins is regulated by clathrin light chain

Jeremy D. Wilbur[1], Chih-Ying Chen[2], Venus Manalo[2] Peter Hwang[3], Robert Fletterick[3] and Frances M. Brodsky[2]

1. Graduate Program in Biophysics

2. The G. W. Hooper Foundation, Departments of Biopharmaceutical Sciences, Microbiology and Immunology, and Pharmaceutical Chemistry

3. Department of Biochemistry and Biophysics University California, San Francisco, California 94143.

Running Title: Clathrin regulates actin binding by Hip proteins

Address Correspondence to: Frances M. Brodsky, Hooper Foundation, Box 0552, UCSF, 513 Parnassus Ave., San Francisco, CA 94143-0552. Email: frances.brodsky@ucsf.edu

The Hip proteins (hip1 and hip1R and Sla2p in yeast) link clathrin-mediated membrane traffic to actin cytoskeleton dynamics. Genetic data in yeast have implicated the light chain subunit of clathrin in regulating this link. To test this hypothesis, the biophysical properties of mammalian hip1 and hip1R and their interaction with clathrin light chain and actin were analyzed. The coiled-coil domains (clathrin light chain-binding) of hip1 and hip1R were found to be stable homodimers with no propensity to heterodimerize *in vitro*. Homodimers were also predominant *in vivo*, accounting for cellular segregation of hip1 and hip1R functions. Coiled coil domains of hip1 and hip1R differed in their stability and flexibility correlating with different affinities for clathrin light chain. Clathrin light chain binding induced a compact conformation of both hip1 and hip1R and markedly reduced actin binding by their THATCH domains. Thus clathrin is a negative regulator of Hip-actin interactions. These observations necessarily change models proposed for Hip protein function.

Regulation of membrane traffic depends on molecules that link membrane vesicle formation to components of the actin cytoskeleton. Huntingtin interacting protein 1 (hip1) and its homolog hip 1-related (hip1R), which both bind clathrin and actin, are implicated in a functional connection between the two. Hip1 and hip1R have the same domain structure consisting of an N-terminal phospholipid binding domain (ANTH), a central coiled coil and a C-terminal actin binding domain (THATCH) (Engqvist-Goldstein and Drubin 2003). The coiled coil domains of both proteins comprise the binding site for clathrin light chains (Legendre-Guillemin, Metzler et al. 2002; Legendre-Guillemin, Metzler et al. 2005),(Chen and Brodsky 2005). The coiled coil domain of hip1 contains binding sites, which are not present in hip1R, whose functions have been linked to caspase activation, androgen receptor activation and the pathogenesis of Huntington's disease (Wanker, Rovira et al. 1997),(Kalchman, Koide et al. 1997),(Hackam, Yassa et al. 2000; Gervais, Singaraja et al. 2002),(Mills, Gaughan et al. 2005). Interactions between hip1 and hip1R have been shown (Legendre-Guillemin, Metzler et al. 2002), though the molecular details remain

unclear. Extensive heterodimerization seems unlikely, as it would mix the functions of hip1 and hip1R, which are distinct in cells, and genetic deletion of each protein produces a markedly different phenotype (Bradley, Hyun et al. 2007),(Metzler, Li et al. 2003). To understand the molecular basis for the functions of hip1 and hip1R, the biophysical properties of the coiled coil domains were analyzed with respect to their self-association and their intramolecular influence on actin binding.

Clathrin coated vesicles mediate major membrane traffic pathways including receptor-mediated endocytosis and protein sorting during lysosome biogenesis (Brodsky, Chen et al. 2001). Clathrin is a triskelion-shaped molecule made up of three heavy chains and three light chains, which polymerizes to form a lattice-like vesicle coat that organizes cargo and proteins that control vesicle budding (Brodsky, Chen et al. 2001),(Kirchhausen 2000),(Traub 2005),(Ungewickell and Hinrichsen 2007). In both mammalian and yeast cells, clathrin-coated vesicle budding also involves actin polymerization (Engqvist-Goldstein, Kessels et al. 1999). In mammalian cells, overexpression of the Hip protein-binding region of clathrin light chain causes an alteration in actin cytoskeleton structure, generating short actin protrusions tipped with cortactin (Chen and Brodsky 2005). The complex of hip1R and cortactin has been shown to slow or block actin assembly and is proposed to mediate the positioning of actin assembly on a budding vesicle (Le Clainche, Pauly et al. 2007). In yeast cells that are deficient in clathrin heavy chain, clathrin light chain can rescue the mobility of membrane patches containing sla2p, the yeast homolog of hip1 and hip1R (Newpher, Idrissi et al. 2006). Regulation of sla2p's actin binding by clathrin light chain was proposed to mediate this rescue (Newpher, Idrissi et al. 2006). Intramolecular regulation of actin binding by the hip1 and hip1R THATCH domains has also been shown (Senetar, Foster et al. 2004; Brett, Legendre-Guillemin et al. 2006). Here we address how intramolecular mechanisms regulating actin binding by Hip proteins are related to the influence of clathrin light chains on Hip-actin interactions.

Many cellular functions of hip1 and hip1R are linked to their coiled coil domains. Dimerization occurs through the coiled coil as does binding to clathrin light chain (4,3,(Ybe, Mishra et al. 2007). The hip1 coiled coil domain also binds huntingtin and has been implicated in binding to androgen receptor (Kalchman, Koide et al. 1997; Wanker, Rovira et al. 1997; Mills, Gaughan et al. 2005). Coiled coils have a canonical heptad repeat that allows the coils to wrap around each other and maintain a hydrophobic core (Burkhard, Stetefeld et al. 2001). Changes within the heptad repeat of coiled coils often signify a region of flexibility where protein partners bind and can potentiate conformational change (Brown, Cohen et al. 1996; Burkhard, Stetefeld et al. 2001),(Singh and Hitchcock-DeGregori 2003; Singh and Hitchcock-DeGregori 2006). Hip1 and hip1R coiled coils have high similarity to each other at the primary sequence level and have virtually identical regions important for clathrin light chain binding (Legendre-Guillemin, Metzler et al. 2005; Ybe, Mishra et al. 2007). The structure of fragments of hip1's coiled coil showed a shift in the heptad repeat in the region where a hydrophobic patch predicted to bind clathrin light chain was located (Ybe, Mishra et al. 2007; Niu and Ybe 2008). Also these structures showed a split in the coiled coil, suggesting increased flexibility in the clathrin light chain-binding region (Ybe, Mishra et al. 2007),(Niu and Ybe 2008), with potential for structural changes upon ligand binding.

To understand the mechanism of their functional importance, the biophysical properties of the coiled coil domains of hip1 and hip1R were quantitatively analyzed. These domains showed strong homodimerization, regions of flexibility, and the ability to induce conformational changes upon clathrin light chain binding, which dramatically reduced actin affinity for the THATCH domains. These biophysical properties of Hip proteins provide a basis for understanding hip1 and hip1R functions in cells and for understanding the regulatory effect of clathrin light chain on Hip protein-actin interaction in endocytosis. Our findings further suggest a model for the sequential function of clathrin and actin in coated vesicle budding.

Experimental Procedures

Expression constructs and peptides- Separate domains of either human hip1 or mouse hip1R were cloned. The coiled coil of hip1 (hip1cc) corresponding to amino acids 361-637, hip1R coiled coil (hip1Rcc) corresponding to amino acids 346-655, hip1 coiled coil and THATCH corresponding to amino acid 400-1038 (hip1ccth), hip1R coiled coil and THATCH corresponding to amino acids 379-1068 (hip1Rccth), hip1 THATCH corresponding to amino acid 600-1038 (hip1th), and hip1R THATCH corresponding to amino acids 600-1068 (hip1Rth), were cloned via PCR amplification of gene fragments with flanking Gateway (Invitrogen) recombination sequences and inserted into pRSF-duet (Novagen) or pET-duet plasmid (Novagen) modified with Gateway sequences. Full length hip1 or hip1R was PCR amplified with a primers containing either the His tag or HA sequence and cloned into pCDNA3.1 plasmid (Invitrogen). Clathrin light chain LCb was expressed from a previously described construct (Liu, Wong et al. 1995). Cortactin SH3 domain was expressed and purified as previously described (17).

A peptide representing the consensus sequence of clathrin light chains (EEDPAAAFLAQQESEIAGIEND) was synthesized commercially. Control peptides of the same length were either the clathrin light chain peptide with three point mutations (bold) that abrogate hip1R binding previously described (Chen and Brodsky 2005) (EE**V**PAAAFLAQQESE**A**AGI**A**ND) or unrelated peptide (FINKPETGAVELESPFILLADKKI) derived from the bacterial protein GroEL.

Protein expression and purification- For all proteins, BL21 (DE3) *E. coli* were transformed with the appropriate plasmid and expression was induced via standard methods. Purification was via standard methods using Talon affinity resin (all Hip protein constructs) or Glutathione Sepharose (GST-cortactin). All Hip protein constructs were further purified over a Superdex 200 column. See supplementary Materials and Methods for details.

Analytical ultracentrifugation- Hip1cc or hip1Rcc were placed in a six channel cell at 1 mg/ml. Concentration filtrate was used for reference cells. Protein was centrifuged at four speeds, 8000, 10000, 15000, 20000 rpms in an AN Ti60 rotor in a Beckman Optima XL-A analytical ultracentrifuge. Data was collected after 8 hours spinning to equilibrium via absorption optics and analyzed using Ultrascan 9.0 (Demeler 2005). Equilibrium data was globally fitted to a monomer-dimer equilibrium or single species of either the calculated monomer or dimer size.

Circular dichroism- Circular dichroism (CD) spectra were collected using a Jasco J-710 spectrapolarimeter. Hip1cc and hip1Rcc were diluted in 20mM pH 7.0 phosphate 150mM NaCl to a concentration of 0.035 mg/ml and 0.036 mg/ml, respectively. Indvidual melting analysis for hip1cc and hip1Rcc was performed starting at 22°C and increasing to 80^{o}C while monitoring ellipticity at 222 nm. Rate of heating was 0.5°C/min.

For melting of the mixture of hip1cc and hip1Rcc the rate of heating was 1°C/min. Annealing time for mixture of hip1cc and hip1Rcc was between 45 minutes and 1 hour, until the ellipticity at 222nm was back to starting baseline. Data was collected and analyzed to 65°C, well past the major melting transitions of both hip1cc and hip1Rcc which should include dimer unfolding. The data was plotted as the mean residue molar ellipticity.

Surface plasmon resonance- Clathrin light chain surface was prepared by immobilizing ~400RU's of protein via amine coupling. For binding experiments, either hip1cc, hip1Rcc, or hip1ccL506A was flowed at concentrations indicted in Fig. 5 with pH 6.7 MES running buffer. Actin surface was prepared by immobilizing ~1000RU's of rabbit muscle F-actin to the chip surface via amine coupling. Hip1ccth or hip1Rccth at indicated concentrations (Fig. 5) and with or without 20mM clathrin light chain peptide or control peptide was flowed over the surface in 20mM Tris pH 8.0, 1mM MgCl2, 150mM KCl running buffer.

GST-cortactin was immobilized with an anti-GST antibody (Biacore) and hip1Rccth at indicated concentrations was flowed over (Fig. S5). Running buffer was HBS-EP (Biacore). All

data was collected on Biacore T100 at 25°C. Steady state binding analysis was performed using BiaEvaluation software (Biacore).

Partial proteolysis- 80 mL of 0.2 mg/ml HIP1 or HIP1R in gel filtration buffer was mixed with 0.2 mg of pure subtilisin (Boehringer Mannheim) in 20mM HEPES pH 7.2, 200mM NaCl. Incubation times are as stated in the figures. At each time point 20µl of protein was taken and the proteolysis reaction stopped by mixing with SDS gel loading buffer and boiling. Proteolysis products were resolved on 12% SDS NuPAGE gel in MES running buffer (Invitrogen).

Mass spectrometry- Spectra of tryptic fragments excised from gels were taken on a Voyager-DE STR (Applied Biosystems) in reflector mode. See supplementary Materials and Methods for details. Spectra were processed and analyzed using the Data Explorer program (Applied Biosystems). Peptide peaks were identified using MS-Fit and MS-Digest within the Protein Prospector suite.

Immunoprecipitation and immunoblotting- HeLa cells were co-transfected with tagged full length hip1 and hip1R using Lipofectamine 2000 (Invitrogen). DNA concentrations used for transfection were adjusted to yield equal levels of tagged protein expression (Fig. 2) or fourfold excess of one tagged protein over the other (Fig. S1). Immunoprecipitation (5µg of each antibody) and immunoblotting were with either 6x His monoclonal antibody (BD biosciences) or HA.11 monoclonal antibody (Covance) using standard methods. For details see supplementary Materials and Methods.

Immunofluorescence- HeLa cells cultured in DMEM containing 10% bovine growth serum (Hyclone), were transfected with a final concentration of 10nM siRNA targeted to clathrin light chain or control using the HiPerfect reagent (Qiagen). Target sequences of LCa and LCb knockdowns were as described (Huang, Khvorova et al. 2004), and both control and the clathrin light chain siRNAs were synthesized by Qiagen. Cells were collected at 72 hours post-siRNA treatment, fixed and stained using rabbit hip1R antiserum (Millpore), monoclonal

antibody to HC (X22), and Alexa Fluor conjugated phalloidin (Invitrogen) to F-actin. The samples were viewed by confocal laser scanning microscopy using a Leica TCS-SP5 operating system.

Electron Microscopy- 5µl of 0.2µM purified hip1ccth or hip1Rccth incubated with 20µM control peptide, clathrin light chain peptide or no peptide was dropped on a carbon-coated grid and stained with uranyl formate. Samples were imaged at 67,000x using a Philips Tecnai T20 electron microscope operated at an acceleration voltage of 120 kV with a defocus range of 1.2 to 1.5 µM. Images were taken using a 4Kx4K CCD (GATAN) with a pixel size of 2.25 Å/pixel. Images were selected and measured using EMAN.

ANS fluorescence- 0.7µM purified Hip coiled coils or coiled coil THATCH domains were incubated with 20µM clathrin light chain peptide or control peptide and 100µM 8-anilino-1-napthalenesulfonic acid (ANS). Excitation of ANS was at 360 nm and emission was scanned from 400 nm to 600 nm. Raw data was averaged over 3 independent experiments and standard errors calculated for each condition. Difference spectra of Hip coiled coil THATCH and Hip coiled coils were plotted with propagated standard error.

RESULTS

Hip1 and Hip1R are stable homodimers.

Homodimers of hip1R have an elongated dumbbell shape, with the N and C terminal globular domains in variable positions (Engqvist-Goldstein, Warren et al. 2001). This implies the primary dimerization determinants are within the coiled coil domain. Interactions between hip1 and hip1R have been shown by GST-pulldown and co-immunoprecipitation (Legendre-Guillemin, Metzler et al. 2002). Whether this interaction arises from direct heterodimerization or another mode of interaction has yet to be explicitly determined and would have different functional implications. Coiled coil domains (hip1 residues 361 to 637 and hip1R residues 346 to 655, Fig.

1) from Hip proteins were expressed and purified and their dimerization properties were analyzed by equilibrium sedimentation analytical ultracentrifugation (Table I). Assuming a monomer-dimer equilibrium model, the data analysis suggested that there was no equilibrium under standard conditions and all of the coiled coils for both hip1 and hip1R are dimeric in solution. The fitted molecular weights were 67.4 and 67.7 kDa for the hip1 and hip1R coiled coils, respectively, corresponding closely to calculated homodimeric sizes of 66.79 kDa, and 73.45 kDa (Table I). Fitting the sedimentation data to a single species with the calculated molecular weight of the dimers (data not shown) also fit well with the experimental data. Attempts to fit the sedimentation data to the true calculated size of the monomer, as either a single species or as a monomer-dimer equilibrium, resulted in poor fits (data not shown). These data indicate that homodimers of hip1 and hip1R are stable and that there is virtually no dissociation to the monomer form of these proteins *in vitro*.

This stability was also evident from biochemical treatments. Dimers of purified hip1 and hip1R coiled coils survived standard denaturing, reducing SDS-PAGE gels (Fig. 2A). Coiled coil domains of hip1 and hip1R were mixed with SDS-PAGE sample buffer and heated to 95^{o}C for increasing time intervals (Fig. 2A). After only short heating times large amounts of hip1 or hip1R coiled coils remained complexed. This SDS resistant behavior is similar to that of the four-helix bundles of SNARE proteins that require an ATP-hydrolyzing chaperone, N-ethylmaleimide sensitive factor (NSF) to dissociate the complex (Chen, Scales et al. 1999). While both hip1 and hip1R coiled coils were very stable, hip1 remained a dimer for longer heating times (Fig. 2A).

Given the high stability of the coiled coil domains of hip1 and hip1R, it is likely that homodimers are the preferred mode of interaction. To test the possibility of heterodimerization in a stringent way, a 1:1 mixture of hip1 and hip1R coiled coils was thermally denatured through their primary melting transition to 65^{o}C, while monitoring the melting transition by circular dichroism (Fig. 2B). The denatured mixture was then allowed to slowly re-anneal. Melting and

annealing should overcome any thermodynamic or kinetic barrier to monomer formation and allow heterodimers to form during re-annealing, if feasible. Following re-annealing, thermal denaturation of the resulting molecules was measured and the melting curve was identical that of the mixture of the hip1 and hip1R coiled coil homodimers (Fig. 2B). Thus, no new stable heterodimers had formed upon re-annealing. These would have generated a melting curve that deviated from that of the original mixture because the heterodimers would contribute differently to the shape of the melting curve.

These biophysical results indicate there is no heterodimerization of hip1 and hip1R through their coiled coil domains *in vitro*. To follow up the biophysical analysis in a cellular context, four constructs encoding full length HA- or His-tagged hip1 and hip1R were produced and expressed in HeLa cells (Fig. 2C). After co-transfection of hip1 and hip1R, each with a different tag (in either combination), immunoprecipitation of either hip1 or hip1R gave negligible amounts of the other protein. However, when hip1 constructs with two different tags or hip1R constructs with two different tags were transfected into cells, homotypic interactions were readily detected by immunoprecipitation. Varying the expression level of each construct relative to the other did not change this result (Fig. S1). Our biochemical analysis and its correlation with cellular data thus suggest that the preferred mode of Hip protein interaction *in vivo* and *in vitro* is homodimeric. These results are compatible with reported functional segregation of hip1 and hip1R in cells.

Flexibility is a mechanism for conformational regulation in hip1 and hip1R. Given the preference for homodimerization and the segregation of biological functions of hip1 and hip1R, further biophysical properties of their coiled coil domains were analyzed to determine whether these might account for functional differences. Separate crystallographic analyses of the N- and C-terminal halves of the coiled coil domain of hip1 showed classical coiled coil segments from residues 371-430 and from residues 540 to 581. These structures also suggested a segment, around residues 445 to 539, of decreased stability and increased flexibility that included residues implicated in binding clathrin light chain, as well as in binding of hip1 protein interactor (HIPPI)

(Ybe, Mishra et al. 2007) (Fig. 1). To investigate whether this region might function in regulatory aspects of the Hip proteins, biophysical properties related to stability of the Hip protein coiled coil domains were analyzed. First, using the program COILS (Lupas, Van Dyke et al. 1991), subdomains of high and low coiled coil propensity were mapped within the coiled coil domains of both hip1 and hip1R (Fig. 3A and 3B). COILS suggested the N-terminal region of both domains have high propensity to form coils. At the D/ELLRKN sequences implicated in binding clathrin light chains (Legendre-Guillemin, Metzler et al. 2005), the predicted propensity dropped for both hip1 and hip1R. However, for hip1 this decrease was slight, while for hip1R there was a significant decrease in predicted coil propensity. C-terminal to the clathrin light chain binding sequence both hip1 and hip1R had predicted regions of significantly lower coiled coil propensity interspersed with regions of high coiled coil propensity, again with hip1R having less coil propensity.

Partial proteolysis was used to establish the regions of stability and flexibility in hip1 and hip1R. The coiled coil domains of hip1 and hip1R were partially digested with subtilisin over time and analyzed by SDS PAGE. Hip1 was resistant to proteolysis for approximately 10-15 minutes while hip1R was almost completely digested in the same time period (Fig. 3C). The coiled coil domains of both proteins had sub-fragments that were resistant to proteolysis. Peptide mass fingerprints from the protease-resistant fragments of hip1 (11 kDa and 20 kDa) and hip1R (10 kDa and 13 kDa) showed that peptides C-terminal to the clathrin light chain binding region were missing from these fragments (Fig. 3D). The increased susceptibility to degradation of the C-terminal subdomains of the hip1 and hip1R coiled coils agrees with the COILS analysis predicting decreased coiled coil propensity in the same region. Thus, while the static crystal structure indicated coiled coil interactions can occur in this region, it is functionally less stable (more flexible) than the coiled coil region that is N-terminal to the clathrin light chain binding site. The differences in the relative rates of partial proteolysis indicate that hip1R has increased flexibility in the C-terminal subdomain compared to hip1.

The relative stability of hip1 and hip1R coiled coils was determined by reversible thermal denaturation while measuring the ellipticity at 222 nm by circular dichroism (Fig. 4). Thermal melting curves for hip1 and hip1R showed significantly different profiles for unfolding. Hip1 had a single smooth melting transition (Tm) at 43°C while hip1R had a multi-step unfolding profile. The first melting transition for hip1R occurred at 32°C much lower than that for hip1; the second hip1R Tm was much higher, near 72°C. The first melting transition for hip1R likely represents most of the protein unfolding given its significant change. The second transition occurs at high temperature and may reflect residual bits of helical structure unfolding or temperature dependent changes in absorption, not related to hip1R denaturation. These melting data correlate with the partial proteolysis, gel denaturation, and COILS data suggesting that hip1 is more stable than hip1R.

Since the flexible region started at the clathrin light chain binding site, we measured whether the differences in flexibility of hip1 and hip1R correlated with their binding affinity for clathrin light chain. Using surface plasmon resonance (SPR) the binding affinity between hip1 or hip1R and full length clathrin light chain b (LCb) was determined (Fig. 5A). Hip1 bound LCb with an affinity of 1.02 μM. Hip1R bound LCb with an affinity of 0.566 μM, about two fold better than hip1, correlating with the greater flexibility of hip1R in the LCb binding region. The D/ELLRKN residues known to affect light chain binding are conserved between hip1 and hip1R with the single difference of an aspartate to glutamate change (Fig. 1). The small hydrophobic patch that was predicted to form a clathrin light chain binding site (Ybe, Mishra et al. 2007) is also conserved between the two proteins. Since sequence determinants for clathrin light chain binding to hip1 and hip1R are virtually identical, other biophysical properties of the coiled coils must contribute to differentially modulate the binding. To test this hypothesis, a single point mutant of the hip1 coiled coil domain (L506A) was made. Leucine 506 makes up part of the hydrophobic core of the hip1 coiled coil, near the predicted clathrin light chain binding site, so

the mutation should decrease the hydrophobic interactions stabilizing the core and increase local flexibility. Partial proteolysis of hip1 L506A indicated decreased stability compared to wild type hip1 coiled coils (data not shown). SPR measurements showed hip1 L506A has a two fold increased affinity for LCb (0.490 μM) compared to the wild type hip1 coiled coil domain (Fig. 5A). These binding affinity measurement again correlate with flexibility in the coiled coil domains of the Hip proteins. The increased flexibility of hip1R and the hip1 L506A mutant may allow it to more frequently sample a conformation amenable to binding clathrin light chain.

Hip1 and hip1R adopt compact conformations in response to clathrin binding. Since flexibility of the coiled coil domains influenced clathrin light chain interaction, we addressed whether their flexibility might influence additional functions of hip1 and hip1R. For the yeast homolog sla2p, sequences in the vicinity of the coiled coil domain were shown by two-hybrid and in vitro binding assays to have the capacity for intramolecular interaction with the actin-binding THATCH domain (Yang, Cope et al. 1999). The flexibility of the coiled coils starting at the clathrin light chain binding region could provide a means to induce this intramolecular interaction, which would affect the surface properties of the THATCH domain. To investigate this possibility we assessed changes in the surface properties of the THATCH domains by monitoring fluorescence of the hydrophobic dye 8-anilino-1-napthalenesulfonic acid (ANS) (Fig. 6A). The C-terminal fragments of hip1 and hip1R including coiled coil and THATCH domain, (hip1 residues 400-1038, hip1ccth; or hip1R residues 379-1068, hip1Rccth) were expressed and ANS fluorescence spectra were collected in the presence of clathrin light chain peptide or control peptide. The clathrin light chain peptide is a 22-residue peptide representing the Hip-binding region of clathrin light chains (common to LCa and LCb) (see Methods for sequence) (4). ANS spectra in the presence of the two peptides were also collected for the coiled coil domains alone. Coiled coil spectra were subtracted from the coiled coil THATCH spectra to get the net change in ANS fluorescence, reflecting changes due to conformational rearrangement of the THATCH domain alone. Hip1R showed a significant change in the ANS fluorescence difference spectrum

for the THATCH domain in the presence of the clathrin light chain peptide compared to the control peptide. The hip1 THATCH domain difference spectrum had a similar trend but was not above the error (Fig 6A). These assays reveal a change in the accessible hydrophobic surfaces in the THATCH domain when clathrin light chains are bound to coiled coil domains, suggesting an intramolecular conformational rearrangement in mammalian Hip proteins similar to that predicted for Sla2p (32). The limited change detected in hip1 is likely due ANS being insensitive to the conformational changes induced by the clathrin light chain peptide since electron microscopy data (see below) clearly show altered conformations for both hip1 and hip1R.

To visualize the conformational changes in hip1 and hip1R that were implied by ANS fluorescence, negative stain electron microscopy (EM) was used. Hip1 and hip1R coiled coil THATCH domains (hip1ccth and hip1Rccth) were combined with clathrin light chain peptide or control peptide, stained with uranyl formate, and EM images collected. The control peptide was the clathrin light chain peptide containing three mutations that abrogate Hip protein binding (4). To avoid bias, random single particles were pre-selected from fields without knowing which peptide had been added, and the longest dimension of these particles was measured. Measurements were binned in groups of 20Å and plotted as a histogram (Fig. 6B and 6C). The length of single particles of both hip1 and hip1R with clathrin light chain peptide bound had a narrower distribution than seen with control peptide and peaked near 120 Å. In the presence of control peptide (and no peptide for hip1R) both proteins showed a much broader length distribution and many more particles with dimensions greater than 120 Å. This indicates clathrin light chain binding in the coiled coil can induce compaction of hip1 and hip1R (Fig. 6B and S2).

Actin binding by hip1 and hip1R is negatively regulated by clathrin binding. Our studies here have shown how clathrin light chain can induce a compact conformation of Hip proteins through binding the flexible regions of their coiled coil domains and also influence the surface properties of their actin-binding THATCH domains. These observations suggest that clathrin light chain binding to the coiled coil domain of Hip proteins could affect the binding of actin to

their THATCH domains. To assess this possibility, SPR was used to determine the binding affinities of hip1 and hip1R for actin in the presence of clathrin light chain peptide or control peptide. For these experiments, hip1ccth or hip1Rccth were flowed over a chip containing immobilized F-actin. The steady state affinity of hip1ccth binding to F-actin was calculated to be 7.66 μM and for hip1R it was 1.03 μM in the presence of control peptide (Fig. 5B). These affinities were very similar to the affinity of Hip proteins for actin with no peptide present (Fig. S3). In agreement with our results, it has been previously reported that hip1 binds actin with a relatively weaker affinity than hip1R, but no specific binding constants were calculated in that study (Legendre-Guillemin, Metzler et al. 2002). Also both affinities calculated from our data are in the micromolar range as previously reported for binding between actin and the isolated THATCH domains of hip1 and hip1R (lacking their regulatory upstream helix) (Senetar, Foster et al. 2004).

Clathrin light chain peptide or control peptide was then introduced into the SPR assay for measuring Hip-actin interaction. In the presence of saturating amounts of clathrin light chain peptide, the determined K_{Dactin} for actin binding was 45 μM for the hip1ccth-peptide complex (Fig. 5B) and >1000 μM for the hip1Rccth-peptide complex (Fig. 5B). Its greater response to clathrin light chain peptide correlates with hip1R being the more flexible of the two Hip proteins. Notably, there was an approximately 7 to >1000 fold decrease in affinity of both hip1 and hip1R for actin when the clathrin light chain peptide was present. Consistent with this, hip1ccth prebound to actin had lower affinity for clathrin light chain, though the change in affinity was less pronounced in this binding sequence (Fig. S6). Purified THATCH domains missing the clathrin-binding coiled coil domains (residue 600-1038 for hip1 and residue 600-1068 for hip1R) had little change in actin-binding affinity in the presence of clathrin light chain peptide or control peptide (Fig. S4). Thus, the clathrin light chain effects on actin binding by Hip proteins are the result of long-range conformational interactions between the coiled coil and THATCH domains. The proline rich domain of Hip1R, found C-terminal to the THATCH domain has been previously

shown to bind cortactin (Le Clainche, Pauly et al. 2007). SPR experiments showed that the presence of clathrin light chain did not alter the affinity hip1Rccth for cortactin (Fig. S5), demonstrating specific regulation of hip1R actin affinity by clathrin light chain. As expected, the equivalent domain of hip1, which does not contain a proline rich domain, did not bind to cortactin (data not shown).

In mammalian cells hip1R primarily colocalizes with punctate clathrin structures that align with but do not overlap with actin (Fig. 7). Knockdown of clathrin light chain shows an increase in colocalization of hip1R with actin and an increase of hip1R containing clathrin structures localizating with actin (Fig 7). Studies in yeast and mammalian cells indicated that clathrin light chain fragments comprising only the Hip-binding or sla2p-binding regions affected actin-dependent membrane dynamics (4(Newpher, Idrissi et al. 2006). Our results explain this effect by demonstrating the profound influence of the clathrin light chain peptide on the Hip proteins' affinity for actin. Consistent with the cellular assays showing functional efficacy of Hip-binding or Sla2p-binding domains, the clathrin light chain peptide used in our studies binds to Hip proteins with similar efficiency as full length purified clathrin light chain (data not shown) (4,18).

DISCUSSION

Hip1 and hip1R promote clathrin assembly through binding to clathrin light chains (Chen and Brodsky 2005; Legendre-Guillemin, Metzler et al. 2005) and hip1R controls assembly of actin through its interaction with cortactin and actin (Brett, Legendre-Guillemin et al. 2006; Le Clainche, Pauly et al. 2007). Hip1 is also an actin-binding protein (Senetar, Foster et al. 2004). In addition to their common properties, hip1 and hip1R seem to have distinct roles within cells, because separate phenotypes result from deletion of the genes encoding either hip1 or hip1R (Bradley, Hyun et al. 2007). Here we demonstrate that homodimerization is the predominant state for the Hip proteins, which allows segregation of their cellular functions. We characterize

further biophysical differences between the coiled coil domains of hip1 and hip1R, but demonstrate that both contain flexible regions that contribute to their common functions. In particular, clathrin light chain binding to their coiled coil regions is shown to change the conformation of both Hip proteins and to reduce their affinity for actin. These observations suggest Hip proteins have separate interactions with clathrin and actin during membrane traffic, rather than the direct bridging function between clathrin and the cytoskeleton proposed previously (Engqvist-Goldstein, Warren et al. 2001; Engqvist-Goldstein, Zhang et al. 2004). They also suggest that clathrin light chain partially rescues endocytosis in yeast, in the absence of clathrin heavy chain, by releasing actin from sla2p to promote asymmetric attachment of actin to the budding vesicle (Newpher, Idrissi et al. 2006).

Differential localization and binding partners suggest hip1 and hip1R have varied and separate functions. Here we show strong tendency for homodimerization of the hip1 and hip1R coiled coil domains in vitro that would prevent the mixing of their separate functions. By co-immunoprecipitation experiments we confirmed preference for homotypic interaction of the full length proteins in cells. It is possible that earlier reports of hip1-hip1R heterodimerization and the very minor (unidirectional) heterotypic interaction in cells that we detect reflect an interaction between hip1 and hip1R that occurs under special conditions of cellular regulation. However, our data establish homodimers as the baseline state for hip1 and hip1R in cells. Consistent with their homodimerization and distinct binding properties, hip1 and hip1R can only partially compensate for each other in genetic knock-down experiments (Bradley, Hyun et al. 2007).

Although strong dimerization is clearly a property of the coiled coil regions of the Hip proteins, there have been suggestions of flexibility within the coiled coil regions from EM and crystallography studies (Engqvist-Goldstein, Warren et al. 2001; Ybe, Mishra et al. 2007; Niu and Ybe 2008). Here we characterized the extent of flexibility within the coiled coil regions of Hip proteins in solution and addressed its function, following up the hypothesis that flexibility might directly modulate binding affinities for partner proteins. First, we showed that differences in

clathrin light chain affinity for hip and hip1R correlated with the extent of the flexible regions in their coiled coil domains. The high sequence identity in the clathrin light chain binding regions predicts little difference in the binding affinities of hip1 and hip1R for clathrin light chain. Nonetheless, the binding constants reported here indicate that hip1 and hip1R exhibit a difference in affinity for clathrin light chain, consistent with previous qualitative predictions about their relative affinities for clathrin light chain (Engqvist-Goldstein, Warren et al. 2001; Chen and Brodsky 2005). The lower melting temperature and higher susceptibility to proteolysis of the hip1R coiled coil domain compared to that of hip1 indicates greater flexibility or conformational entropy, which could explain hip1R's stronger binding of light chain. We propose that intrinsic flexibility modulates binding affinity by influencing the frequency the coiled coil is in the clathrin light chain binding-competent state.

We further investigated whether coiled coil flexibility and clathrin light chain binding play a role in actin binding by Hip proteins because it had been previously suggested that the THATCH domain might be subject to intramolecular regulation (18,19,32). Earlier studies of actin binding by Hip proteins focused on isolated THATCH domains and identified elements regulating actin binding. Senetar *et al.* observed that the actin binding affinity of the Hip proteins' THATCH domain was considerably reduced in the presence of an N-terminal upstream helix (USH) (Senetar, Foster et al. 2004). In studies by Brett *et al.* a C-terminal LATCH sequence was found to increase hip1R THATCH binding affinity for actin, in the presence of the USH (Brett, Legendre-Guillemin et al. 2006). Here we measured actin binding of extended domains of hip1 and hip1R, fragments comprising the coiled coil and THATCH domains with both the USH and the LATCH sequences. In the absence of the clathrin light chain peptide, hip1 and hip1R bound actin in the micromolar range, as though the inhibitory USH were inactive. In the presence of the light chain peptide, the K_{Dactin} of both proteins was significantly decreased, with a 6.4-fold change for hip1 and a >1000 fold change for hip1R. Thus the binding of clathrin light chain to the coiled coil of Hip proteins seems to reduce the affinity of both proteins for actin,

with a greater effect on the more flexible hip1R. These effects of clathrin light chain peptide on actin binding by the Hip proteins correlated with conformational changes induced upon peptide binding. Both hip proteins are able to adopt compact structures as well as extended structures. Upon clathrin light chain binding, both proteins assumed a more compact state. That these conformational changes involved the THATCH domain was confirmed by ANS spectra, which showed a change in THATCH surface contacts upon binding of clathrin light chain peptide to hip1R, with the same tendency for hip1. These data suggest that in the compact state, the Hip protein actin binding sites are masked by an intramolecular interaction. The compact conformation may also involve repositioning of the USH (Fig. 7A). In cells depleted of clathrin light chain hip1R has increased colocalization with actin, consistent with hip1R being in the extended state with higher affinity to actin. In cells treated with control siRNA hip1R, likely in the compact state with low affinity for actin, is primarily colocalized with clathrin containing structures and segregated from actin. This indicates hip1R interactions with clathrin and actin are mutally exclusive. It remains to be seen whether binding of other partner proteins to the hip1 coiled coil affect its coiled coil properties and whether this in turn influences other binding partner interactions. In contrast to the multiple partners of hip1, clathrin light chain is the only binding partner identified so far for the hip1R coiled coil region and suggests hip1R primarily functions in clathrin mediated endocytosis. Notably, clathrin light chain binding did not affect cortactin binding, which interacts with hip1R C-terminal to the THATCH domain. In addition to demonstrating the specificity of clathrin light chain binding in regulating actin binding, this result indicates that cortactin should remain bound to hip1R whether or not clathrin is bound. The latter observation has implications for the influence of clathrin and Hip proteins on actin organization during coated vesicle formation.

Our data indicate that clathrin light chain binding to hip1 and hip1R regulates their ability to interact with actin and explains how the N-terminal fragment of clathrin light chain could affect the mobility of sla2p patches in yeast as shown by Newpher *et al.* According to our

findings, binding of hip1 and hip1R to actin would be reduced when they are bound to clathrin lattices. Thus, Hip interactions with clathrin and actin may be sequential (Fig. 7B). As the clathrin-coat forms around a budding vesicle, it can capture Hip proteins without being encumbered by Hip-actin interactions, which are weaker when Hips are bound to clathrin. Consistent with this prediction, electron micrographs localizing Hip proteins relative to clathrin and actin show that clathrin-associated Hip does not co-localize with actin, except at the neck of the budding vesicle, near the clathrin lattice edge (Engqvist-Goldstein, Warren et al. 2001). At the neck, Hip proteins can interact directly with membrane via their ANTH domains and dissociate from clathrin light chains, enhancing their affinity for actin and thereby regulating actin filament growth at this site. As suggested by Le Clainche *et al.*, the binding of cortactin by hip1R could control the localization of actin polymerization at the budding vesicle neck and contribute to vesicle scission. Since cortactin binding is not affected by clathrin light chain, it could be constitutively bound to lattice-associated hip1R. In yeast, the Sla2p-binding region of clathrin light chain could facilitate endocytosis in the absence of clathrin heavy chain by stimulating release of actin from sla2p bound to the membrane.

In summary, our findings show that while overall stability of the coiled coil domains of Hip proteins favors homodimerization and segregation of hip1 and hip1R function, their intrinsic flexibility allows regulation of partner protein binding through conformational changes. In particular these changes mandate sequential interaction of Hip proteins with clathrin and actin, re-defining the proposed mechanism of action of Hip proteins during membrane traffic.

REFERENCES

Acton, S. L., D. H. Wong, et al. (1993). "Alteration of clathrin light chain expression by transfection and gene disruption." Mol Biol Cell **4**(6): 647-60.

Bradley, S. V., T. S. Hyun, et al. (2007). "Degenerative phenotypes caused by the combined deficiency of murine HIP1 and HIP1r are rescued by human HIP1." Hum Mol Genet **16**(11): 1279-92.

Brett, T. J., V. Legendre-Guillemin, et al. (2006). "Structural definition of the F-actin-binding THATCH domain from HIP1R." Nat Struct Mol Biol **13**(2): 121-30.

Brodsky, F. M., C. Y. Chen, et al. (2001). "Biological basket weaving: formation and function of clathrin-coated vesicles." Annu Rev Cell Dev Biol **17**: 517-68.

Brown, J. H., C. Cohen, et al. (1996). "Heptad breaks in alpha-helical coiled coils: stutters and stammers." Proteins **26**(2): 134-45.

Burkhard, P., J. Stetefeld, et al. (2001). "Coiled coils: a highly versatile protein folding motif." Trends Cell Biol **11**(2): 82-8.

Chen, C. Y. and F. M. Brodsky (2005). "Huntingtin-interacting protein 1 (Hip1) and Hip1-related protein (Hip1R) bind the conserved sequence of clathrin light chains and thereby influence clathrin assembly in vitro and actin distribution in vivo." J Biol Chem **280**(7): 6109-17.

Chen, Y. A., S. J. Scales, et al. (1999). "SNARE complex formation is triggered by Ca2+ and drives membrane fusion." Cell **97**(2): 165-74.

Demeler, B. (2005). Ultrascan a comprehensive data analysis software package of analytical ultracentrifugation experiments. . Modern Analytical

Ultracentrifugation: Techniques and Methods. S. E. H. D.J. Scott, A.J. Rowe, Royal Society of Chemistry**:** 210-229.

Engqvist-Goldstein, A. E. and D. G. Drubin (2003). "Actin assembly and endocytosis: from yeast to mammals." Annu Rev Cell Dev Biol **19**: 287-332.

Engqvist-Goldstein, A. E., M. M. Kessels, et al. (1999). "An actin-binding protein of the Sla2/Huntingtin interacting protein 1 family is a novel component of clathrin-coated pits and vesicles." J Cell Biol **147**(7): 1503-18.

Engqvist-Goldstein, A. E., R. A. Warren, et al. (2001). "The actin-binding protein Hip1R associates with clathrin during early stages of endocytosis and promotes clathrin assembly in vitro." J Cell Biol **154**(6): 1209-23.

Engqvist-Goldstein, A. E., C. X. Zhang, et al. (2004). "RNAi-mediated Hip1R silencing results in stable association between the endocytic machinery and the actin assembly machinery." Mol Biol Cell **15**(4): 1666-79.

Gervais, F. G., R. Singaraja, et al. (2002). "Recruitment and activation of caspase-8 by the Huntingtin-interacting protein Hip-1 and a novel partner Hippi." Nat Cell Biol **4**(2): 95-105.

Hackam, A. S., A. S. Yassa, et al. (2000). "Huntingtin interacting protein 1 induces apoptosis via a novel caspase-dependent death effector domain." J Biol Chem **275**(52): 41299-308.

Huang, F., A. Khvorova, et al. (2004). "Analysis of clathrin-mediated endocytosis of epidermal growth factor receptor by RNA interference." J Biol Chem **279**(16): 16657-61.

Kalchman, M. A., H. B. Koide, et al. (1997). "HIP1, a human homologue of S. cerevisiae Sla2p, interacts with membrane-associated huntingtin in the brain." Nat Genet **16**(1): 44-53.

Kirchhausen, T. (2000). "Clathrin." Annu Rev Biochem **69**: 699-727.

Le Clainche, C., B. S. Pauly, et al. (2007). "A Hip1R-cortactin complex negatively regulates actin assembly associated with endocytosis." Embo J **26**(5): 1199-210.

Legendre-Guillemin, V., M. Metzler, et al. (2002). "HIP1 and HIP12 display differential binding to F-actin, AP2, and clathrin. Identification of a novel interaction with clathrin light chain." J Biol Chem **277**(22): 19897-904.

Legendre-Guillemin, V., M. Metzler, et al. (2005). "Huntingtin interacting protein 1 (HIP1) regulates clathrin assembly through direct binding to the regulatory region of the clathrin light chain." J Biol Chem **280**(7): 6101-8.

Liu, S. H., M. L. Wong, et al. (1995). "Regulation of clathrin assembly and trimerization defined using recombinant triskelion hubs." Cell **83**(2): 257-67.

Lupas, A., M. Van Dyke, et al. (1991). "Predicting coiled coils from protein sequences." Science **252**(5010): 1162-4.

Metzler, M., B. Li, et al. (2003). "Disruption of the endocytic protein HIP1 results in neurological deficits and decreased AMPA receptor trafficking." Embo J **22**(13): 3254-66.

Mills, I. G., L. Gaughan, et al. (2005). "Huntingtin interacting protein 1 modulates the transcriptional activity of nuclear hormone receptors." J Cell Biol **170**(2): 191-200.

Newpher, T. M., F. Z. Idrissi, et al. (2006). "Novel function of clathrin light chain in promoting endocytic vesicle formation." Mol Biol Cell **17**(10): 4343-52.

Niu, Q. and J. A. Ybe (2008). "Crystal structure at 2.8 A of Huntingtin-interacting protein 1 (HIP1) coiled-coil domain reveals a charged surface suitable for HIP1 protein interactor (HIPPI)." J Mol Biol **375**(5): 1197-205.

Pauly, B. S. and D. G. Drubin (2007). "Clathrin: an amazing multifunctional dreamcoat?" Cell Host Microbe **2**(5): 288-90.

Schafer, D. A., S. A. Weed, et al. (2002). "Dynamin2 and cortactin regulate actin assembly and filament organization." Curr Biol **12**(21): 1852-7.

Senetar, M. A., S. J. Foster, et al. (2004). "Intrasteric inhibition mediates the interaction of the I/LWEQ module proteins Talin1, Talin2, Hip1, and Hip12 with actin." Biochemistry **43**(49): 15418-28.

Singh, A. and S. E. Hitchcock-DeGregori (2003). "Local destabilization of the tropomyosin coiled coil gives the molecular flexibility required for actin binding." Biochemistry **42**(48): 14114-21.

Singh, A. and S. E. Hitchcock-DeGregori (2006). "Dual requirement for flexibility and specificity for binding of the coiled-coil tropomyosin to its target, actin." Structure **14**(1): 43-50.

Toret, C. P. and D. G. Drubin (2006). "The budding yeast endocytic pathway." J Cell Sci **119**(Pt 22): 4585-7.

Traub, L. M. (2005). "Common principles in clathrin-mediated sorting at the Golgi and the plasma membrane." Biochim Biophys Acta **1744**(3): 415-37.

Ungewickell, E. J. and L. Hinrichsen (2007). "Endocytosis: clathrin-mediated membrane budding." Curr Opin Cell Biol.

Wanker, E. E., C. Rovira, et al. (1997). "HIP-I: a huntingtin interacting protein isolated by the yeast two-hybrid system." Hum Mol Genet **6**(3): 487-95.

Yang, S., M. J. Cope, et al. (1999). "Sla2p is associated with the yeast cortical actin cytoskeleton via redundant localization signals." Mol Biol Cell **10**(7): 2265-83.

Ybe, J. A., S. Mishra, et al. (2007). "Crystal structure at 2.8 A of the DLLRKN-containing coiled-coil domain of huntingtin-interacting protein 1 (HIP1) reveals a surface suitable for clathrin light chain binding." J Mol Biol **367**(1): 8-15.

FIGURE LEGENDS

Figure 1. Schematic diagram of the functional domains of Hip proteins showing amino acid numbers that mark domain boundaries. The sequences of residues influencing clathrin light chain binding regions as determined by mutagenesis are shown in alignment (Legendre-Guillemin, Metzler et al. 2005). The structure (PDB 2NO2) of a fragment of the hip1 coiled coil is shown in blue. Red amino acids make up the clathrin light chain binding site predicted by Ybe *et al*. (Ybe, Mishra et al. 2007). Yellow amino acid is L506 mutated in this study.

Figure 2. Hip1 and hip1R are stable homodimers. (A) Hip1cc or hip1Rcc, amino acids 361-637 and 346-655, respectively, were expressed and purified. Purified Hip proteins were diluted into reducing SDS PAGE loading buffer and run on an SDS PAGE gel after boiling for indicated times. The migration positions of monomers and dimers and molecular weight markers (in kDa) are indicated. **(**B**)** Circular dichroism melting analysis of a 1:1 mixture of hip1cc and hip1Rcc. Shown is the molar ellipticity versus

temperature as determined by monitoring the change in helical content of proteins at 222 nm. Homodimers of hip1 and hip1R were mixed at 1:1 ratio then a thermal denaturation profile was obtained (corresponding to melting step 1 in the diagram. open circles O). The mixture of denatured proteins was allowed to anneal slowly back to 20°C (step 2 or 2' in the diagram). The thermal denaturation profile was collected for the newly annealed population (step 3 in the diagram, filled triangles ▲). Both melting curves precisely overlap showing path 2 (not 2') was the primary route followed. (C) Co-immunoprecipitation of His or HA tagged full length hip1 or hip1R. HeLa cells were co-transfected with the two constructs indictated in the table above the blot to achieve equivalent levels of expression of each construct. HeLa cells were lysed and tagged Hip proteins were immunoprecipated with either anti-His antibody or anti-HA antibody as shown. Immunoprecipitates were separated by SDS PAGE and transferred to nitrocellulose for immunoblotting. Nitrocellulose membranes were probed with anti-His, stripped and then probed with anti-HA. Black dot (•) indicate homotypic transfection combinations. Grey dots (•) indictate heterotypic transfection combinations.

Figure 3. Flexibility within the hip1 and hip1R coiled coils. (A) Coiled coil propensity of hip1 and hip1R coiled coil domains as predicted by the program COILS (Lupas, Van Dyke et al. 1991). * indicates the approximate position of residues known to be important for light chain binding as determined by mutagenesis (Legendre-Guillemin, Metzler et al. 2005). (B) Partial proteolysis of hip1 and hip1R coiled coil domains. 12.5 μg of hip1 or hip1R was digested with 0.01 μg/ml of subtilisin protease for indicated times. Reactions were stopped by boiling in SDS PAGE loading buffer. Proteolysis products were separated by SDS PAGE. Fragment sizes in kDa are indicated. (C) Whole hip1 or hip1R coiled coils or subtilisin proteolysis products indicated by arrows (←) in B were excised from the SDS PAGE gel and further digested with trypsin. Peptides from the tryptic digests were identified by MALDI mass spectrometry. Black regions of the diagram indicate tryptic peptides identified by mass spectrometry in each digestion reaction, labeled according to the molecular weight of the band digested. Tryptic peaks

missing (gray) in the subtilisin fragments indicate which regions were digested during subtilisin proteolysis. CLC marks the approximate position of clathrin light chain binding.

Figure 4. Thermal stability of hip1 and hip1R. Purified coiled coil domains of hip1 and hip1R were subjected to increasing temperature while monitoring the change in ellipticity at 222 nm by circular dichroism. Heating rate was 0.5°/min. The mean residue molar ellipticity versus temperature was plotted. The transition for the thermal denaturation of hip1 coiled coil domain had a Tm = 43°C (O). The thermal denaturation profile of hip1R had a Tm = 32°C (▲) for the first and primary transition.

Figure 5. SPR binding of hip1 and hip1R to clathrin light chain and actin. Representative SPR plots for steady state binding between Hip proteins and clathrin light chains or actin. Fitted binding affinities are indicated on each plot, with a vertical line also indicating fitted binding affinity. (A) Purified clathrin light b was immobilized on CM5 chip by amine coupling. Purified coiled coil domains of hip1 or hip1R or hip1 with residue 506 mutated from leucine to alanine (Hip1 L506A) at the concentrations indicated were flowed over the clathrin light chain surface until a steady state was reached and the steady state response units (RU) were plotted. (B) Actin purified from rabbit muscle was assembled and immobilized on a CM5 chip by amine coupling. Purified hip1 or hip1R fragments containing the coiled coil and THATCH domains (hip1ccth or hip1Rccth) were flowed over the F-actin surface with saturating amounts of clathrin light chain peptide (CLC) or GroEL control peptide. All data was collected on a Biacore T100. Biaevaluation software (Biacore) was used for steady state analysis of binding data. Curves with clathrin light chain peptide have no line indicating binding affinity because fitted binding affinity was off the scale.

Figure 6. Hip1 and hip1R change conformation upon clathrin light chain binding. (A) Purified Hip coiled coils (hip1cc and hip1Rcc) or Hip coiled coil plus THATCH fragments (hip1ccth and hip1Rccth) were incubated with ANS in the presence of clathrin light chain (CLC) or control peptide. ANS

fluorescence spectra of the coiled coils were subtracted from ANS spectra of the coiled coil THATCH fragments. The difference between the spectra of the two fragments is plotted, indicating ANS fluorescence due to the THATCH domains alone in the presence of either peptide. (B) Representative electron microscopy images of hip1Rccth in the presence of control or clathrin light chain peptide, stained with uranyl formate. Box edge is 270Å. (C) Length distributions of particles of hip1ccth and hip1Rccth in the presence of clathrin light chain, control, or no peptide from images of uranyl formate stained samples. Distribution was determined by measuring approximately 100 particles for each condition. The control peptide is the clathrin light chain peptide with three mutations that abrogate Hip protein binding (4).

Figure 7. Immunofluorescence of hip1R, actin and clathrin. HeLa cells were transfected with siRNAs for control or both clathrin light chain a and clathrin light chain b for 72 hours. (A) HeLa cells were collected for analysis at 72 hours post-siRNA and subjected to immunoblotting analysis. The proteins were detected by a rabbit polyclonal antiserum against the conserved region of clathrin light chains (LC) (Acton, Wong et al. 1993), monoclonal antibodies to clathrin heavy chain (HC) (TD.1, Convance), β actin (Sigma). Actin is shown for loading control. (B) Control or clathrin light chain knockdown cells were stained for clathrin heavy chain (HC) (green), Hip1R (red), and actin (grey). Bar, 10μm.

Figure 8. Model of Hip protein conformations and effect on vesicle budding. (A) Model of the mechanism for regulation of actin binding to Hip proteins by clathrin light chain. Flexibility in the coiled coil promotes large scale bending of the coiled coil allowing intramolecular binding to a newly exposed site, affecting actin binding directly or repositioning the USH to an inhibitory conformation. (B) The decreased affinity of hip1 and hip1R for actin while bound to clathrin light chain suggests that Hips do not interact with actin while incorporated into the clathrin coat. Instead Hip proteins may interact with

actin at the neck of the budding vesicle or edge of the clathrin coat promoting development of a budding vesicle.

SUPPLEMENTAL FIGURE LEGENDS

Figure S1. Co-immunoprecipitation of hip1 and hip1R after transfection with unequal ratios of expression constructs. Co-immunoprecipitation of His or HA tagged full length hip1 or hip1R. HeLa cells were co-transfected with the two constructs at the ratios indictated. HeLa cells were lysed and tagged Hip proteins were immunoprecipated with either anti-His antibody or anti-HA antibody as shown. Immunoprecipitates were separated by SDS PAGE and transferred to nitrocellulose for immunoblotting. Nitrocellulose membranes were probed with anti-His, stripped and then probed with anti-HA. Anti-His IP shows slight amounts of heterotypic interactions while the same experiment immunoprecipitating with anti-HA show no heterotypic interactions. Both antibodies co-immunoprecipitate large amounts of homotypically interacting protein.

Figure S2. Clathrin light chain binding induces compact a conformation of hip1. Representative electron microscopy images of hip1ccth in the presence of control or clathrin light chain peptide, stained with uranyl formate. Box edge is 270Å.

Figure S3. SPR binding affinity measurements of hip1 and hip1R coiled coil and THATCH domains for F-actin. Actin purified from rabbit muscle was assembled and immobilized on a CM5 chip by amine coupling. Purified hip1 or hip1R fragments containing the coiled coil and THATCH domains (hip1ccth or hip1Rccth) were flowed over the F-actin surface with no peptide present. All data was collected on a Biacore T100. Biaevaluation software (Biacore) was used for steady state analysis of binding data.

Figure S4. SPR binding affinity measurements of hip1 and hip1R THATCH domains for F-actin. Actin purified from rabbit muscle was assembled and immobilized on a CM5 chip by amine coupling. Purified hip1 or hip1R fragments containing THATCH domains (hip1th or hip1Rth) were flowed over the F-actin surface. All data was collected on a Biacore T100. Biaevaluation software (Biacore) was used for steady state analysis of binding data. The control peptide is the clathrin light chain peptide with three mutations that abrogate Hip protein binding (Chen and Brodsky, 2005).

Figure S5. SPR binding affinity measurement of hip1R coiled coil and THATCH domain for cortactin. GST-cortactin fusion protein was immobilized by capturing with anti-GST antibody coupled to a CM5 chip. Purified hip1R fragments containing the coiled coil and THATCH domain (hip1ccRth) were flowed over the cortactin surface with control or clathrin light chain peptide. All data was collected on a Biacore T100. Biaevaluation software (Biacore) was used for steady state analysis of binding data. The control peptide is the clathrin light chain peptide with three mutations that abrogate Hip protein binding (Chen and Brodsky, 2005).

Figure S6. SPR binding affinity measurement of hip1 coiled coil and THATCH domains for clathrin light chain preincubated with F-actin. Full length neuronal clathrin light chain was coupled to a CM5 chip. Purified hip1 fragments containing the coiled coil and THATCH domain (hip1ccth) were flowed over the clathrin light chain surface with F-actin or no actin already bound. All data was collected on a Biacore T100. Biaevaluation software (Biacore) was used for steady state analysis of binding data.

Table I.

Hip coiled coil molecular weight determined by analytical ultracentrifugation fitted to monomer-dimer equilibrium model.

Protein	Calculated Dimer MW (kDa)	Calculated Monomer MW (kDa)	Experimental MW (kDa)	Fitting Variance
Hip1cc	66.79	33.39	67.4	2.51×10^{-5}
HIP1Rcc	73.45	36.73	67.7	2.57×10^{-5}

Figure 1.

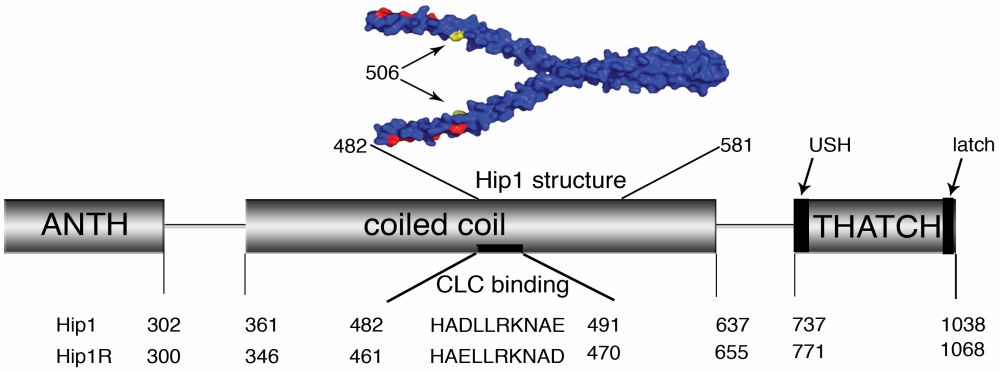

506
482
581
USH
latch
Hip1 structure
ANTH
coiled coil
THATCH
CLC binding
Hip1 302 361 482 HADLLRKNAE 491 637 737 1038
Hip1R 300 346 461 HAELLRKNAD 470 655 771 1068

Figure 2.

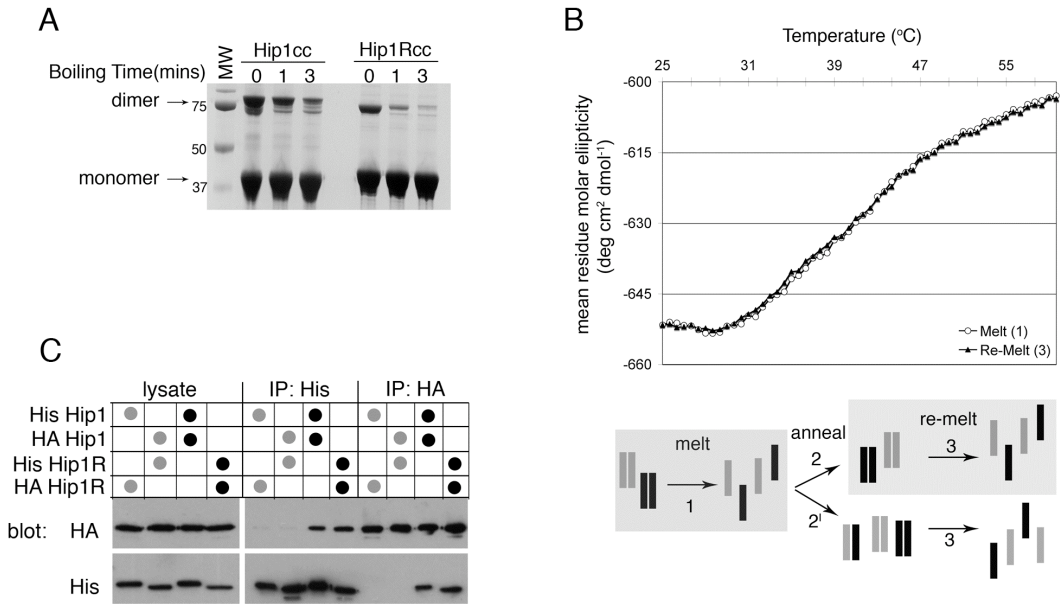

A
Boiling Time(mins)
MW
Hip1cc
0 1 3
Hip1Rcc
0 1 3
dimer
75
50
monomer
37
B
Temperature (°C)
25 31 39 47 55
mean residue molar ellipticity (deg cm² dmol⁻¹)
-600
-615
-630
-645
-660
Melt (1)
Re-Melt (3)
melt
1
anneal
2
2'
re-melt
3
3
C
lysate
IP: His
IP: HA
His Hip1
HA Hip1
His Hip1R
HA Hip1R
blot: HA
His

Figure 3.

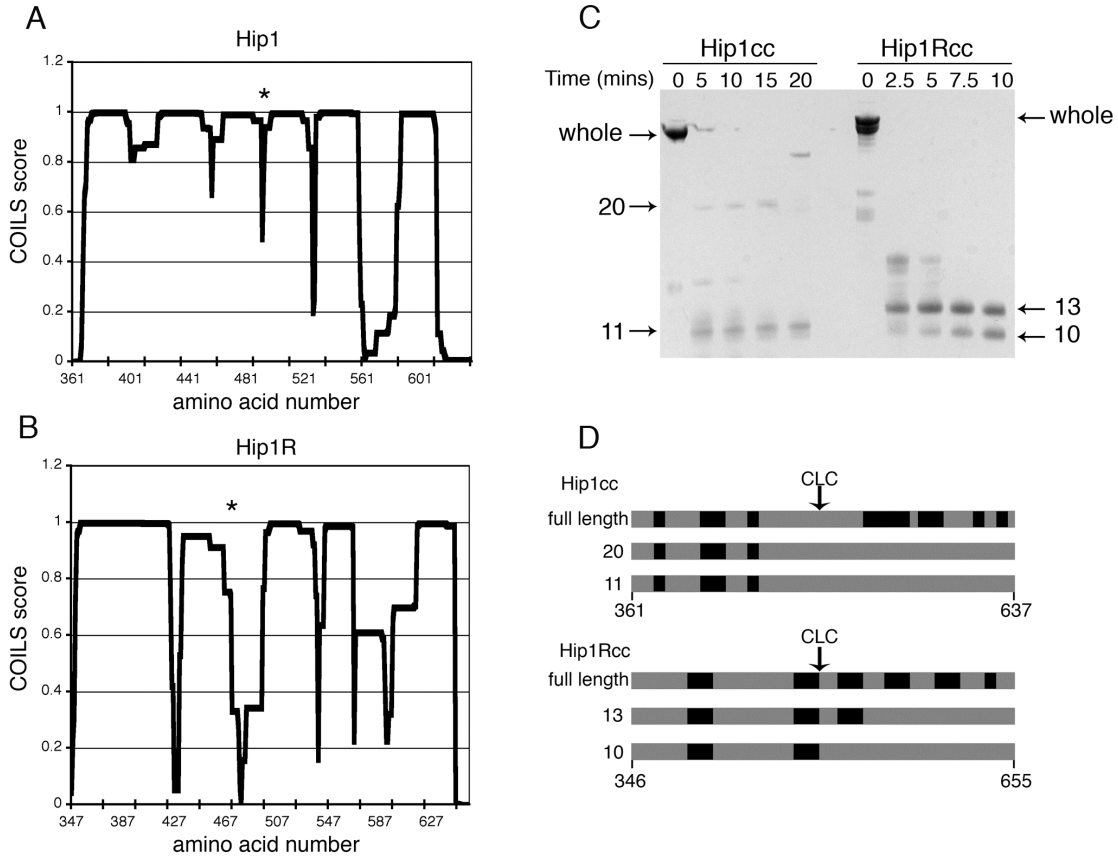

A
Hip1
COILS score
amino acid number
B
Hip1R
COILS score
amino acid number
C
Hip1cc
Hip1Rcc
Time (mins)
0 5 10 15 20
0 2.5 5 7.5 10
whole
20
11
13
10
D
Hip1cc
CLC
full length
20
11
361
637
Hip1Rcc
CLC
full length
13
10
346
655

Figure 4.

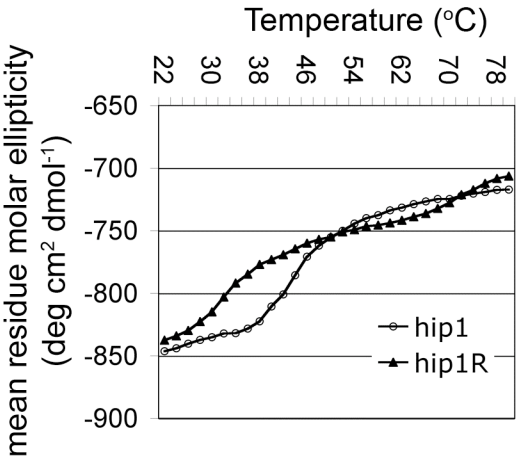

Temperature (°C)
22
30
38
46
54
62
70
78
mean residue molar ellipticity
(deg cm2 dmol-1)
-650
-700
-750
-800
-850
-900
hip1
hip1R

Figure 5.

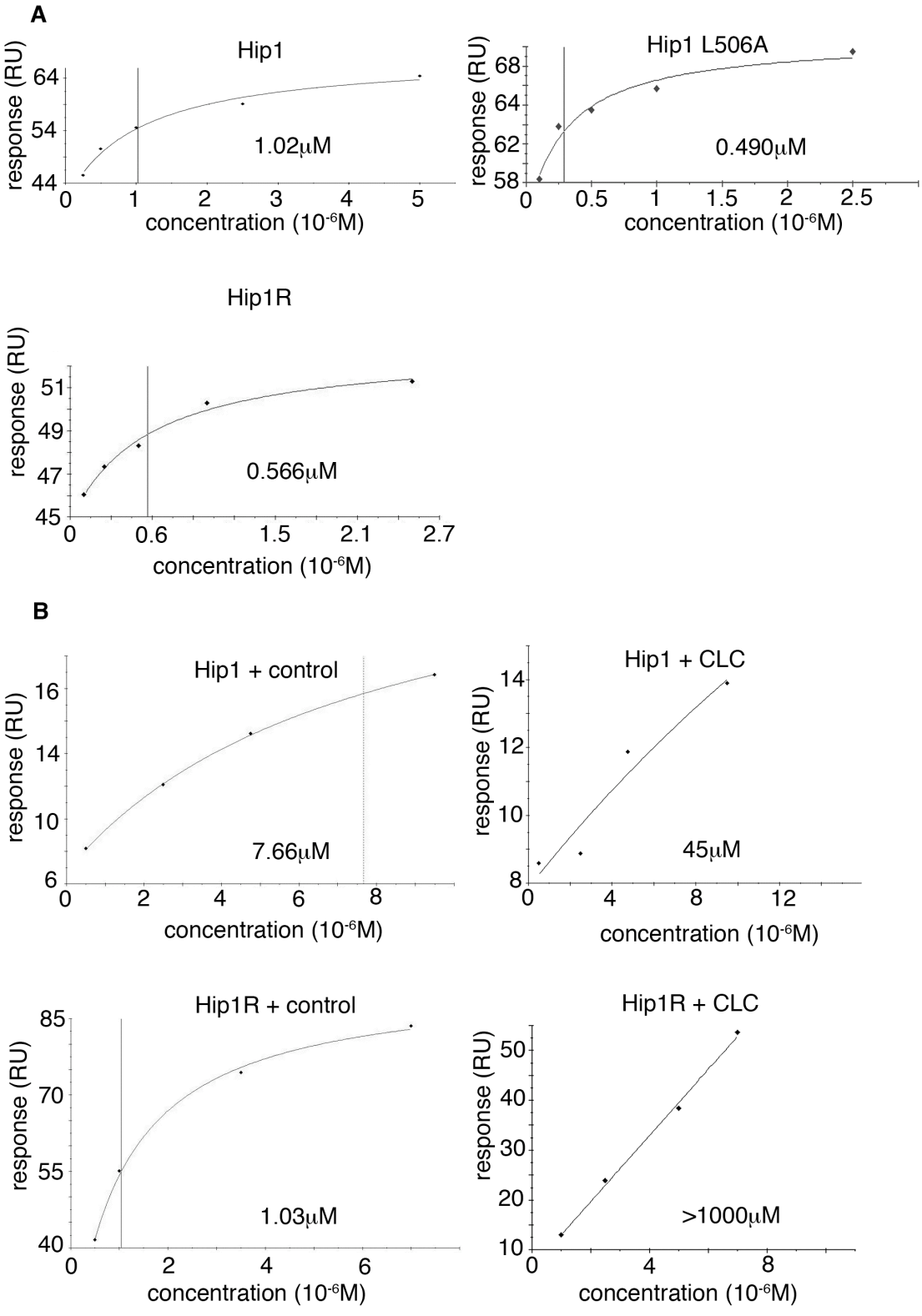

A
Hip1
response (RU)
64
54
44
0
1
2
3
4
5
1.02μM
concentration (10-6M)
Hip1 L506A
response (RU)
68
64
62
58
0
0.5
1
1.5
2
2.5
0.490μM
concentration (10-6M)
Hip1R
response (RU)
51
49
47
45
0
0.6
1.5
2.1
2.7
0.566μM
concentration (10-6M)
B
Hip1 + control
response (RU)
16
14
10
6
0
2
4
6
8
7.66μM
concentration (10-6M)
Hip1 + CLC
response (RU)
14
12
10
8
0
4
8
12
45μM
concentration (10-6M)
Hip1R + control
response (RU)
85
70
55
40
0
2
4
6
1.03μM
concentration (10-6M)
Hip1R + CLC
response (RU)
50
40
30
20
10
0
4
8
>1000μM
concentration (10-6M)

Figure 6.

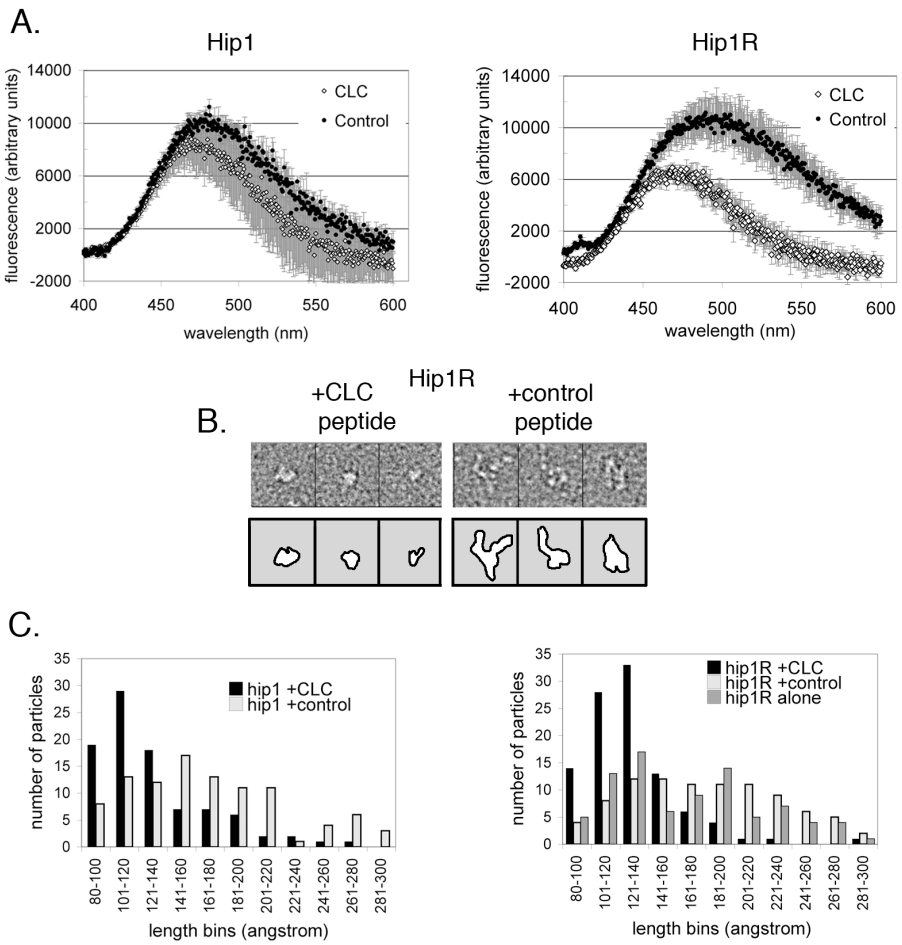

A.
Hip1
Hip1R
fluorescence (arbitrary units)
14000
10000
6000
2000
-2000
400
450
500
550
600
wavelength (nm)
CLC
Control
B.
Hip1R
+CLC peptide
+control peptide
C.
number of particles
hip1 +CLC
hip1 +control
hip1R +CLC
hip1R +control
hip1R alone
80-100
101-120
121-140
141-160
161-180
181-200
201-220
221-240
241-260
261-280
281-300
length bins (angstrom)

Figure 7.

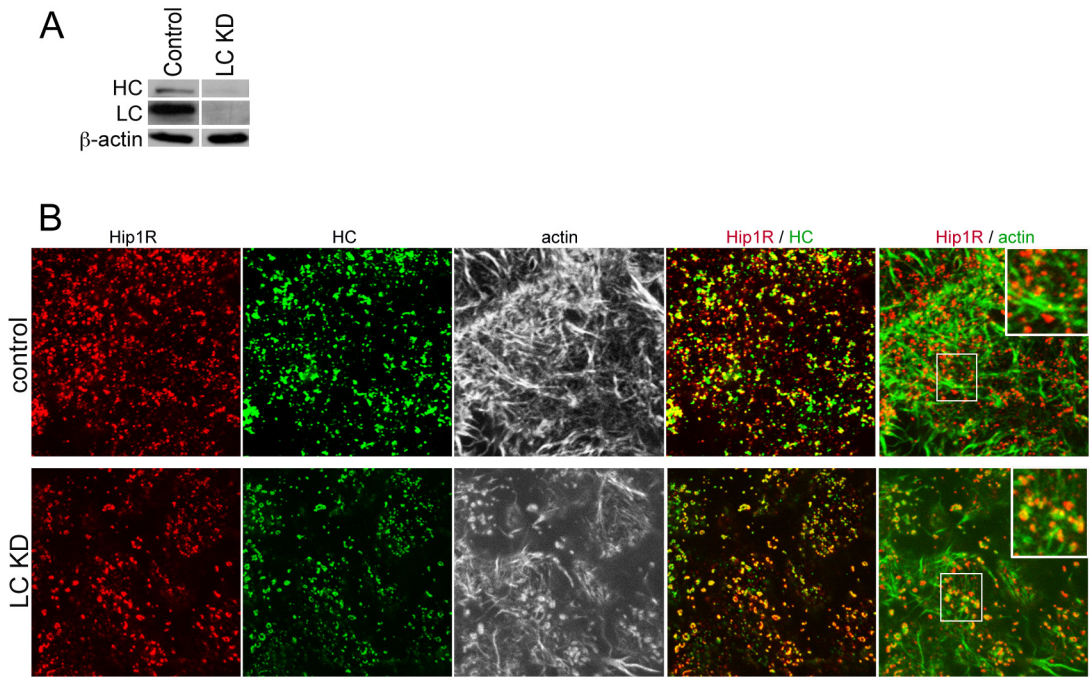
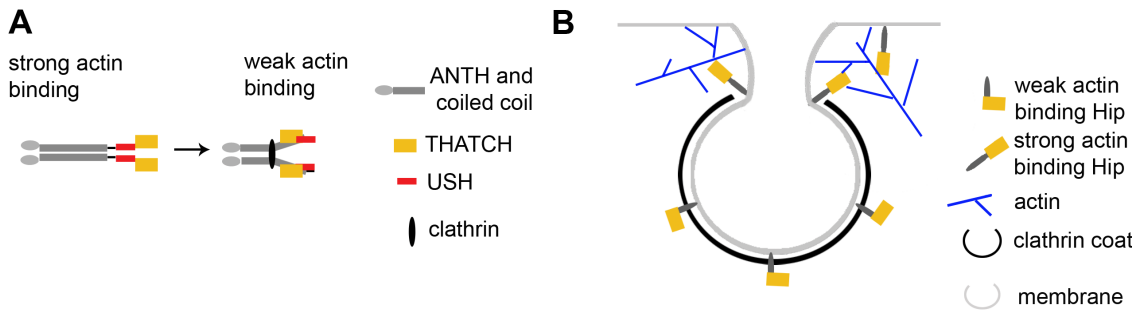
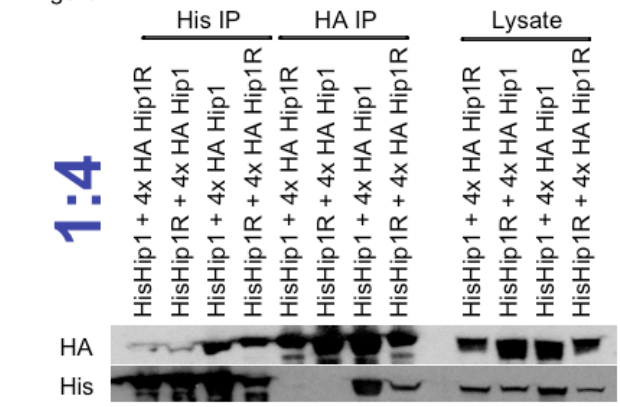

A
Control
LC KD
HC
LC
β-actin
B
Hip1R
HC
actin
Hip1R / HC
Hip1R / actin
control
LC KD

Figure 8.


A
strong actin binding
weak actin binding
ANTH and coiled coil
THATCH
USH
clathrin
B
weak actin binding Hip
strong actin binding Hip
actin
clathrin coat
membrane

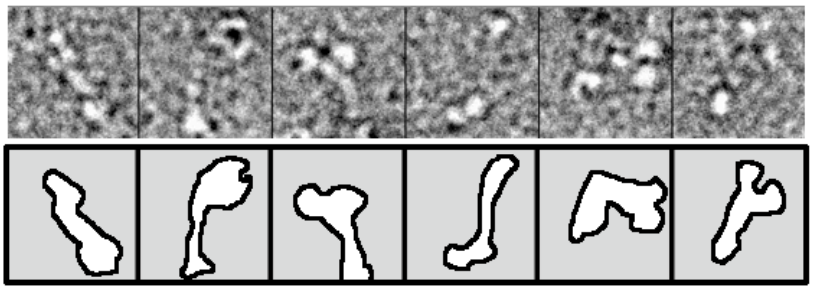
SUPPLEMENTAL FIGURES

Supplemental Figure 1.

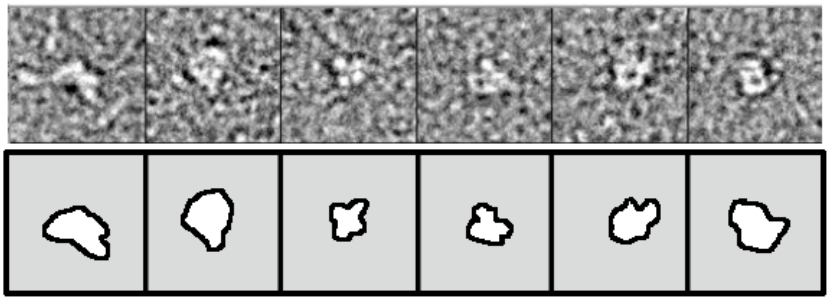


Suppletmental
Figure 2.

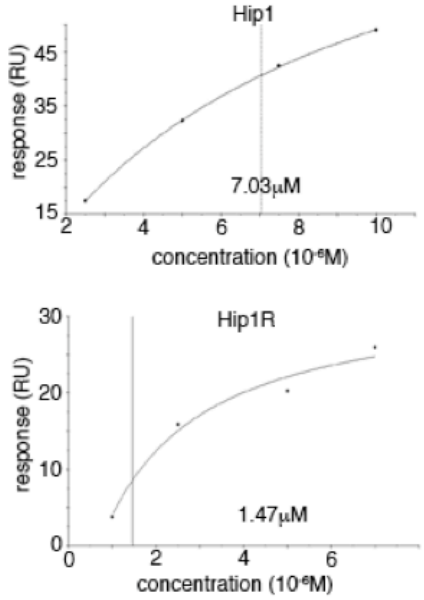
hip1ccth + control peptide



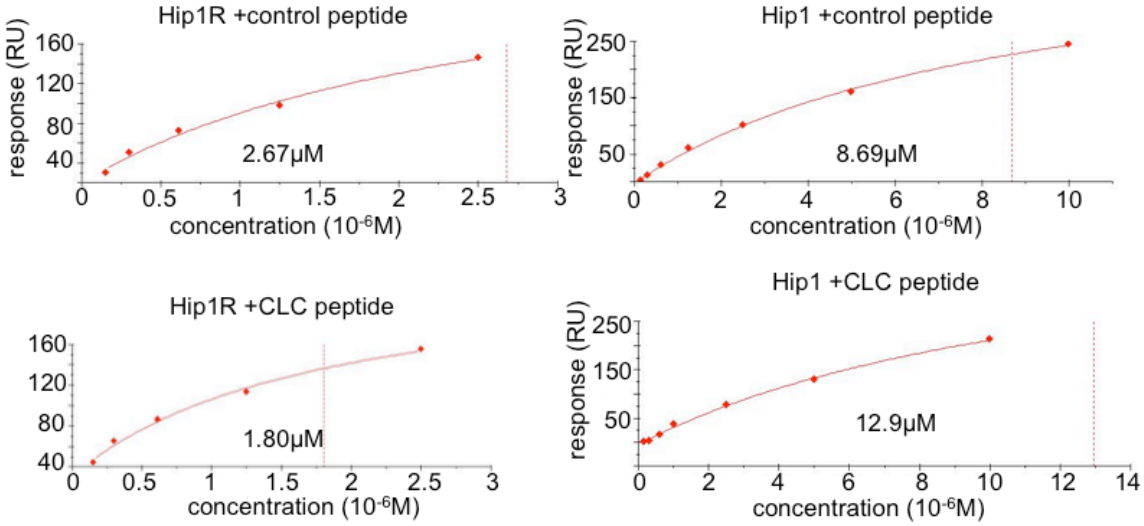
hip1ccth + CLC peptide



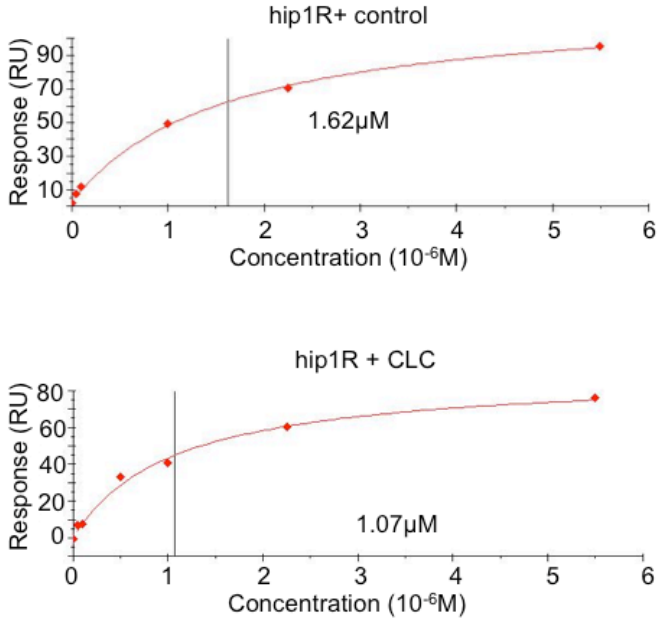
Supplemental
Figure 3.



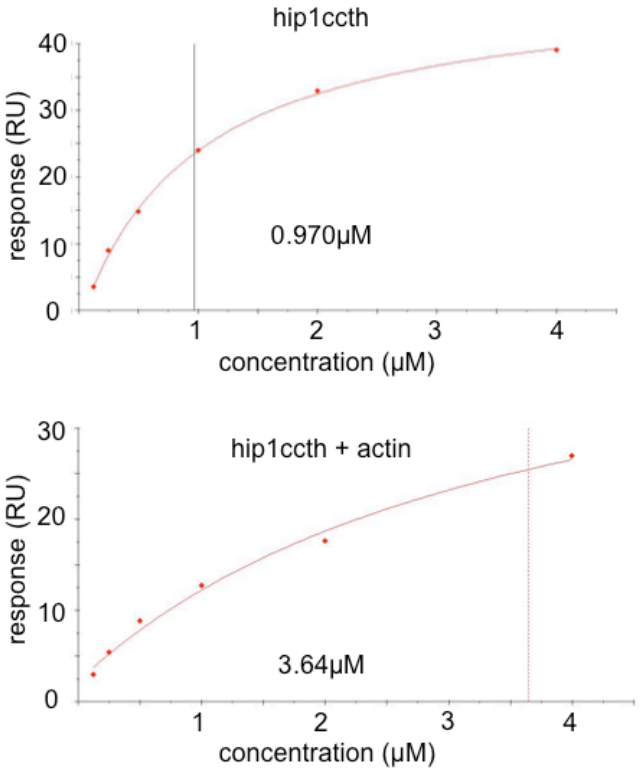
Supplemental
Figure 4.



Supplemental
Figure 5.



Supplemental figure 6



Chapter 3

Regulation of clathrin lattice assembly by tyrosine phosphorylation

Phosphorylation is an efficient means of transferring information from one part of the cell to another. Signaling events occurring at the plasma membrane initiate long cascades of phosphorylation to transmit information to the nucleus for transcriptional reprogramming of the cell. Phosphorylation can change protein conformations, and mediate or prevent protein interactions. Phosphorylation likely has evolved as a means for dramatically but reversibly changing of the chemical environment around certain amino acid residues within a protein. The attachment of a negatively charged phosphate group through the hydroxyl of serine, threonine, or tyrosine, switches polar but uncharged residues to very negatively charged residues at physiological pH.

Many proteins involved with clathrin mediated endocytosis are phosphorylated during or after clathrin coat formation. Proteins destined to become cargo are phosphorylated on tail domains within the cytoplasm to regulate the interaction with adaptor proteins (Marion, Fralish et al. 2007). Adaptor proteins are phosphorylated themselves to regulate multiple protein interactions (Flett, Semerdjieva et al. 2005; Fessart, Simaan et al. 2005; Huang, Jiang et al. 2003). Clathrin accessory proteins in eps15, amphiphysin and dynamin are all phosphorylated to achieve highly regulated spatial and temporal functioning (Hill, Drickamer et al. 1988; Chen, Slepnev et al. 1999; Wilde, Beattie et al. 1999; Tomizawa, Sunada et al. 2003). Proteins involved with remodeling the actin cytoskeleton around budding clathrin vesicles are also phosphorylated (Chen, Slepnev et

al. 1999). Clathrin itself is phosphorylated on both clathrin light chain and clathrin heavy chain (Crotzer, Mabardy et al. 2004). The numerous phosphorylation events that occur during clathrin mediated endocytosis likely occur in a highly coordinated manner to promote the appropriate timing of protein interactions.

In the case of clathrin, the molecular mechanisms of regulation by phosphorylation are unknown. Cellular functions have been attributed to phosphorylation but a cohesive mechanistic model has yet to be determined. Signaling through the epidermal growth factor receptor (EGFR), the T-cell receptor (TCR) and the B-cell receptor (BCR) all induce phosphorylation of the clathrin heavy chain (Crotzer, Mabardy et al. 2004). This phosphorylation is achieved through the action of Src family kinases and for the case of signaling through the EGFR is correlated with a redistribution of clathrin to the cell periphery. The signaling receptor at the cell surface can then be internalized by clathrin mediated endocytosis and signaling can be down regulated. Interestingly in the case of the TCR the phosphorylation of clathrin heavy chain must occur to get recruitment of clathrin to the cell surface but to internalize the TCR clathrin must also be dephosphorylated (Crotzer, Mabardy et al. 2004). By blocking dephosphorylation with pervanadate clathrin appears to be restricted to the cell surface and there are increased levels of clathrin phosphorylation. The BCR case is such that phosphorylation of clathrin achieves integration of two pathways for endocytosis (Stoddart, Dykstra et al. 2002). That is phosphorylation appears to act on signaling events in lipid raft associated receptors targeting them for internalization through clathrin mediated endocytosis.

Mutants of clathrin that cannot be phosphorylated at tyrosine residue 1477 function poorly during internalization (Wilde, Beattie et al. 1999). It is clear that phosphorylation of tyrosine 1477 is important for the spatial regulation of clathrin in the cell. For both the EGFR and TCR phosphorylation is critical in the early stages of organization of the clathrin machinery and for the TCR case it is compelling that dephosphorylation is necessary for clathrin mediated endocytosis. This would suggest two possible functions for the phosphorylation of clathrin heavy chain. One is as a localization signal and the other is as a modulator of clathrin assembly state.

To help determine the molecular mechanism of phosphorylation of clathrin I sought to define the most fundamental aspect to clathrin phosphorylation in detail. Though it was known that clathrin heavy chain tyrosine residue 1477 was phosphorylated in cells, it was unclear whether this was the only sight of phosphorylation by the Src family kinases. In an effort to unequivocally determine the positions of phosphorylation on clathrin heavy chain in response to Src family kinases I developed and *in vitro* phosphorylation reaction to phosphorylate clathrin heavy chain and used mass spectrometry to determine the position of phosphorylated residues within the clathrin sequence. This method led to the discovery of a novel point of phosphorylation at tyrosine residue 1487.

Methods

Phosphorylation of clathrin heavy chain

Purified bovine clathrin heavy chain residues 1074-1675 (hub) or 1-1073 (TDD) was expressed and purified as previously described. In a total volume of 80 μl, 100-500 μg of

purified hub or TDD was incubated for 1-4 hours with 1mM $MgCl_2$, 100μM $MnCl_2$, 1mM ATP, and 1μl commercially available Src, Lck, or Lyn at room temperature.

Mass Spectrometry

Phosphorylation reactions were digested with proteomics grade procine trypsin (Sigma) by addition of 1 μl of 20 μg/ml. Digestion was allowed to occur for 4-16 hours at 37°C. Digestion reactions were desalted with C18 ZipTips (Millipore) following the manufacturers protocol and eluted directly onto the MALDI mass spectrometry sample plate. Sample spectra were collected in a Voyager STR MALDI/TOF mass spec, or a Voyager MS/MS mass spec. This method yielded excellent spectra with many identifiable peptide mass peaks. Tryptic peptides containing phosphorylated tyrosine residues were determined by an 80 Da increase in the mass compared to the calculated mass of the peptide. Peptides were identified using MS-Digest in the Protein Prospector suite (UCSF).

Results

Identification of peptides resulting from *in vitro* phosphorylation and tryptic digestion of clathrin yield 30%-40% coverage of the proteins (hub or TDD). Residue 1477 fell within a tryptic peptide of ~2356 Da. Easily identifiable was another tryptic peptide exactly 80 Da increased in mass. Both peptide existed in the sample due to incomplete phosphorylation of clathrin hub. Sequencing the peptide by MS/MS confirmed the identity of the peptide and further showed that tyrosine 1477 was modified by 80 Da, equivalent to a phosphate. Searching through other peptides from clathrin hub yielded

another precise 80 Da difference between the peptide of mass ~1942 and peptide of mass 2022. This suggested another phosphorylation event. Identification of the tryptic peptide containing this phosphorylation event showed it contained tyrosine 1487 and no other tyrosines. Sequencing the peptide containing tyrosine 1487 again confirmed the identity and the modification of tyrosine 1487. TDD yielded no peptides of unexpected mass (data not shown).

By developing a robust *in vitro* phosphorylation reaction, I was able to determine that phosphorylation of clathrin heavy chain occurs on two residues, tyrosines 1477 and 1487. Phosphorylation on these residues was confirmed for Src, Lck, Lyn tyrosine kinases.

Discussion

Identification of the residues that can be modified by phosphorylation is a first step in understanding the molecular mechanism regulating the function of protein. For clathrin heavy chain, though it was known that phosphorylation on tyrosine 1477 was important for complete function, all the data was not accounted for by this single phosphorylation event. To help understand the role of phosphorylation on clathrin heavy chain I identified a second phosphorylation sight that likely plays important roles within the cell.

Mapping the positions of phosphorylation onto the known structures of clathrin shows that both phosphorylation sites occur on the same helix on the proximal leg approaching the trimerization domain. Both phosphorylation sites also occur on the interaction interface between two clathrin triskelia in the assembled state. This suggests that

phosphorylation maybe incompatible with assembly or alternatively compatible with only the flat lattice assembly state. If phosphorylation of clathrin is completely incompatible with assembly is its clear why both phosphorylation and dephosphorylation are necessary for productive endocytosis. Given that the pucker of a clathrin triskelia is relatively rigid there must be change in the angle between legs of interacting triskelia upon changing from flat to curved lattices. In the flat lattice proximal legs interact primarily in the their central regions at a steep angle. This steep angle of interaction would position the phosphorylation sites outside interaction interface. The leg of an adjacent triskelia would pass below the phosphorylation site and maybe fixed at a steep angle of interaction by the phosphorylation sites. Upon dephosphorylation, adjacent legs could interact with a less steep interaction angle and could adopt a curved lattice type of interaction. This proposed mechanism is consistent with the TCR data that suggests both phosphorylation and dephosphorylation are necessary for productive clathrin mediated endocytosis.

Phosphorylation may, alternatively, localize clathrin to the membrane. Upon signaling through a cell surface receptor clathrin gets phosphorylated via a Src family kinase. This phosporylation signal is the initiation of endocytosis of the receptor. Phosphorylation may recruit clathrin to the membrane by interacting with a phospho-specific membrane localized binding protein. This model may also fit with the model of complete inhibition of assembly in the phosphorylated state. As clathrin is recruited it is simultaneously prevented from assembling until it is properly localized on the membrane.

Though both functions of phosphorylation of clathrin fit well with the current data they clearly provide separate testable predictions. On prediction born from the regulation of lattice curvature model is that phosphorylated clathrin should not form curved lattices well. A prediction of the localization model is that phosporylated clathrin interacts with a cellular component that unphosphorylated clathrin does not interact with.

Finally, if one considers each model closely, neither is mutally exclusive. It is possible that phosphorylation maybe perform both of these functions or functions not yet conceived. This work however, provides the important first step of defining primary phosphorylation sites within clathrin heavy chain.

References

Chen, H., V. I. Slepnev, et al. (1999). "The interaction of epsin and Eps15 with the clathrin adaptor AP-2 is inhibited by mitotic phosphorylation and enhanced by stimulation-dependent dephosphorylation in nerve terminals." J Biol Chem **274**(6): 3257-60.

Crotzer, V. L., A. S. Mabardy, et al. (2004). "T cell receptor engagement leads to phosphorylation of clathrin heavy chain during receptor internalization." J Exp Med **199**(7): 981-91.

Fessart, D., M. Simaan, et al. (2007). "Src-dependent phosphorylation of beta2-adaptin dissociates the beta-arrestin-AP-2 complex." J Cell Sci **120**(Pt 10): 1723-32.

Flett, A., S. Semerdjieva, et al. (2005). "Regulation of the clathrin-coated vesicle cycle by reversible phosphorylation." Biochem Soc Symp(72): 65-70.

Hill, B. L., K. Drickamer, et al. (1988). "Identification of the phosphorylation sites of clathrin light chain LCb." J Biol Chem **263**(12): 5499-501.

Huang, F., X. Jiang, et al. (2003). "Tyrosine phosphorylation of the beta2 subunit of clathrin adaptor complex AP-2 reveals the role of a di-leucine motif in the epidermal growth factor receptor trafficking." J Biol Chem **278**(44): 43411-7.

Marion, S., G. B. Fralish, et al. (2007). "N-terminal tyrosine modulation of the endocytic adaptor function of the beta-arrestins." J Biol Chem **282**(26): 18937-44.

Stoddart, A., M. L. Dykstra, et al. (2002). "Lipid rafts unite signaling cascades with clathrin to regulate BCR internalization." Immunity **17**(4): 451-62.

Tomizawa, K., S. Sunada, et al. (2003). "Cophosphorylation of amphiphysin I and dynamin I by Cdk5 regulates clathrin-mediated endocytosis of synaptic vesicles." J Cell Biol **163**(4): 813-24.

Wilde, A., E. C. Beattie, et al. (1999). "EGF receptor signaling stimulates SRC kinase phosphorylation of clathrin, influencing clathrin redistribution and EGF uptake." Cell **96**(5): 677-87.

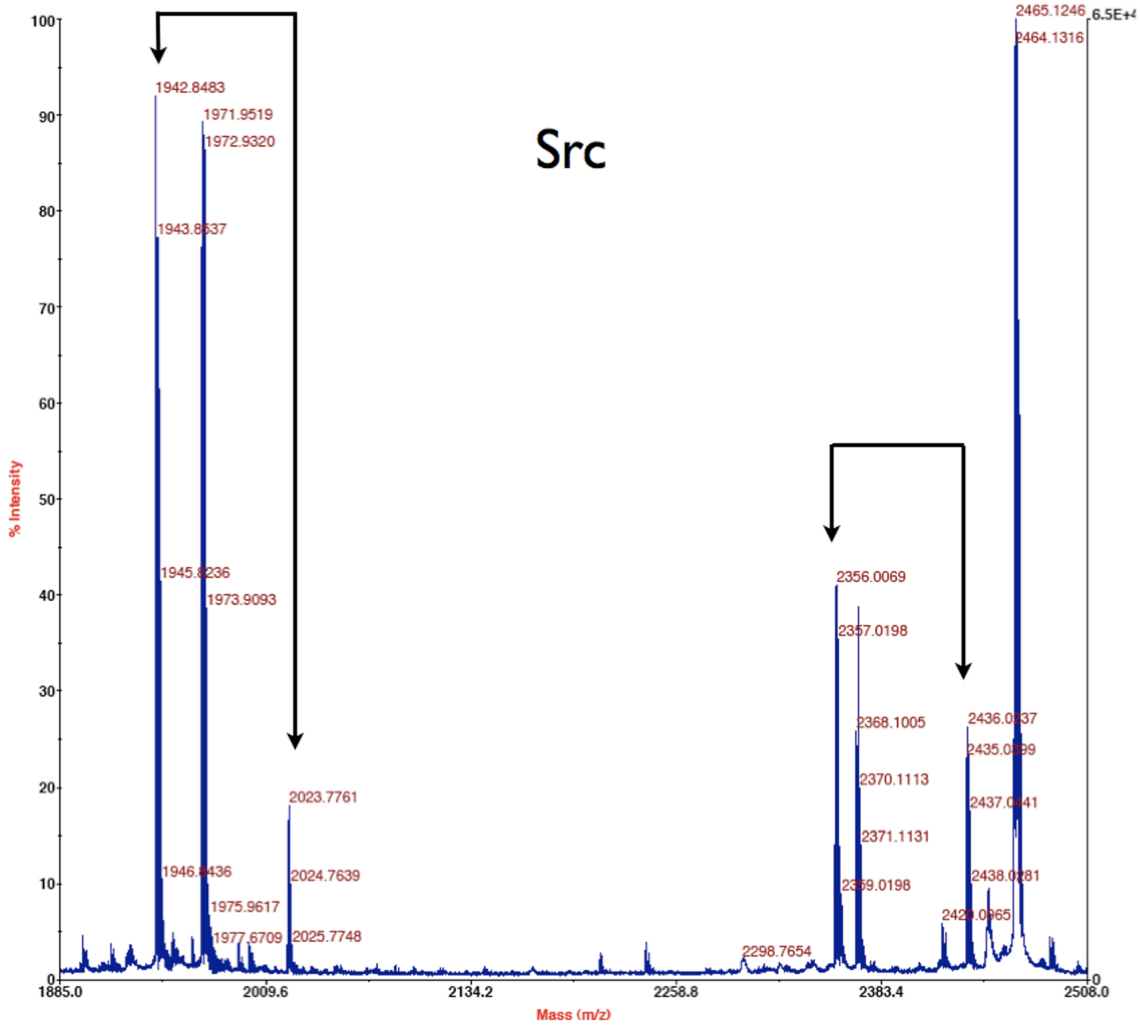
Figure 1. MALDI mass spectra of phosphorylated clathrin hub. Purified clathrin hub was incubated with either Src, Lck, or Lyn kinase in buffer containing ATP. Reactions were allowed to proceed for 1-3 hours at which point 0.5ug trypsin protease was added. The reactions were digested for 16hrs at room temperature and spotted onto a MALDI mass spec plate after desalting with ZipTip C18 tips. Spectra were collected with grid voltage of 20,000 mV, grid spacing of 65.5. Arrows point to peaks 80 Da apart.

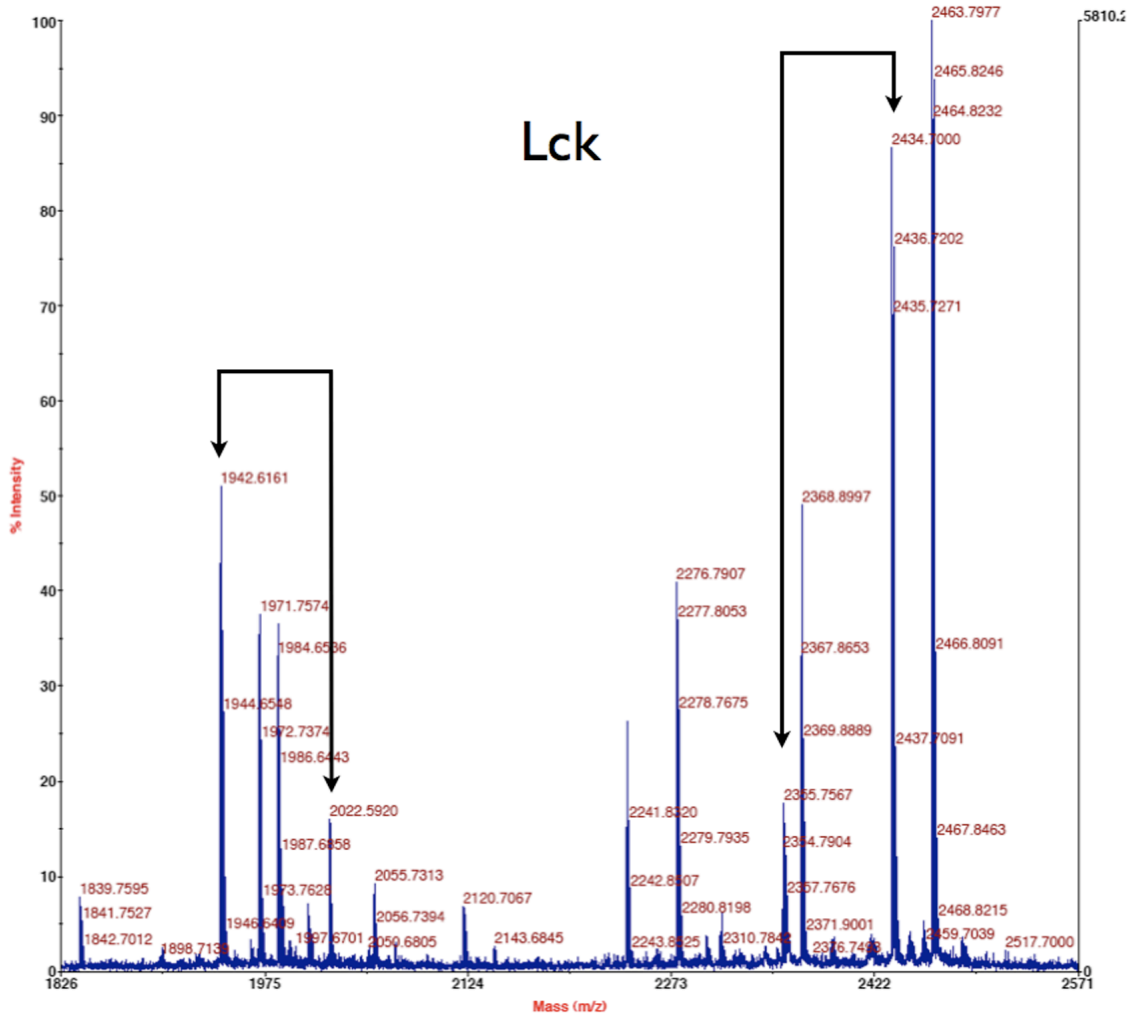
Figure 2. MS/MS spectra gated on the peptide of MW 2436. This peptide contains tyrosine 1477 of clathrin hub. Arrows point to peaks 80 Da different. Indicating phosphorylated and nonphorylated peptide fragments.

Figure 3. MS/MS spectra gated on the peptide of MW 2022. This peptide contains tyrosine 1487 of clathrin hub. Arrows point to peaks 80 Da different. Indicating phosphorylated and nonphorylated peptide fragments.

Figure 4. Schematic view of interacting clathrin triskelia with phosphorylation site indicated. The hypothesized mechanism of preventing curve lattice formation is shown by legs clashing with phosporylation sites.

Figure 1.


Src
% Intensity
Mass (m/z)
1885.0
2009.6
2134.2
2258.8
2383.4
2508.0
1942.8483
1971.9519
1972.9320
1943.8537
1945.8236
1973.9093
1946.8436
1975.9617
1977.6709
2023.7761
2024.7639
2025.7748
2298.7654
2356.0069
2357.0198
2368.1005
2370.1113
2371.1131
2359.0198
2436.0537
2435.0399
2437.0441
2438.0281
2429.0965
2465.1246
2464.1316
6.5E+4


Lck
% Intensity
Mass (m/z)
1826
1975
2124
2273
2422
2571
1942.6161
1971.7574
1984.6586
1944.6548
1972.7374
1986.6443
2022.5920
1987.6858
1839.7595
1841.7527
1842.7012
1973.7628
2055.7313
2056.7394
2059.6805
2120.7067
2143.6845
2276.7907
2277.8053
2278.7675
2241.8320
2279.7935
2242.8407
2280.8198
2243.8825
2310.7842
2368.8997
2367.8653
2369.8889
2365.7567
2364.7904
2357.7676
2371.9001
2316.7468
2463.7977
2465.8246
2464.8232
2434.7000
2436.7202
2435.7271
2466.8091
2437.7091
2467.8463
2468.8215
2459.7039
2517.7000

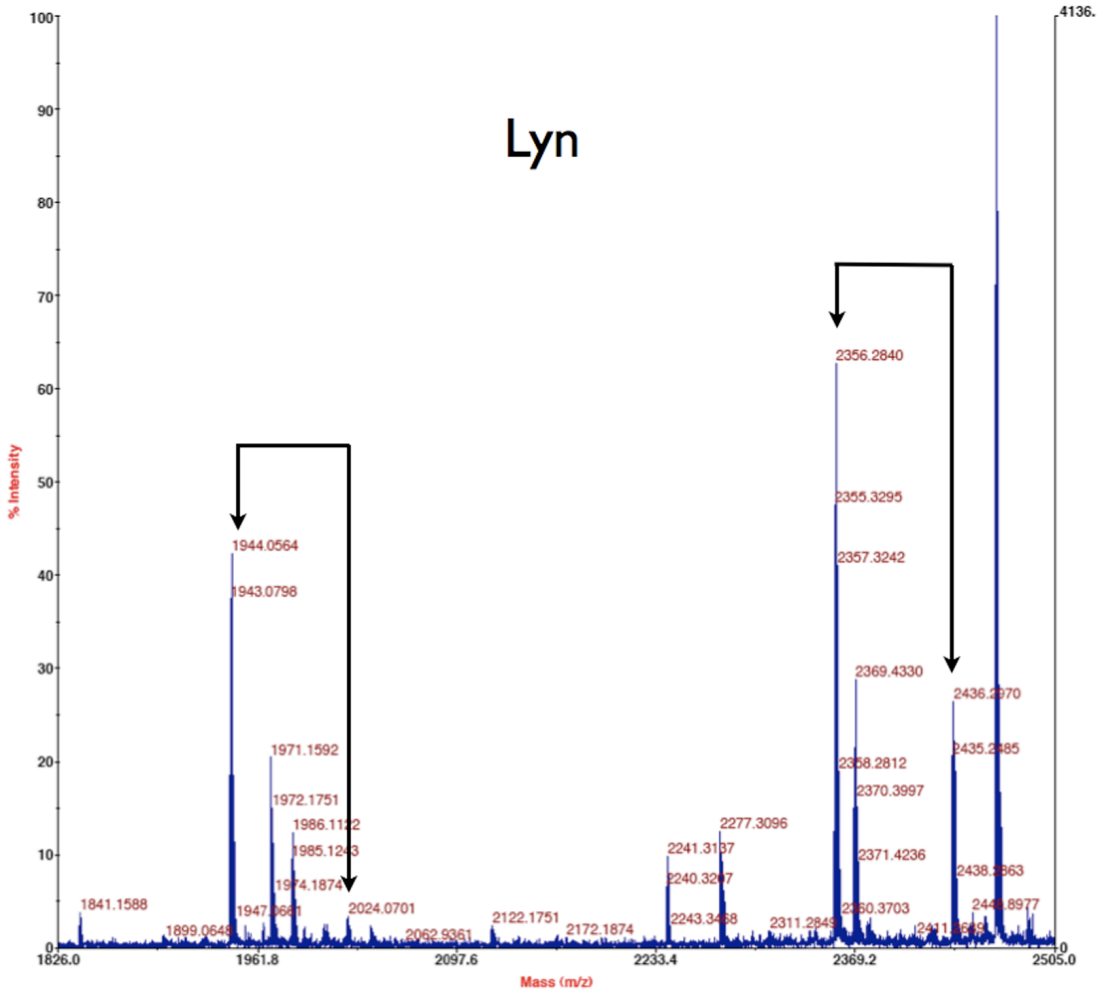

Lyn
% Intensity
Mass (m/z)
1826.0
1961.8
2097.6
2233.4
2369.2
2505.0
4136.
1944.0564
1943.0798
1971.1592
1972.1751
1986.1122
1985.1243
1974.1874
1841.1588
1899.0648
2024.0701
2062.9361
2122.1751
2172.1874
2277.3096
2241.3137
2240.3247
2243.3458
2311.2849
2356.2840
2355.3295
2357.3242
2369.4330
2358.2812
2370.3997
2371.4236
2360.3703
2436.2970
2435.2985
2438.3363
2448.8977

Figure 2.

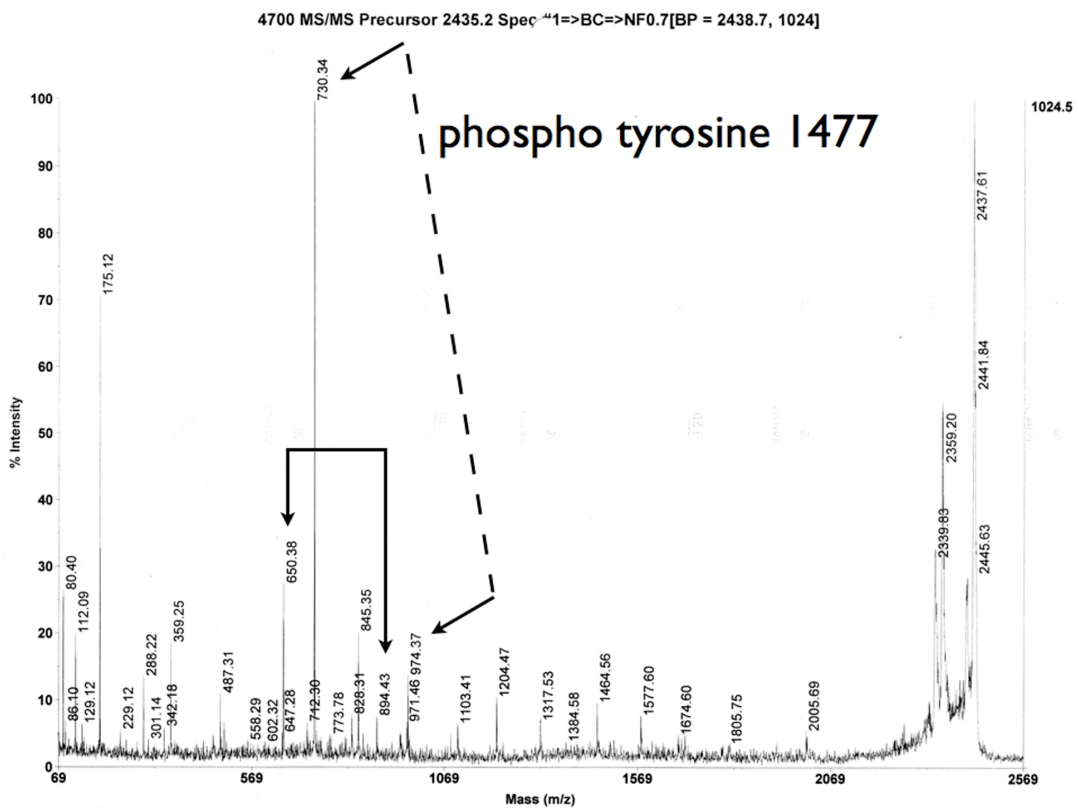

4700 MS/MS Precursor 2435.2 Spec "1=>BC=>NF0.7[BP = 2438.7, 1024]
phospho tyrosine 1477
% Intensity
Mass (m/z)
730.34
1024.5
175.12
2437.61
2441.84
2359.20
2339.83
2445.63
650.38
845.35
80.40
112.09
86.10
129.12
229.12
288.22
301.14
342.18
359.25
487.31
558.29
602.32
647.28
712.30
773.78
828.31
894.43
971.46
974.37
1103.41
1204.47
1317.53
1384.58
1464.56
1577.60
1674.60
1805.75
2005.69

Figure 3.

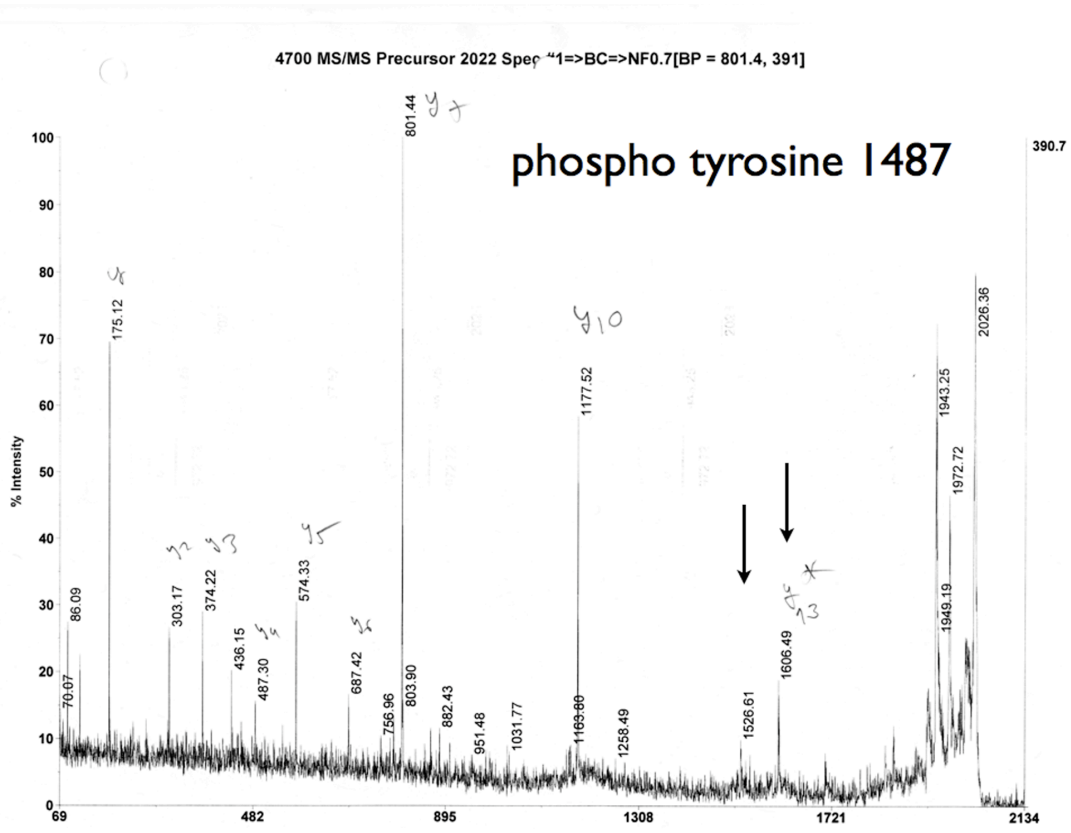
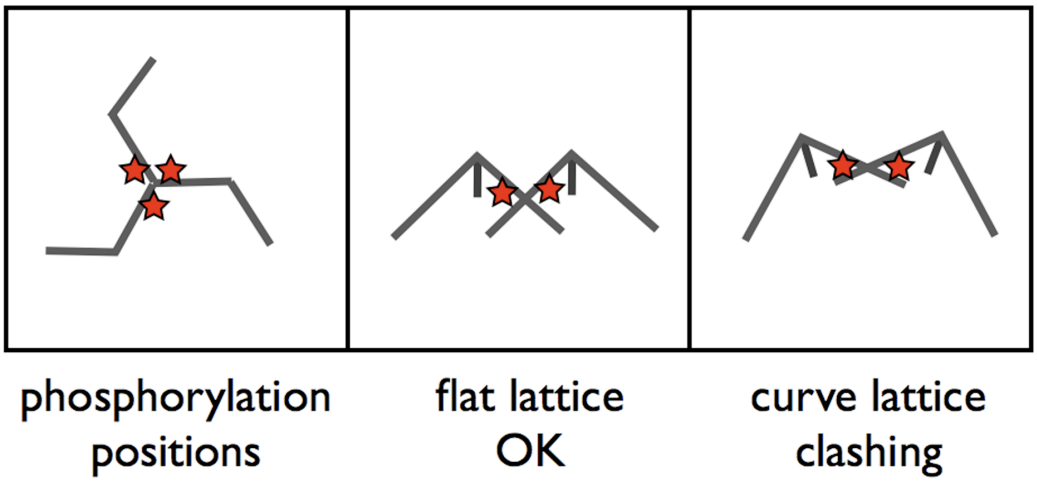

4700 MS/MS Precursor 2022 Spec "1=>BC=>NF0.7[BP = 801.4, 391]
phospho tyrosine 1487
% Intensity
801.44
175.12
1177.52
2026.36
1943.25
1972.72
1949.19
390.7
303.17
374.22
436.15
487.30
574.33
687.42
756.96
803.90
882.43
951.48
1031.77
1163.60
1258.49
1526.61
1606.49
86.09
70.07
100
90
80
70
60
50
40
30
20
10
0
69
482
895
1308
1721
2134

Figure 4.


phosphorylation positions
flat lattice OK
curve lattice clashing

Chapter 4: On the phasing of low-resolution crystal structures by molecular replacement.

In collaboration with Ben Sellers and Peter Hwang

Introduction

X-ray crystallography continues to be among the most important methods for determining the three dimensional structure of biological macromolecules. Despite 96 years of X-ray crystallography experiments numerous difficulties still exist within a standard X-ray crystallography experiment. The determination of the phases of X-rays that generate diffraction images from a crystal continue to be the significant challenge. Modern computation and the advent of tune-able synchrotron radiation has provided better means for solving the phase problem and enormously impacted our ability to determine macromolecular structures. Tuning the X-ray energy to absorption edges of certain atoms has allowed the development of multiple anomalous dispersion (MAD) and pushed this method to the forefront of novel structure determination. Multiple isomorphous replacement (MIR) continues to a powerful technique but is limited to obtaining isomorphous crystal derivatives which can be challenging. Molecular replacement uses previous structural and sequence information to calculate approximate phases from a model of the macromolecule in question. Considering over 51,000 structures have been determined and submitted to the protein databank, molecular replacement as a method of solving the phase problem will continue to grow. Currently it is responsible for the majority of new structure solutions.

As increasingly difficult structural problems are being undertaken, the challenge of solving structures with lower quality data is also increasing. Larger proteins and large complexes have greater heterogeneity, membrane proteins have decreased solubility, and many interesting structural targets contain natively disordered regions. Each of these problems often leads to lower quality crystals and poor or low resolution diffraction from these crystals (Brunger 2005). Considering X-ray crystallography can achieve the highest resolution of all means of determining macromolecular structures, data considered low resolution for X-ray crystallography may still provide important insights into protein structures. If large conformational changes occur within the protein, low resolution data (data in the range of 5-20 Å) may be able to detect this. The position of major structural features and differences between two similar proteins or complexes can often be determined at low resolution(Rossmann, Bernal et al. 2001).

Molecular replacement as a means to solve the phase problem in crystallography requires the least amount of diffraction data. A single quality dataset may suffice for the solution of a macromolecular structure (Dodson 2008). Molecular replacement uses a structurally similar previously determined structure, or a model of the expected protein structure to determine the phases of the X-ray diffraction dataset. Conceptually determination of the phases from a model structure is best understood by the original methodology used for molecular replacement. Approximate phases are achieved by correlation of a calculated patterson function from the model and the experimental patterson function. A six dimensional search with three rotational and three translational degrees of freedom is

used to match the calculated and experimental patterson functions and orient the model within the unit cell of the crystal. Then the phases are calculated from the positioned model and combined with the experimental amplitudes to generate the initial electron density maps. Modern maximum likelihood methods are similar in that the comparison between the model and the experimental data are the metric for determining the best solutions, but instead of a patterson correlation function, the statistical likelihood that each test solution is correct is used (Read 2001).

At low resolution the number reflections in the data set is decreased and therefore less data is available for correlation to calculated data. In addition in the range of 8-15 Å the contribution by the bulk solvent diffraction to the overall diffraction intensities is at a maximum(Guo, Blessing et al. 2000). The differentiation of correct molecular replacement solutions from incorrect solutions by patterson correlation becomes more difficult at lower resolution compared to high resolution due to the inter-atomic vectors becoming more prominent than the intra-atomic vectors. Intra-atomic vectors are necessary for correct rotation solution of the model that is subsequently the basis for the translational positioning. The combination of increased error in the experimental data, and greater error in the patterson correlation functions can lead to difficulties in finding correct molecular replacement solutions.

Improving the likelihood of determining correct phases by molecular replacement continues to be challenging. Of the parameters that directly contribute to high quality molecular replacement phasing few are easily improved. The inherent noise within a

particular dataset cannot easily be changed especially if crystals are difficult to obtain. Finding better ways to score patterson correlation functions (or maximum likelihood functions) is possible but difficult. There is however, one way improvement of molecular replacement searches can be easily achieved. The one major contribution to error within a molecular replacement search that the crystallographer has easy control over is the choice and development of the molecular replacement model.

Model choice is well established to be an essential part of the scientific process for the molecular replacement experiment. Improved bioinformatics algorithms and homology modeling can have great effect on finding a correct molecular replacement solution and the quality of initial electron density maps. The attributes of a good molecular replacement model at high resolution have been well established and in all cases the best model is the exactly correct structure. Given that at low resolution no interesting structural questions will have an exactly correct model the question remains as to what empirically constitutes a good molecular replacement model for low resolution datasets.

To determine the attributes of a good molecular replacement model at low resolution we can address the solution of the structure of the minimal self-assembling region clathrin. As a case study this structure provides some unique opportunities, first the data aside from being low resolution are of excellent quality (Table 1, Chapter 1). Secondly, the solved structure has very good refinement statistics for the resolution. Finally, the structure contained multiple subunits adopting different conformations, which provided a challenging molecular replacement situation.

Methods

Data collection and processing

Data was collected from a single crystal in two passes. First, diffraction images were collected to a high resolution of 8.3Å. Then the detector distance was increased and the exposure time optimized for accurately collecting the low resolution reflections without overloading the detector. All data was collected at the Advanced Light Source beamline 8.3.1 at Lawrence Berkeley National Laboratories equipped with an ADSC Q315 detector. Data was integrated with Denzo and scaled with Scalepack (Table 1, Chapter 1). The spacegroup was determined to be $I4_122$ with cell dimensions a = b = 228.5 Å and c = 710.3 Å.

Molecular replacement model development

Molecular replacement searches were performed using Phaser. The number of allowed clashes was between 10-50 for each search roughly depending on the expected completeness of the model. The estimated sequence identity was set to 90%. Initial forays into molecular replacement relied heavily on the previously determined structure of the proximal leg of clathrin to 2.9Å. Searching using simply this structure provided no solutions though this search may not have been completely exhaustive and may have used programs other than Phaser for molecular replacement searches. Additionally a false trimer of proximal legs yielded no good results, again this may have suffered from similar problems as the proximal leg attempts.

To develop a more accurate model for exhaustive molecular replacement searches more rigorous model building was undertaken. Using the proximal leg as a template, each of the separate repeated domains that make up the clathrin leg, called the clathrin heavy chain repeat (CHCR), was modeled based on a large multiple sequence alignment of CHCRs from many species. Each CHCR model was then aligned with its respective region based on the low resolution cryoEM structure determined by Fotin *et al*. (1XI4). In this way we built nearly the full leg. An alpha helix was used as a template for the trimerization domain and attached to the end of the modeled CHCRs. At each CHCR junction the loops were modeled using a previously described modeling method.

Molecular replacement searches were performed using either residues 1074-1630, 1280-1630, or the same regions as a poly-alanine sequence, or with randomized side chains of this model. In addition these models were used with B factors of all atoms set to either 99.99 or 50.

Results

Many initial models based simply on the known structure of clathrin did not provide results by molecular replacement suitable for complete structure solution with low resolution data. Even if the molecular replacement search with the proximal fragment of clathrin correctly identified a solution it is likely the low resolution of the data, 8.3Å, prevented the proximal leg from rising above the noise during scoring of molecular replacement solutions. Even with false trimers no solutions were identified (Peter Hwang

unpublished data). This may have simply been due to the amount of error within the falsely trimerized proximal legs.

Due to the low resolution of the crystal data, initial searches with the full 1074-1630 model were attempted with a poly-alanine version of the model. The poly-alanine model was used to minimize model bias but all the residues 1074-1630 were used in an attempt to maximize signal during the molecular replacement searches. This initial poly-alanine model also had the default B factor value of 99.99 for each atom. This model provided low Z-scores on the order of 4-6 for the Phaser translation search (Table 2). Scores in this range are unlikely to be correct solutions. Though the 1074-1630 seems like a reasonable model the trival answer for no correct solutions determined is due to altered protein conformations leading to increased packing clashes. The possibility remains that the correct solution was determined but then subsequently removed due to too many clashes with crystal symmetry mates.

The high solvent content of the crystals hindered development of a correct molecular replacement search. The expected range of solvent content for protein crystals is a normal distribution of approximately 35% to 65%. The Matthew's coefficient of the crystal based on the normal distribution suggested two molecules per asymmetric unit. Searching for two clathrin hub molecules with any of the models provided poor results. The reason for this is clear in retrospect. The unique shape of the clathrin molecule created a very "open" crystal lattice and allowed the formation of a protein crystal that

was 85% solvent. There are only 11 different proteins crystallized with a solvent content of 85% or higher, and these include viruses and ribosomes.

The next molecular replacement attempts used the complete model N-terminally truncated at residue 1280 and contained all side chain information. The side chain positions had been energy minimized and the loops at junctions between the CHCR were optimized as described in the methods. The B factors of this model were set to the default 99.99. Three solutions were returned after the molecular replacement search with Phaser. These scored highly with a rotation function Z-score of 4.1 and a translation function Z-score of 10 (Table 2). The log likelihood gain was 70 for both rotation and translation (Table 2). There were no clashes between backbone trace atoms with this correct solution. Initial maps from this solution were highly correlated with the model and additional density corresponding to a large portion of the clathrin heavy chain leg was visible. Small regions of clathrin light chain density were also resolved.

To test whether the it was the truncation of the model or the addition of side chain information that contributed to finding the correct solution, a poly-alanine model truncated at residue 1280 was used for the molecular replacement search. This model, though only varying from the model that found the correct solution by the lack of side chain information, produced no good results. The rotation function score was significantly lower with a Z-score of 2.5 (Table 2). The translation function Z-score was similar to the correct solution at 10.1 (Table 2). This suggested that the correct solution may have been found but during the packing tests the best solutions with this truncated

poly-alanine model all had large numbers of clashes between trace atoms. The highest scoring solution with the truncated poly-alanine model had 128 clashes between trace atoms (Table 2).

It appeared that while N-terminally truncating the model at residue 1280 may have been important, the side chain information was playing some role in the determining the correct solution. At the low resolution of the dataset, the side chain information is completely unresolved. In addition, the side chains in the model were expected to be incorrectly positioned. This suggests the side chain information was contributing to the molecular replacement search by generally contributing to scattering but not necessarily to precisely correct phases. The increase in scattering atoms in the all atom 1280-1630 model compared to the polyalanine 1280-1630 model may have contributed to finding the correct solution by increasing the signal above the bulk solvent scattering noise. Without the side chain information the calculated scattering of the model was too weak for high correlation between the calculated and experiment diffraction intensity.

To test whether side chains were contributing significantly by their correct positions or simply by increasing the amount of scatting atoms within the model, the 1280-1630 model with random side chains were used in the phaser molecular replacement search. This search found a solution that was overall correct but closer to solutions within the noise. The rotation function Z-score was 4.31 and the translation function Z-score was 10.17 (Table 2). The log likelihood gain was 53.38 but there were 2 packing clashes

(Table 2). This sidechain information apparently contributed primarily by increasing scattering of the model, but using correct side chains positions was of course better.

Since parameters within the model that affected calculated scattering seemed to affect the quality of the molecular replacement solutions, a test of how the model B factors contributed to molecular replacement solution was performed. The 1280-1630 model with side chains or without side chains were used for molecular replacement searches with all atom B factors set to either 50 or 99.99. From previous experiments it was known that side chains contributed positively to finding correct solutions and that the 1280-1630 polyalanine model with B factors of 99.99 did not provide correct solutions. Molecular replacement with the 1280-1630 polyalanine with B factors of 50 provided a solution with rotational and translational Z scores of 3.8 and 13.7, respectively (Table 2). The log likelihood gain was 59 and there were no packing clashes (Table 2). This solution is correct based on the determined structure and demonstrates that in addition to side chain information, B factors have significant contribution to finding molecular replacement solutions at low resolution. Lower B factors in the polyalanine model may have again increased calculated scattering (or confidence in the calculated scattering) and contributed by differentiating the protein scattering from the high bulk solvent noise.

Interestingly decreasing the B factors of the randomized side chain model containing residues 1280-1630 has a negative effect on the ability of this model to lead to a correct solution. Random side chains and B factors of 50 resulted in a rotation function Z-score of 3.4 and a translation function Z-score of 9.9 (Table 2). Both of these numbers are

slight worse than for the same model with B factors of 99.99. The log likelihood gain, on the other hand, was improved to 62 (Table 2). In addition the best solution has 12 packing clashes instead of 2 for the same model with B factors set to 99.99 (Table 2). This suggests that lowering the B factors in an incorrect model may be detrimental to finding correct molecular replacement solutions.

Discussion

The contribution of the model in molecular replacement phasing of X-ray crystallographic data cannot be understated. The more closely all model parameters match the experimental data the greater likelihood of determining correct phases for the crystallographic data. Additionally, the model used in molecular replacement is likely the easiest parameter to change or improve on. Incorporating all known structural information into a molecular replacement model maybe allow the determination of structures that are otherwise challenging.

Low resolution crystallographic data has generally been challenging to work with. The low resolution means there is less data to correlate with structural models, thereby decreasing the both the ease of phasing and the accuracy of the final model. Low resolution data also has increase bulk solvent scattering contributing to the scattering intensity causing increased noise in the data that one collects. Since, within molecular replacement, the model is the easiest parameter to adjust to help overcome inherent problems with low resolution data, this work sought to determine features of a model that positively contributes to correct phasing of low resolution data by molecular replacement.

The use of all atom models and the decreasing of the molecular replacement model B factor were found to positively contribute to phasing under certain circumstances. While decreasing the B factors of an incorrect model negatively contributed to determining molecular replacement solutions.

I propose that the positive contribution of both side chains and decreased B factors was due to increased confidence in the calculated scattering from the model. This increased scattering (or increased confidence in scattering) was necessary to boost the signal in the molecular replacement search above the higher noise in the low resolution dataset. Compared to high resolution data where the most correct solution with the least bias is a good start for phasing, at low resolution finding any molecular replacement solution that is even close to correct is the challenge. The contribution to scattering by bulk solvent decreases the contrast between the solvent and the protein. This suggests even though low resolution data in all data sets provides strong positioning of the model during molecular replacement by defining the protein envelop, changes to the molecular replacement search model that increase contrast will be most important with low resolution data.

References

Brunger, A. T. (2005). "Low-resolution crystallography is coming of age." Structure ***13****(2): 171-2.*

Dodson, E. (2008). "The befores and afters of molecular replacement." Acta Crystallogr D Biol Crystallogr ***64****(Pt 1): 17-24.*

Guo, D. Y., R. H. Blessing, et al. (2000). "Bulk-solvent correction in direct-methods phasing." Acta Crystallogr D Biol Crystallogr ***56****(Pt 4): 451-7.*

Read, R. J. (2001). "Pushing the boundaries of molecular replacement with maximum likelihood." Acta Crystallogr D Biol Crystallogr ***57****(Pt 10): 1373-82.*

Rossmann, M. G., R. Bernal, et al. (2001). "Combining electron microscopic with x-ray crystallographic structures." J Struct Biol ***136****(3): 190-200.*

Table 2. Molecular replacement scoring at low resolution of five different models

Name	B factors	Rotation function Z-score	Translation function Z-score	Log Likelihood Gain	Packing clashes
All atom 1280-1630	99.99	4.1	10.0	70	0
Poly Ala 1280-1630	99.99	2.35	10.1	ND*	128
Random side chain 1280-1630	99.99	4.31	10.17	53.38	2
Poly Ala 1280-1630	50.0	3.8	13.7	59	0
Random side chain 1280-1630	50.0	3.4	9.9	62	12

*ND, not determined

Conclusions

Determining the structure and functions of proteins is a fundamental role in many fields of bioscience research. X-ray crystallography and *in vitro* biochemistry play critically important roles in determining mechanisms of protein function in cells and adapting proteins for technological uses.

Two of the crystal structures determined in this work are at the forefront of protein structure research. First the low resolution structure of the minimal self assembling region of clathrin addresses many of the difficulties of solving structures of large molecular complexes. The methods used in this dissertation research to determine this structure are generally applicable and may well provide a means for solving other structures. The structure of RcdA, an adaptor type protein in bacteria, elucidated a novel fold. The three helix bundle is a common motif, represented in may functional classes of proteins. Rarely, however, have three helix proteins dimerized and never has a three helix bundle protein dimerized in the way RcdA dimerizes. This novel fold will likely be the prototype of a new class of proteins and will serve as a template for understanding the structure and mechanism of proteins in this class.

The biochemical experiments described within this work are necessary to address the confomational flexibility and dynamics of these proteins and understand how these attributes lead to function. By using a variety of structural and biochemical methods many insights beyond the simple interpretation of experiments can be achieved.

Virtually all this work opens many more possibilities for future research. The most obvious are obtaining higher resolution structures of those proteins described here. In particular a high resolution structure clathrin complexed with clathrin light chain is still necessary to fully determine the mechanisms involved in clathrin's function. The structure of full length Hip proteins or large fragments in alternate conformations would be a dramatic achievement. Similarly, a structure by single particle reconstruction of the Hip proteins with various bound ligands would be very informative. Should high ambitions prevail, an electron tomographic reconstruction of a budding vesicle could allow the fitting of all known protein structures into a tomographic density maps and effectively build a huge complex structure at high resolution. More straightforward experiments like determining the crystal structure of clathrin light chain bound to Hip proteins may still provide significant insights into these proteins functions and be useful tools for cell biological studies.

Overall though there are many creative and interesting options available for further study, I hope this work took many steps toward a complete understanding of clathrin mediated endocytosis.

Appendix 1. The Disordered C-terminus of RcdA Is Necessary for Efficient CtrA Proteolysis in *Caulobacter crescentus*

In collaboration with James A. Taylor and Kathleen Ryan.

Introduction

A hallmark of eukaryotic differentiation is asymmetric cell division to create distinct cell types. In bacteria, the prime example of differentiation coupled to the cell cycle is found in *Caulobacter crescentus*. *Caulobacter* divides asymmetrically to produce morphologically distinct daughter cells. The stalked cell (ST) can immediately reinitiate chromosome replication, while the swarmer cell must first differentiate into a stalked cell before beginning a new cycle of chromosome replication and division (reviewed in (1)). Underlying the differences between the daughter cells is the activity of a master transcriptional regulator, CtrA (2-3). When phosphorylated, this response regulator directly controls the expression of 95 genes (4), as well as binding to five sites in the chromosomal replication origin preventing replication initiation (5). Thus, for S-phase to occur, CtrA activity must be eliminated from the developing SW cell. This is achieved redundantly by dephosphorylation and directed proteolysis (3).

CtrA is degraded specifically at the G1-S phase transition by the AAA+ protease ClpXP (6). Within this complex, the ClpP subunits have protease activity, while the ClpX

subunits mediate substrate recognition and unfolding (reviewed in (7)). CtrA phosphorylation and proteolysis are coordinately regulated by the same phosphorelay (8), in which the hybrid histidine kinase phosphorylates the histidine phosphotransferase ChpT. ChpT can then pass a phosphoryl group to CtrA or to the single-domain response regulator CpdR. When the phosphorelay is inactive, unphosphorylated CpdR promotes the degradation of CtrA (9). In addition to ClpXP and CpdR, at least one other protein of unknown function, RcdA, is required for CtrA proteolysis. Unlike CpdR, RcdA is specifically required for CtrA degradation by ClpXP and does not influence the degradation of other substrates such as McpA (10). RcdA is a hypothetical protein that is conserved specifically in alpha-proteobacteria (11), but whose fold cannot be predicted from its primary sequence.

Despite the fact that ClpXP, CtrA, RcdA, and CpdR are all cytoplasmic proteins, CtrA degradation occurs at a specific location in the *Caulobacter* cell (12-13). At the nascent stalked pole in differentiating swarmer cells, and at the stalked pole of late predivisional cells, all five proteins accumulate transiently and CtrA is proteolyzed (9-10). There is a linear dependency among these proteins for polar localization, such that CtrA requires the presence of RcdA, RcdA requires ClpXP, and ClpXP requires CpdR (9-10). Co-immunoprecipitation from *Caulobacter* lysates demonstrates that CpdR associates with ClpXP, and RcdA associates with both ClpXP and CtrA *in vivo* (9-10).

Adaptor proteins can enhance the degradation of individual substrates or classes of substrates by ClpXP (reviewed in (14)). For example, incomplete proteins synthesized from damaged mRNAs on stalled ribosomes are tagged with the SsrA peptide, and these tagged proteins are recognized and degraded by ClpXP (15-16). The dimeric adaptor

protein SspB binds simultaneously to the ssrA peptide and to ClpX to increase the local substrate concentration near the protease (17-18). Structural studies reveal SspB contacts ClpX through its intrinsically disordered C-terminus (19), while a conserved cleft in the globular region of the protein binds the SsrA peptide tag (20).

It was proposed that RcdA could function as an adaptor for CtrA, receiving cell cycle information and determining when CtrA is a target for degradation by ClpXP (10). However, while RcdA is required for cell cycle-regulated proteolysis of CtrA *in vivo*, the rate of CtrA proteolysis by purified ClpXP *in vitro* is unaffected by the addition of RcdA, and RcdA does not interact stably with ClpX (21). Thus RcdA seems not to function as a direct adaptor between CtrA and ClpXP, and its role in CtrA proteolysis remains unclear.

We have crystallized RcdA and determined the structure to 2.9Å. RcdA forms an elongated crescent-shaped dimer with three flexible disordered regions. We created a series of structure-based mutants and tested their ability to restore CtrA proteolysis to an *rcdA* deletion strain. We show here that the intrinsically disordered C-terminal region of RcdA is required for efficient cell cycle-regulated CtrA degradation, as well as the accumulation of CtrA and RcdA at the cell pole.

Results

RcdA structure. Crystallization of RcdA yielded rod-shaped tetragonal crystals, and the structure was determined using multiple anomalous dispersion from selenomethionine-substituted crystals (Fig. 1A, Supporting Information (SI) Table 1, PDB NNNN). The asymmetric unit contained one RcdA homodimer. RcdA dimers form a crescent shape with approximate dimension of 63Å × 41Å × 21Å. The radius of curvature of the

concave side is approximately 36 Å. At each end of the crescent is a flat, paddle-shaped region containing several highly conserved residues (Fig. 1D, circled). Each monomer of RcdA consists of a three-helix bundle with three intrinsically disordered regions (Fig. 1B-C).

The first twenty amino acids (aa) of each monomer are disordered (Fig. 1B, gray dashed line), and the first defined residues are a random coil leading into helix 1 (Fig. 1B, arrows). Helix 1 has a distortion in helical geometry associated with a Gly-Ala-Gly sequence (aa 46-48) just before the turn leading to helix 2. Helix 2 contains a series of hydrophobic residues starting in the dimerization interface and continuing along its solvent-exposed side into the paddle region (Fig. 1D, arrows). Following helix 2 is an unstructured internal loop of 29 residues (Fig. 1B-C, red dashed lines), initiated by a Pro-Pro motif (aa 93-94) and containing two more prolines. Because of the disordered internal loop, RcdA has an unusual topology for three-helix bundles with helices 2 and 3 packing parallel to each other (Fig. 1C). At the end of helix 3 is a short section of random coil and approximately 22 unstructured residues (Fig. 1B-C, green dashed lines).

RcdA dimerization buries a large hydrophobic patch (~2000 $Å^2$), composed primarily of residues Met34, Val37, Met75, Ala78, Leu82, and Met92 from each monomer in a two-fold symmetric orientation. The helical interfaces from each monomer pack at an angle near 90º. This is significantly different from the more parallel interaction of BAR domains, the best known three-helix bundle dimers (22). Database searches using the protein structure comparison service SSM (23) yield many three-helix bundle proteins, but none that dimerize similarly to RcdA. Dimerization in solution was confirmed using calibrated gel filtration (Fig. 2D). A large acidic patch on the convex

side is the most prominent electrostatic feature of the RcdA dimer (supporting information (SI) Fig. 4).

The disordered C-terminus of RcdA is required for efficient CtrA degradation. RcdA is required for CtrA degradation *in vivo*, but the molecular function of RcdA is unknown. We sought to identify regions of RcdA that could bind to other proteins involved in CtrA degradation. An alignment of RcdA with its nearest homologs in different genera (SI Fig. 5) showed that much of the sequence conservation in RcdA lies along the dimer interface and is unlikely to contribute to interactions with other proteins. However, extended paddle regions at either end of the RcdA dimer (Fig. 1D, circled) also contain highly conserved sequences, including a surface-exposed hydrophobic patch composed of L61, A64 and M68 (Fig. 1D, arrows). We introduced the point mutations L61S and A64S (Fig. 1E, black arrows) into pPM53, which encodes P_{xyl}-*rcdA::GFP*, to determine the effects of these substitutions on CtrA proteolysis during cell cycle progression. We integrated the resulting plasmid (pJT41) into the *xylX* locus (24) of *ΔrcdA* strain KR899 (10) to create a strain expressing the mutant protein RcdApatch (KR2268). In KR2268, as in the strain *ΔrcdA/xylX::rcdA::GFP* (KR2209), both CtrA and the chemoreceptor McpA were degraded normally during the cell cycle (schematized in Fig. 2A), as observed Western blotting of samples from synchronous cultures (Fig. 2B and data not shown).

Next, we hypothesized that protein-protein interactions could be mediated by conserved charged residues (D45, R49, K53 and R57, Fig. 1E, red arrows) that protrude from edges of the paddle regions. Using site-directed mutagenesis of pJT41, each of these

residues was replaced by alanine in the polypeptide already containing the *rcdApatch* mutations to yield pJT50. We integrated pJT50 at the *xylX* locus of KR899 to create a mutant expressing the protein RcdDpaddle (KR2269) and again observed normal CtrA proteolysis during swarmer cell differentiation (Fig. 2C). Thus, these conserved residues on the surface of the paddle region are not required for cell cycle-regulated CtrA degradation.

The C-terminus of RcdA beyond residue 150 is unstructured in the crystal, although it contains several amino acids conserved in RcdA homologs (SI Fig. 5). We deleted residues 151-169 to create RcdAΔC and expressed this protein in *ΔrcdA* cells from the *xylX* locus (KR2210). During synchronous growth, KR2210 cells failed to degrade CtrA during swarmer cell differentiation (Fig. 2D). In the same experiment, McpA was proteolyzed at the expected time, indicating that passage through the cell cycle was unperturbed. A Western blot probed with anti-RcdA serum (10) demonstrated that RcdA-GFP in KR2209 and RcdAΔC in KR2210 were expressed at similar levels (Fig. 2E). We used size exclusion chromatography to determine if deleting the C-terminus disrupts the gross structure of RcdA (Fig. 2F). Purified RcdA and RcdAΔC eluted from a Superdex-200 column at 15.4 ml and 16.1 ml, respectively, corresponding to molecular masses of 42.1 kDa and 34.1 kDa. These values are close to the calculated dimer masses of RcdA (38.1 kDa) and RcdAΔC (34.3 kDa), so it is likely that RcdAΔC folds and dimerizes correctly.

We performed pulse-chase assays on unsynchronized cultures to quantify overall rates of CtrA proteolysis. CtrA half-lives were determined to be 204 ± 23 min, 50 ± 5 min, and 89 ± 12 min in strains *ΔrcdA*, *ΔrcdA/xylX::rcdA* and *ΔrcdA/xylX::rcdAΔC*,

respectively. The presence of RcdAΔC increases the rate of CtrA proteolysis ~two-fold over *ΔrcdA*, but this rate is also ~two-fold slower than CtrA degradation in cells containing wild-type RcdA. This modest effect on the degradation rate is enough to prevent the complete clearing of CtrA during swarmer cell differentiation (Fig. 2D).

CtrA activity is regulated by redundant mechanisms, proteolysis and phosphorylation. Wild-type cells arrest in the G1 phase of the cell cycle if they overexpress a CtrA variant that both mimics the phosphorylated state and cannot be degraded (3). We asked if the slowed CtrA degradation mediated by RcdAΔC is sufficient to allow cell cycle progression in the presence of the phosphomimetic protein CtrA-D51E (SI Methods). Wild-type cells overexpressing CtrA-D51E have normal morphology and DNA content, because CtrA activity is still modulated by proteolysis (SI Fig. 6A-C), whereas *ΔrcdA* cells overexpressing CtrA-D51E become filamentous and predominantly contain only one chromosome (SI Fig. 6D-F). Coexpression of wild-type *rcdA* in this strain restores normal morphology and DNA content (SI Fig. 6G-I), while coexpression of *rcdAΔC* causes filamentation and G1 arrest similar to cells lacking all RcdA protein (SI Fig. 6J-L). Thus, the disordered C-terminal peptide of RcdA plays a key role in cell cycle-regulated CtrA proteolysis.

The RcdA C-terminus is required for polar localization of RcdA and CtrA. An RcdA-GFP fusion protein is functional in CtrA degradation (Fig. 2B) and exhibits dynamic localization during the cell cycle (10). In developing swarmer cells, RcdA accumulates at the nascent stalked pole during CtrA proteolysis. In early predivisional cells, RcdA migrates to the midcell, then in late predivisional cells, RcdA accumulates at

the pole in the stalked compartment, where CtrA proteolysis occurs (Fig. 3A-B). When we examined the subcellular location of RcdAΔC-GFP (KR2211), we found that it does not localize at the cell pole or the midcell of *Caulobacter* (Fig. 3C-D). Western blots probed with anti-RcdA and anti-GFP antibodies demonstrated that RcdAΔC-GFP is not cleaved *in vivo* and is expressed at levels similar to RcdA-GFP (data not shown).

RcdA is required for the polar localization of CtrA, and it can be co-immunoprecipitated from *Caulobacter* lysates along with CtrA and ClpX (10). To determine if CtrA can accumulate at the cell pole even when RcdA does not, we expressed YFP-CtrARD+15, a fusion protein which contains the signals sufficient for CtrA localization and proteolysis (12), in cells expressing either RcdA (KR2256) or RcdAΔC (KR2247). YFP-CtrARD+15 is properly localized in cells containing wild-type RcdA (Fig. 3E-F), but is diffuse in cells containing RcdAΔC (Fig. 3G-H). Thus the disordered C-terminus of RcdA is required for the polar localization of CtrA.

To determine if the C-terminal peptide of RcdA is sufficient for polar or midcell localization of an unrelated passenger protein, we fused all of the residues following helix 3 (aa 143-169) to the C-terminus of YFP. We observed no localization of this fusion protein in wild-type cells (Fig 3I-J), so the C-terminal peptide of RcdA is not sufficient to direct RcdA to the pole or midcell.

RcdA and CpdR together do not stimulate ClpXP-mediated proteolysis of CtrA. Purified ClpXP degrades CtrA *in vitro*, without additional proteins, and RcdA does not increase the rate of degradation (21). Since RcdA and CpdR are each necessary for regulated CtrA proteolysis *in vivo*, we asked if CpdR and RcdA together could stimulate

CtrA proteolysis *in vitro* (SI Methods). *Caulobacter* ClpXP degraded CtrA at a rate of 0.3 molecules/$ClpX_6$·min. When CpdR and RcdA were added (~2 CpdR:1 $RcdA_2$:1 CtrA), the rate of proteolysis was unchanged (SI Fig. 7). Thus, RcdA and CpdR together do not function as a proteolytic adaptor for CtrA.

Discussion

Using structure-based mutations, we have shown that the intrinsically disordered C-terminus of RcdA is required for efficient CtrA degradation *in vivo* as well as subcellular localization of CtrA and RcdA itself. Several amino acid substitutions at conserved positions in the paddle region of the folded RcdA domain allowed CtrA to be degraded properly during the cell cycle. It is possible that the paddle regions are not actually surface-exposed, because the disordered N- and C-termini of RcdA exit the folded structure in this area. We anticipate that some surface residues in the folded portion of RcdA mediate interactions with proteins involved in CtrA proteolysis. Future mutagenesis studies will address the roles of other surfaces in CtrA localization and proteolysis, including the negatively charged convex face (SI Fig. 4) and the dimerization interface.

Intrinsically disordered proteins often play important roles as hubs of protein interaction networks (25). Intrinsic disorder may facilitate adaptability to multiple binding partners and the ability to maintain large interaction surfaces using fewer amino acid residues. It is suggested that bacteria have fewer proteins with intrinsic disorder than eukaryotes, possibly due to protease accessibility or protein interaction networks that do not required the complexity of eukaryotic organisms (26). Interestingly, the interaction

between SspB and ClpX is mediated by an intrinsically disordered SspB peptide (19). RcdA contains both well-structured three-helix bundles and multiple disordered regions. The combination of order and disorder may allow RcdA the flexibility to mediate multiple protein interactions while maintaining a compact protease-resistant structure.

Superficially, RcdA resembles SspB, as both proteins form stable head-to-head dimers, and each protein contains an intrinsically disordered peptide at the C-terminus that is required to enhance degradation of a substrate. However, functional studies show that SspB and RcdA do not share a common molecular mechanism. The C-terminal peptide of SspB binds to the N-terminal domain of ClpX (19, 27), while the SsrA tag binds in a conserved cleft within the structured region of SspB (20, 28). SspB-mediated tethering of SsrA-tagged substrates to ClpX increases the local substrate concentration and allows more efficient proteolysis (17). In contrast, RcdA does not increase the rate of CtrA proteolysis by ClpXP or bind directly to either ClpX or CtrA (21) and our unpublished data). Nonetheless, RcdA is required for regulated CtrA proteolysis during the *Caulobacter* cell cycle, and RcdA can be coimmunoprecipitated from *Cauolobacter* lysates along with ClpX and CtrA (10). Direct binding partners of RcdA in this complex or at the cell pole have not been identified.

Integrating structural information about RcdA with previous studies, is it possible to distinguish among models of RcdA function? Chien, *et al*. (2007) suggested that RcdA could block the action of a negative regulator of CtrA proteolysis. CtrA may bind to a factor which protects it from ClpXP-mediated degradation during part of the cell cycle. In cells preparing for S phase, this factor would be removed by RcdA, allowing CtrA to be proteolyzed. Our data do not exclude this model. RcdAΔC can promote some CtrA

degradation without being localized to the cell pole. Rather than removing the inhibitor of CtrA proteolysis while colocalized with ClpXP, RcdAΔC could remove the inhibitor at all points in the cell. CtrA could then be proteolyzed, or it could rebind the inhibitor before reaching ClpXP for degradation.

CtrA consists of an N-terminal receiver domain and a C-terminal DNA-binding domain. The DNA-binding domain of CtrA is degraded by purified ClpXP *in vitro*, at a rate comparable to full-length CtrA (21). These findings suggested that an inhibitor of CtrA proteolysis could bind to the N-terminal receiver domain before being removed by RcdA (21). If this model is correct, then the DNA-binding domain should be degraded more quickly *in vivo* than full-length CtrA, since it lacks the region necessary for binding to an inhibitor. However, the DNA-binding domain of CtrA is not degraded during swarmer cell differentiation (12), and pulse-chase analyses of unsynchronized cultures reveal that this domain is degraded at least two times more slowly than full-length CtrA (data not shown). These results suggest that the receiver domain of CtrA plays a positive role in proteolysis, rather than binding to an inhibitor.

In another model, RcdA could facilitate CtrA degradation by blocking the recognition of another class of substrates also degraded by ClpXP (21). Since there are at most ~1000 ClpXP complexes per cell (29), CtrA proteolysis may be blocked at specific times because ClpXP is occupied with other substrates. When RcdA is colocalized with ClpXP at the cell pole or midcell, it could block the access of some substrates, but not others, to the protease. Notably, in this model, CtrA is recognized for degradation without a positively acting adaptor protein. Although our current studies do not preclude this model, and substrate choice could affect CtrA stability, we think it unlikely that this is the

only mechanism regulating the timing of CtrA degradation. First, while cell cycle-regulated CtrA proteolysis is abolished in *ΔrcdA* cells, degradation of the chemoreceptor McpA by ClpXP is not affected (10). If RcdA were to promote CtrA degradation solely by inhibiting recognition of an unrelated class of proteins, we would anticipate that McpA proteolysis, which occurs at the same time in the cell cycle, might be affected by the same mechanism.

Second, even if CtrA cannot be proteolyzed by ClpXP, it can still localize to the pole at the appropriate time in the cell cycle (13). Changing the C-terminal AA residues of CtrA to DD blocks degradation *in vivo* and *in vitro* (3, 21), but does not prevent the transient accumulation of CtrA at the cell pole (13). This suggests that CtrA binds to another protein at the pole in a cell cycle-dependent manner, even if CtrA cannot be engaged by ClpXP. Thus, we hypothesize that CtrA proteolysis involves a positively acting localization factor *in vivo*, even though no adaptor protein is strictly required for ClpXP-mediated degradation.

We favor a third model, in which RcdA modulates the activity of a bona fide adaptor or localization factor for CtrA. The adaptor may be unable to bind ClpX or CtrA until contacted by RcdA at the cell pole. In this model, the partial stimulation of CtrA proteolysis by RcdAΔC would be due to transient interaction of RcdAΔC with the polar degradation complex. We speculated that the CtrA adaptor could be CpdR, since this protein is required for CtrA proteolysis *in vivo*, colocalizes with the protease and substrate, and can be coimmunoprecipitated with ClpXP (9). However, we observed no stimulation of CtrA proteolysis by ClpXP when both RcdA and CpdR were present in the reaction. We therefore believe that other, unidentified proteins play a role in regulating

CtrA proteolysis *in vivo*. We are currently using genetic screens and yeast two-hybrid assays to identify components that interact with either RcdA or the N-terminal receiver domain of CtrA.

Methods

Bacterial Strains, Media and Plasmids. Strains and plasmids used are listed in SI Table 2. All experiments were performed using derivatives of *Caulobacter crescentus* strain CB15N grown to mid-logarithmic phase. CB15N strains were grown in peptone-yeast extract (PYE) or M2G medium (30) at the indicated temperatures, supplemented with antibiotics as required. Xylose was added to PYE (0.03%) and M2G (0.3%) media where indicated. *E. coli* strains were grown in Luria Broth (LB) at 37°C, supplemented with antibiotics as required. Mutations were introduced using QuikChange (Stratagene), with pPM53 or pES51 as template (oligonucleotide sequences available on request). The mutations introduced into pPM53 to generate pJT50 encoding *rcdApaddle::GFP* were D45A, R49A, K53A, R57A, L61S, and A64S. To generate *rcdAΔC*, a BamHI site and a stop codon were introduced after codon 150 of *rcdA*. The BamHI site was used to fuse *rcdAΔC* to *GFP*. pJT52 was generated by moving P_{xyl}- *YFP::ctrARD+15* from pEJ146 into pMR20.

Protein expression and purification. For crystallization, *E. coli* BL21 harboring pES53 were grown in LB to an OD_{600} of 0.6-0.8 and induced to express His6-RcdA with 1 mM isopropyl-thiogalactoside (IPTG) for 4-5 hours. The cell pellet was harvested by centrifugation and frozen at -80ºC. Cell pellets were resuspended in binding buffer

(20mM Tris pH 8.0, 500mM NaCl, 5mM imidazole) and lysed by sonication. Lysates were cleared by centrifugation at 40,000*g* for 45 min, and the supernatant was incubated with Talon affinity resin (Clontech). After 30 min, Talon resin was pelleted, resuspended in fresh binding buffer, and transferred to a column. The resin was washed with 60 ml of wash buffer (20mM Tris pH 8.0, 500mM NaCl, 25mM imidazole), followed by elution of His6-RcdA with 20mM Tris pH 7.0, 350mM imidazole. The His6 tag was removed by incubating the eluted protein with 20 units of thrombin (Sigma) overnight at 4°C. Cleaved RcdA was isolated by gel filtration on a Superdex 200 column in 10mM HEPES pH 7.0, 100mM NaCl, then concentrated to 6 mg/ml. Selenomethionine derivatives of RcdA were obtained using the same protocol, except that cell growth and protein expression occurred in M9 minimal medium supplemented with 50 mg/l selenomethionine as previously described (31).

For calibrated gel filtration and proteolysis assays, RcdA and RcdAΔC were purified as follows. BL21 Tuner cells were grown to an OD_{600} of 0.4 and induced to express His6-RcdA or His6-RcdAΔC with 0.5 mM IPTG for 3 hrs. Cells harvested by centrifugation were frozen at -80ºC. Cell pellets were resuspended in buffer A (50 mM Tris-HCl pH8.2, 100 mM KCl, 1 mM $MgCl_2$, 1 mM DTT, 10% glycerol, 20 mM imidazole) and lysed for 1 hr on ice by the addition of lysozyme (1 mg/ml). Cells were disrupted by sonication and the lysate cleared by centrifugation at 12,000*g* for 20 min at 4ºC. Ni-NTA agarose was added to the lysate, and proteins allowed to bind at 4ºC. The slurry was transferred to a column and washed three times with 30 ml buffer A. Ni-NTA agarose was resuspended in 5 ml buffer A and incubated overnight with 15 U thrombin. Cleaved RcdA or RcdAΔC was eluted from the resin with two additional washes of 10 ml

buffer A and added directly to a mono-Q column equilibrated in buffer A. RcdA was eluted using a 40 ml gradient of 0-100% buffer B (buffer A + 1M KCl).

Crystallization and structure determination. Native or selenomethionine-containing crystals were grown by hanging drop vapor diffusion at room temperature in 1:1 mixtures of protein and 1.5-1.75 M sodium formate, 0.1M HEPES pH 7.5 or 0.1M Tris pH 8.5. Tetragonal crystals grew overnight in clusters from a single nucleation point and were separated prior to screening and data collection. Native data were collected at the Advanced Light Source Beamline 5.0.2, (Lawrence Berkeley National Lab) and a 3-wavelength dataset from a single selenomethionine derivative crystal was collected at the Advanced Light Source Beamline 8.3.1. Data was processed with HKL2000 (32), and selenium sites were located using multiple anomalous dispersion using SHELX (33). Eight selenium sites were found, and interpretable maps were generated upon density modification in SHELX. Poly-alanine helices were placed into the maps for all six helices in the asymmetric unit. These were not refined but instead the used as a molecular replacement model for the higher-quality native dataset. Molecular replacement solutions were easily found using PHASER (34), and the coordinates were built using COOT (35) and refined in REFMAC (36). Initial rounds of refinement were constrained by the selenomethionine experimental phases. Coordinates and structure factors have been submitted to the pdb (PDB CODE). Images of RcdA coordinates were generated using PYMOL (Delano Scientific LLC).

Size exclusion chromatography. Gel filtration was performed on a Superdex-200 column (GE Healthcare) that had been calibrated using albumin, ovalbumin, chymotrypsin and blue dextran (GE Healthcare) as standards. The column was run in buffer A + 18% buffer B, and elution peaks measured by absorbance at 280 nm. Molecular masses were calculated according to K_{av} vs. $\ln(M_w)$.

CtrA half-life measurement. Pulse-chase and immunoprecipitation to determine CtrA half-life was performed as described (37), except that cultures were grown in M2G supplemented with 0.3% xylose, and cell lysates were precleared using 30µl Pansorbin (Calbiochem).

Microscopy. Fluorescence and DIC microscopy were performed as described (37).

Synchronization and immunoblots. Isolation of G1-phase swarmer cells and synchronous growth were performed as described (2), except that cells were grown in PYE/0.03% xylose. Immunoblots were probed with CtrA (2) and McpA (38) antisera at 1:10,000 and 1:30,000 dilutions, respectively.

Acknowledgements. We thank Pascal Egea for technical assistance with crystallography. We thank all members of the Ryan lab for discussion. This work was supported in part by NSF grant 0543801 to K.R.

References

1. Collier J, Shapiro L (2007) Spatial complexity and control of a bacterial cell cycle. *Curr Opin Biotechnol* **18**: 333-340.

2. Quon KC, Marczynski GT, Shapiro L (1996) Cell cycle control by an essential bacterial two-component signal transduction protein. *Cell* **84**: 83-93.

3. Domian IJ, Quon KC, Shapiro L (1997) Cell type-specific phosphorylation and proteolysis of a transcriptional regulator controls the G1-to-S transition in a bacterial cell cycle. *Cell* **90**: 415-424.

4. Laub MT, Chen SL, Shapiro L, McAdams HH (2002) Genes direcly controlled by CtrA, a master regulator of the *Caulobacter* cell cycle. *Proc Natl Acad Sci USA* **99**: 4632-4637.

5. Quon KC, Yang B, Domian IJ, Shapiro L, Marczynski GT (1998) Negative control of bacterial DNA replication by a cell cycle regulatory protein that binds at the chromosome origin. *Proc Natl Acad Sci USA* **95**: 120-125.

6. Jenal U, Fuchs T (1998) An essential protease involved in bacterial cell-cycle control. *EMBO J* **17**: 5658-5669.

7. Licht S, Lee I (2008) Resolving individual steps in the operation of ATP-dependent proteolytic molecular machines: From conformational changes to substrate translocation and processivity. *Biochemistry* **47**: 3595-3605.

8. Biondi EG, et al. (2006) Regulation of the bacterial cell cycle by an integrated genetic circuit. *Nature* **444**: 899-904.

9. Iniesta AA, McGrath PT, Reisenauer A, McAdams HH, Shapiro L (2006) A phospho-signaling pathway controls the localization and activity of a protease complex critical for bacterial cell cycle progression. *Proc Natl Acad Sci USA* **103**: 10935-10940.

10. McGrath PT, Iniesta AA, Ryan KR, Shapiro L, McAdams HH (2006) A dynamically localized protease complex and a polar specificity factor control a cell cycle master regulator. *Cell* **124**: 535-47.

11. Kainth P, Gupta RS (2005) Signature proteins that are distinctive of alpha proteobacteria. *BMC Genomics* **6**: 94.

12. Ryan KR, Judd EM, Shapiro L (2002) The CtrA response regulator essential for *Caulobacter crescentus* cell-cycle progression requires a bipartite degradation signal for temporally controlled proteolysis. *J Mol Biol* **324**: 443-455.

13. Ryan KR, Huntwork S, Shapiro L (2004) Recruitment of a cytoplasmic response regulator to the cell pole is linked to its cell cycle-regulated proteolysis. *Proc Natl Acad Sci USA* **101**: 7415-7420.

14. Ades SE (2004) Proteolysis: Adaptor, adaptor, catch me a catch. *Curr Biol* **12**: R924-926.

15. Keiler KC, Waller PR, Sauer RT (1996) Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* **271**: 990-993.

16. Gottesman S, Roche E, Zhou Y, Sauer RT (1998) The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes Dev* **12**: 1338-1347.

17. Levchenko I, Seidel M, Sauer RT, Baker TA (2000) A specificity-enhancing factor for the ClpXP degradation machine. *Science* **289**: 2354-2356.

18. Wah DA, Levchenko I, Baker TA, Sauer RT (2002) Characterization of a specificity factor for an AAA+ ATPase: assembly of SspB dimers with ssrA-tagged proteins and the ClpX hexamer. *Chem Biol* **9**: 1237-1245.

19. Wah DA, et al. (2003) Flexible linkers leash the substrate binding domain of SspB to a peptide module that stabilizes delivery complexes with the AAA+ ClpXP protease. *Mol Cell* **12**: 355-363.

20. Song HK, Eck MJ (2003) Structural basis of degradation signal recognition by SspB, a specificity-enhancing factor for the ClpXP proteolytic machine. *Mol Cell* **12**: 75-86.

21. Chien P, Perchuk BS, Laub MT, Sauer RT, Baker TA (2007) Direct and adaptor-mediated substrate recognition by an essential AAA+ protease. *Proc Natl Acad Sci USA* **104**: 6590-6595.

22. Peter BJ, et al. (2004) BAR domains as sensors of membrane curvature: the amphiphysin BAR structure. *Science* **303**: 495-499.

23. Krissinel E, Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr* **D60**: 2256-2268.

24. Meisenzahl AC, Shapiro L, Jenal U (1997) Isolation and characterization of a xylose-dependent promoter from *Caulobacter crescentus*. *J Bacteriol* **179**: 592-600.

25. Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN (2005) Flexible nets. The roles of intrinsic disorder in protein interaction networks. *FEBS J* **272**: 5129-5148.

26. Dunker AK, Obradovic Z, Romero P, Garner EC, Brown CJ (2000) Intrinsic protein disorder in complete genomes. *Genome Inform Ser Workshop Genome Inform* **11**: 161-171.

27. Dougan DA, Weber-Ban E, Bukau B (2003) Targeted delivery of an ssrA-tagged substrate by the adaptor protein SspB to its cognate AAA+ protein ClpX. *Mol Cell* **12**: 373-380.

28. Levchenko I, Grant RA, Wah DA, Sauer RT, Baker TA (2003) Structure of a delivery protein for an AAA+ protease in complex with a peptide degradation tag. *Mol Cell* **12**: 365-372.

29. Osteras M, Stotz A, Schmid Nuoffer S, Jenal U (1999) Identification and transcriptional control of the genes encoding the *Caulobacter crescentus* ClpXP protease. *J Bacteriol* **181**: 3039-3050.

30. Ely B (1991) Genetics of *Caulobacter crescentus*. *Methods Enzymol* **204**: 372-384.

31. Van Duyne GD, Standaert RF, Karplus PA, Schreiber SL, Clardy J (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J Mol Biol* **229**: 105-124.

32. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**: 307-326.

33. Sheldrick GM (2008) A short history of SHELX. *Acta Crystallogr* **A64**: 112-122.

34. McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr* **D63**: 32-41.

35. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr* **D60**: 2126-2132.

36. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr* **D53**: 240-255.

37. Reisinger SJ, Huntwork S, Viollier PH, Ryan KR (2007) DivL performs critical cell cycle functions in *Caulobacter crescentus* independent of kinase activity. *J Bacteriol* **189**: 8308-8320.

38. Alley MR, Maddock JR, Shapiro L (1993) Requirement of the carboxyl terminus of a bacterial chemoreceptor for its targeted proteolysis. *Science* **259**: 1754-1757.

Figure Legends

Figure 1. The crystal structure of RcdA shows novel dimerization and multiple disordered regions. (*A*) Stereo view of the 2fo-fc electron density (1 sigma) of helix 3 near the C-terminal disordered region. (*B*) Overview of the RcdA dimer. One monomer is blue and the other is gold. Approximate locations of the N-terminal disordered peptide (gray), the internal disordered loop (red) and the C-terminal disordered peptide (green) are shown. (*C*) Topology of the RcdA three-helix bundle. Helices 2 and 3 are parallel to each other, and two identical bundles make up the dimer. (*D*) Conservation of surface residues including the dimer interface and paddle region (circled) of RcdA. Highly conserved (magenta) and moderately conserved (pink) residues are shown on the gold monomer. Surface-exposed hydrophobic residues in helix 2 are indicated by blue arrows. Additional conserved residues occur in the C-terminal disordered peptide (SI Fig. 5). (*E*) Residues altered by site-directed mutagenesis are shown in cyan. Hydrophobic residues that were replaced with serines (black arrows) and charged residues that were replaced with alanines (red arrows) are indicated on the gold monomer.

Figure 2. The C-terminus of RcdA is necessary for cell cycle-regulated CtrA degradation. (*A*) Schematic of *Caulobacter* cell cycle progression. Wavy and straight lines indicate polar flagella and pili, respectively. Circles indicate nonreplicating chromosomes, while theta structures indicate replicating chromosomes. (*B-D*) Western blots of lysates from equal numbers of cells withdrawn from synchronized populations at the indicated times (minutes). Blots were probed with antiserum against CtrA or McpA. (*B*) KR2209, expressing RcdA-GFP. (*C*) KR2269, expressing RcdApaddle. (*D*) KR2210,

expressing RcdAΔC. (*E*) Anti-RcdA Western blot of lysates from equal cell numbers of CB15N (lanes 1 and 2), KR2209 (lanes 3 and 4), and KR2210 (lanes 5 and 6). Lanes 1, 3, 5: samples from cells grown in PYE. Lanes 2, 4, 6: samples from cells grown in PYE/0.03% xylose. An unidentified cross-reacting band is indicated (*). (*F*) Size-exclusion chromatography of RcdA (thin trace) and RcdAΔC (thick trace) on a Superdex-200 column. The elution peaks of protein standards used to calibrate the column are indicated above.

Figure 3. The RcdA C-terminus is necessary for RcdA and CtrA localization, but is not sufficient to localize a passenger protein. Left panels are DIC images matching the fluorescence images on the right. All strains were grown in PYE/0.03% xylose. (*A-B*) KR2209, RcdA-GFP. Filled arrows highlight polar localization, while the hollow arrow indicates midcell localization. (*C-D*) KR2211, RcdAΔC-GFP. (*E-F*) KR2256, YFP-CtrARD+15 in *ΔrcdA*/*xylX*::*rcdA*. (*G-H*) KR2247, YFP-CtrARD+15 in *ΔrcdA*/*xylX*::*rcdAΔC*. (*I-J*) KR2254,YFP-C.

Figure S1. Qualitative surface electrostatic representation of the RcdA dimer, where red indicates negative charge and blue indicates positive charge. Convex (left), side (middle) and concave (right) faces are shown.

Figure S2. Sequence alignment of RcdA homologs from the indicated species, which represent the genera *Caulobacter, Rhodopseudomonas, Bradyrhizobium, Nitrobacter, Methylobacterium, Stappia, Maricaulis, Oceanicaulis, Parvibaculum,* and *Hyphomonas*

of the α-proteobacteria. Black bars above the sequence indicate helices 1-3. The gray, red, and green bars indicate the N-terminal, internal, and C-terminal disordered peptides, respectively. Asterisks below the sequence indicate completely conserved residues, and two dots indicate highly conserved residues.

Figure S3. *ΔrcdA* cells expressing RcdAΔC display CtrAD51E-induced G1 arrest. The left and middle columns are DIC images of cells grown for 8 hrs in PYE supplemented with 0.2% glucose or 0.3% xylose, respectively. The right column shows chromosome content of imaged cultures as measured by flow cytometry of Sytox green-labeled cells. In each graph, the data for cells grown in PYE/glucose (red trace) and PYE/xylose (blue trace) are overlaid. (A-C) KR1759, CB15N + pIDC42. (D-F) KR1757, *ΔrcdA* + pIDC42. (G-I) KR2272, *ΔrcdA/xylX::rcdA* + pIDC42. (J-L) KR2273, *ΔrcdA/xylX::rcdAΔC* + pIDC42.

Figure S4. CpdR and RcdA do not alter the rate of CtrA degradation by ClpXP *in vitro*. Fraction of CtrA remaining as determined by SDS-PAGE and Coomassie blue staining of samples incubated for the indicated times at 30ºC. All reactions contained an ATP regeneration system, 1.16 μM CtrA (initially), 0.3 μM $ClpX_6$, and 0.8 μM $ClpP_7$ (red). Some reactions also contained 3 μM RcdA and 3 μM CpdR (blue). Curves are fits of the data to an exponential decay. Results are the average of three experiments; error bars show one standard deviation.

SI Table 1. Crystallographic data collection and refinement for RcdA

	Native data	MAD data		
Space group	$P4_3$	$P4_3$		
Cell dimensions (Å)	a=b=75.5 c=81.6	a=b=75.7 c=81.4		
		Peak	Inflection	Remote
Wavelength (Å)	1.1	0.97957	0.97972	1.01986
Resolution (Å)	44-2.9		44.6-3.2	
R_{merge} (%)	8.5 (29.9)	16.4 (55.0)	15.4 (60.5)	12.5 (48.7)
I/σI	18.1	9.6	8.9	10.8
Completeness (%)	99.1 (89.6)	99.9	99.9	99.9
Redundancy	4.7 (3.6)	4.3	4.3	4.2
Total # Obs. Reflections	44,107	32,822	32,851	32,190
Refinement				
Resolution (Å)	44-2.9 (2.975-2.90)			
# reflections used during refinement	9,630			
R_{work}	27.5 (30.9)			
R_{free}	31.2 (33.0)			
R.m.s deviations				
Bond lengths (Å)	0.011			
Bond angles(°)	1.64			

SI Table 2. Plasmid and strain descriptions.

Plasmid	Description	Reference
pMR10	Broad host-range, low copy-number vector	(1)
pMR20	Broad host-range, low copy-number vector	(1)
pPM34	P_{xyl}-*rcdA* for integration at *xylX* locus	P. McGrath, personal communication
pPM53	P_{xyl}-*rcdA::GFP* for integration at *xylX* locus	P. McGrath, personal communication
pJT41	P_{xyl}-*rcdApatch::GFP* for integration at *xylX* locus, derived from pPM53	This work
pJT46	P_{xyl}-*rcdAΔC* for integration at *xylX* locus, derived from pPM53	This work
pJT49	P_{xyl}-*rcdAΔC::GFP* for integration at *xylX* locus, derived from pJT46	This work
pJT50	P_{xyl}-*rcdApaddle::GFP* for integration at *xylX* locus, derived from pJT41	This work
pES53	pET28d-*rcdA*	This work
pJT51	pET28d-*rcdAΔC*, derived from pES53	This work
pEJ146	pMR10-P_{xyl}-*YFP::ctrARD+15*	(2)
pJT52	pMR20-P_{xyl}-*YFP::ctrARD+15*	This work
pJT56	pMR10-P_{xyl}-*YFP-C(rcdA143-169)*	This work
pCJ6	pET21a-*ctrA*	C. Jacobs-Wagner, personal communication
pKR14	pET28a-*clpX*	This work
pJT20	pET29-*clpP*	This work
pET-RR7	pET28a-*cpdR*	P. McGrath, personal communication
Caulobacter strains		
KR684	Wild type strain CB15N	(3)
KR899	*ΔrcdA*	(4)
KR2207	*ΔrcdA/xylX::rcdA*	This work
KR2209	*ΔrcdA/xylX::rcdA::GFP*	This work
KR2210	*ΔrcdA/xylX::rcdAΔC*	This work
KR2211	*ΔrcdA/xylX::rcdAΔC::GFP*	This work
KR2268	*ΔrcdA/xylX::rcdApatch::GFP*	This work
KR2269	*ΔrcdA/xylX::rcdApaddle::GFP*	This work
KR2254	CB15N/*xylX::YFP::C(rcdA143-169)*	This work
KR2247	*ΔrcdA/xylX::rcdAΔC* + pJT52	This work
KR2248	*ΔrcdA* + pJT52	This work
KR2256	*ΔrcdA/xylX*::*rcdA* + pJT52	This work

E. coli strains		
BL21	Protein overexpression strain	Novagen
BL21 Tuner	Protein overexpression strain	Novagen
BL21 Rosetta	Protein overexpression strain	Novagen
KR2278	BL21 Tuner/pES53	This work
KR2279	BL21 Tuner/pJT51	This work
KR207	BL21/pCJ6	This work
KR31	BL21/pKR14	This work
KR912	BL21 Rosetta/pET-RR7	P. McGrath, personal communication
KR1838	BL21 Tuner/pJT20	This work

References

1. Roberts RC, et al. (1996) Identification of a *Caulobacter crescentus* operon encoding hrcA, involved in negatively regulating heat-inducible transcription, and the chaperone grpE. *J Bacteriol* **178**: 1829-1841.

2. Ryan KR, Judd EM, Shapiro L (2002) The CtrA response regulator essential for *Caulobacter crescentus* cell-cycle progression requires a bipartite degradation signal for temporally controlled proteolysis. *J Mol Biol* **324**: 443-455.

3. Evinger M, Agabian N (1977) Envelope-associated nucleoid from *Caulobacter crescentus* stalked and swarmer cells. *J Bacteriol* **132**: 294-301.

4. McGrath PT, Iniesta AA, Ryan KR, Shapiro L, McAdams HH (2006) A dynamically localized protease complex and a polar specificity factor control a cell cycle master regulator. *Cell* **124**: 535-47.

Figure 1.

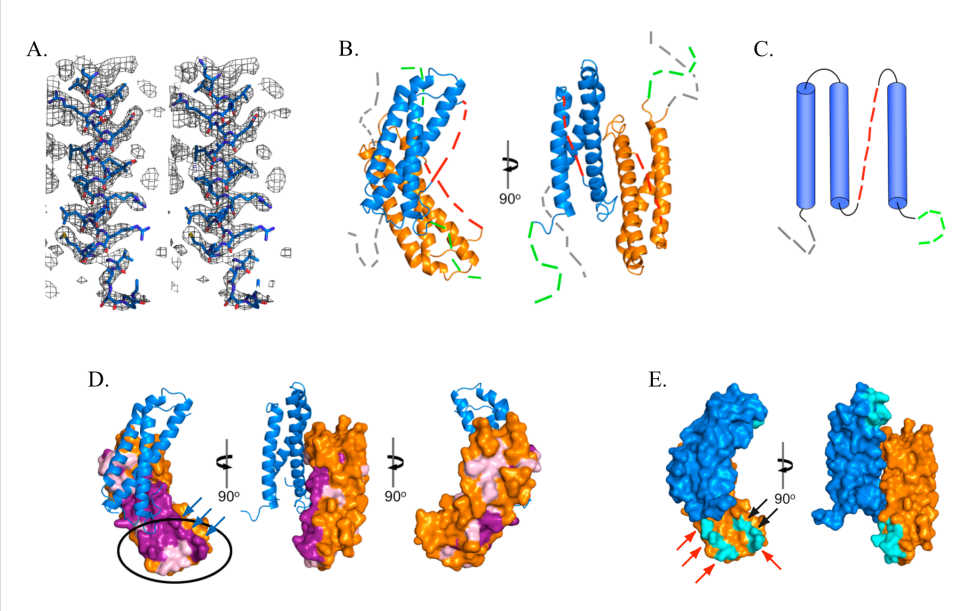


Figure 2

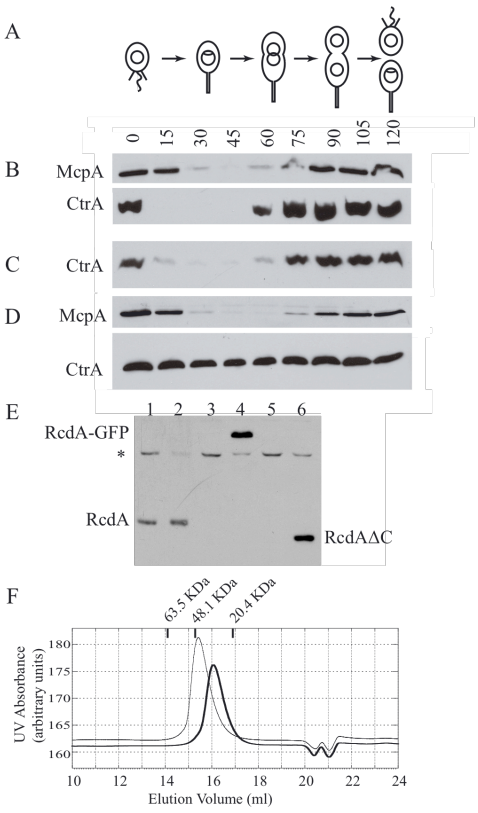

A
B
McpA
CtrA
C
CtrA
D
McpA
CtrA
0
15
30
45
60
75
90
105
120
E
1 2 3 4 5 6
RcdA-GFP
*
RcdA
RcdAΔC
F
63.5 KDa
48.1 KDa
20.4 KDa
UV Absorbance (arbitrary units)
180
175
170
165
160
10 12 14 16 18 20 22 24
Elution Volume (ml)

Figure 3.

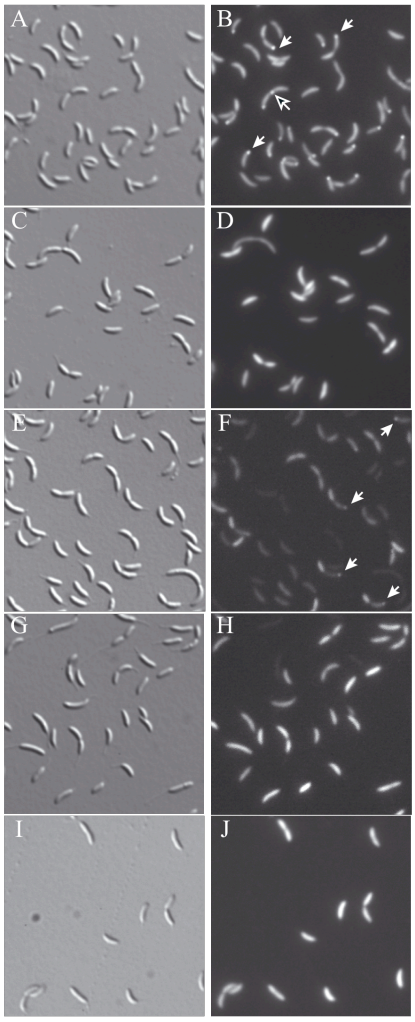


Figure S1.

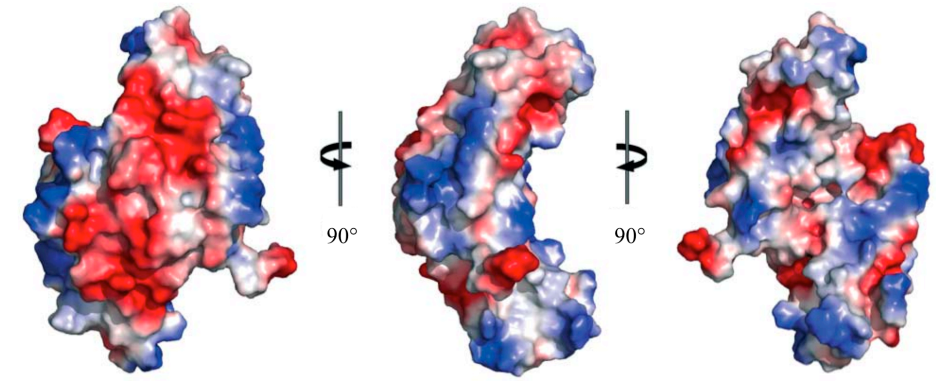


Figure S2.

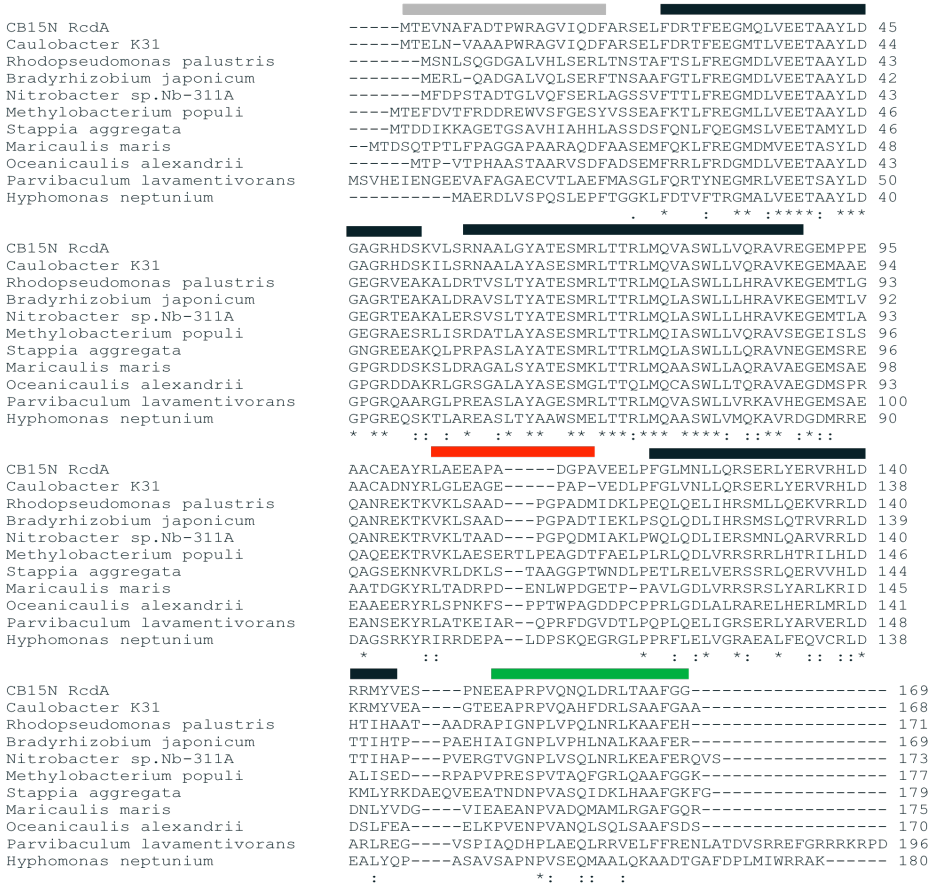

CB15N RcdA
Caulobacter K31
Rhodopseudomonas palustris
Bradyrhizobium japonicum
Nitrobacter sp.Nb-311A
Methylobacterium populi
Stappia aggregata
Maricaulis maris
Oceanicaulis alexandrii
Parvibaculum lavamentivorans
Hyphomonas neptunium

Figure S3.

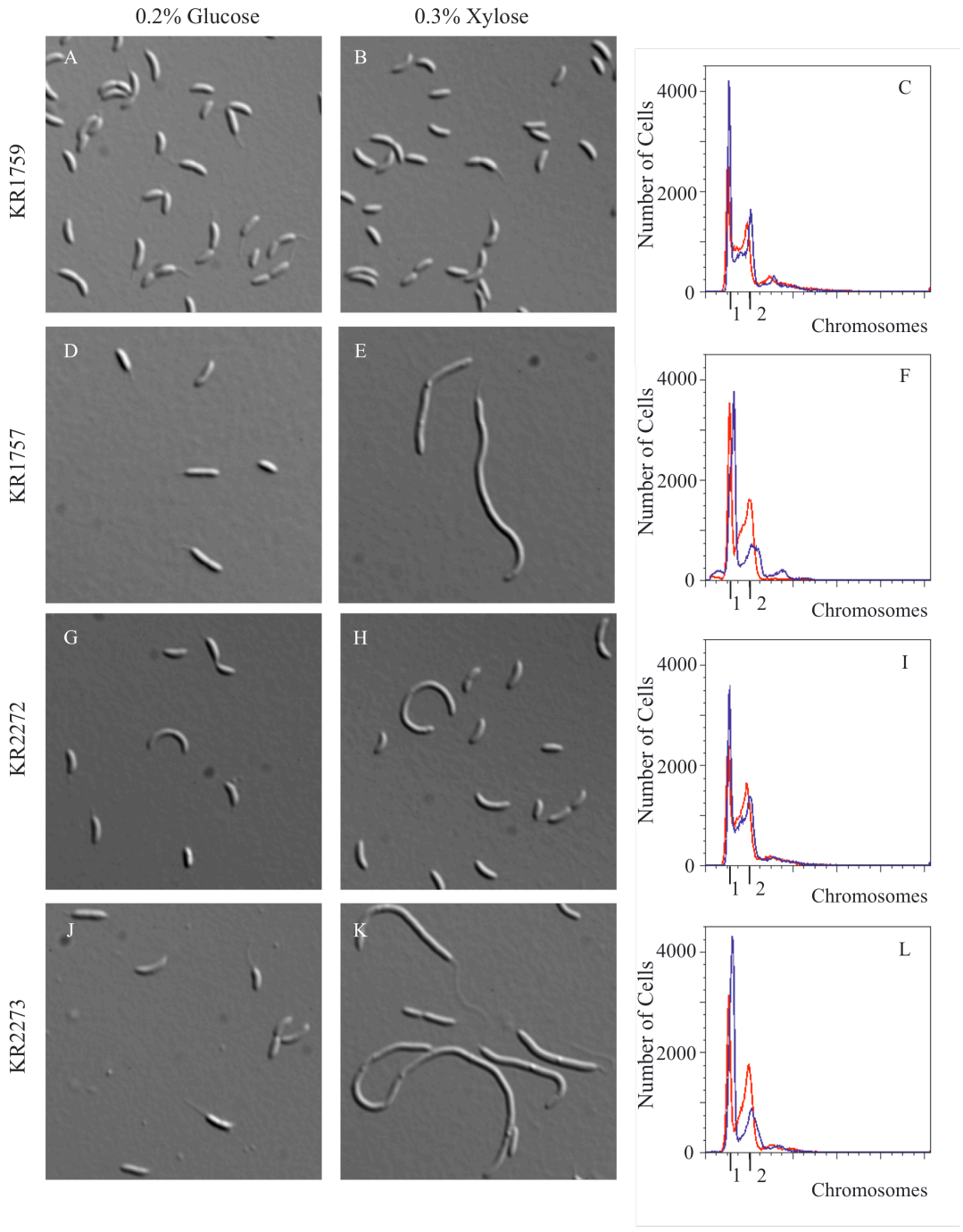
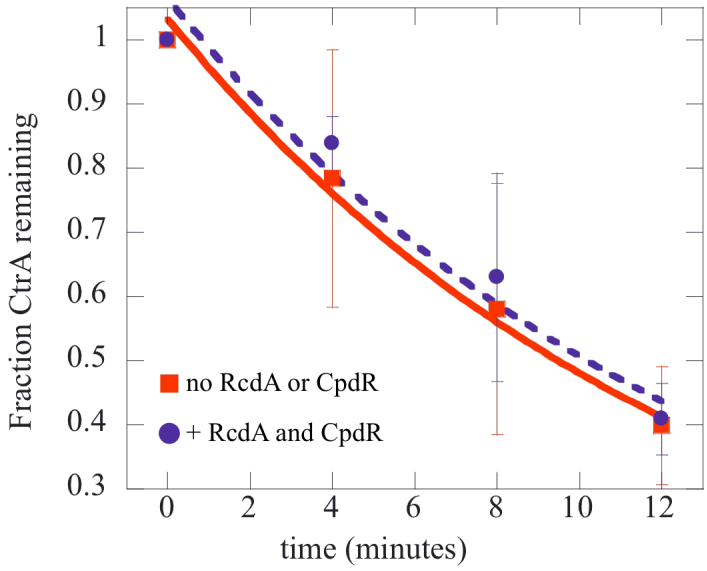

0.2% Glucose
0.3% Xylose
KR1759
KR1757
KR2272
KR2273
A
B
C
D
E
F
G
H
I
J
K
L
Number of Cells
4000
2000
0
1
2
Chromosomes

Figure S4.


Fraction CtrA remaining
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0
2
4
6
8
10
12
time (minutes)
no RcdA or CpdR
+ RcdA and CpdR

Appendix 2. Experimental methods and future directions

1) TDD crystallization

Goal: Determine the crystal structure of distal leg of clathrin and complexes of distal leg with protein interactors.

The distal leg of clathrin remains one of the few fragments of clathrin heavy chain to have no high resolution structure determined. Trimerization and very C terminus are the others. The distal leg is expected to adopt a structure similar to the proximal leg and does so in the clathrin basket (based on cryoEM reconstructions). Purification of TDD is straight forward and crystallization was attempted though not successful.

Expression and purification is identical to clathrin hub region, including the same buffers and similar yields. Briefly,

Bacteria harboring pET23d-TDD were grown at 37°C until OD600 = ~0.6.

Cultures were cooled to 30°C and induced for ~4 hours with 1mM final concentration of IPTG

Bacteria were pelleted and frozen at -80°C

Purification was with NiNTA (or TALON) resin via standard methods followed by concentration and further separation on superdex 75 or superdex 200.

Purified TDD was concentrated to ~4mg/ml and set into crystal screens.

Best crystallization conditions were:
0.2M Tartrate +1.0% TFE , grown under oil to slow equilibration.

Continuation of the these experiments would be best suited by starting with freshly puried TDD concentrated to ~6mg/ml and screening the most recent crystallization condition libraries using robotic drop setting. This would allow a broad range of crystallization space to sampled and increase the likelihood of generating high quality crystals.

2) Chemical cross linking as a probe for conformational changes in clathrin light chain.

Goal: To determine the whether clathrin light chain undergoes conformational changes correlating with clathrin assembly.

Conformational changes in clathrin light chain appear to have importance consequences on the regulation of clathrin lattice assembly. Conformational changes especially in large self-assembling molecule can prove challenging to elucidate. The self-assembling nature

of clathrin means methods often used to measure protein dynamics likely will not work. NMR for example would be highly sensitive to changes in the assembly state of the clathrin lattice and even the smallest fragments of clathrin are too large for most NMR experiments. Crystallography is difficult with self-assembling fragments of clathrin though as can be seen in chapter 1 is possible. Improving crystals used in chapter 1 may be one way to pursue conformational change information. The FRET based methods described in chapter 1 also provide some insight into conformational changes. One method that is higher resolution than the FRET methodology is to combine chemical cross linking and mass spectrometry to identify interacting regions between clathrin light chain and clathrin heavy chain. Upon conformational changes sights of cross linking should be different. Identifying the regions of cross linking should allow relatively positioning of the light chain on the heavy chain.

The selection of chemical cross linkers will define the complexity of the experiment and data interpretation. A single chemical cross link between a mobile region of clathrin light chain and clathrin heavy chain would lead to a single change. No specific cross linking could achieve similar or higher resolution but rapidly becomes complex with multiple regions of clathrin light chain cross linking to clathrin heavy change and therefore many changes to account for. Caution is advised to use photo activated cross linkers or multi set cross linkers to minimize unintentional cross linking between heavy chain residues.

One photoactivatable cross linker was attempted. This was maleimide linked to a photoactivatable aryl azide. This was first reacted with cysteine residues on clathrin light

chain (nLCb) via the manufacturers protocol (Pierce). The buffer was extensively exchanged and cross linker removed. The clathrin light chains were then added to clathrin heavy chains and incubated from 1 hour to overnight. Complexes of clathrin light chain and heavy chain were either assembled or not by addition of pH 6.2 MES buffer. The photoactivation was initiated by exposure to UV light. Crosslinking under either condition was inefficient as measured by a shift in SDS-PAGE gel bands. Attempts to identify cross linked positions by subsequent digestion with trypsin and collecting peptide mass data by mass spectrometry lead only to ambiguous results.

Expression of clathrin light chain was weak for the duration of these experiments and increased expression levels may be one simple way to improve upon these experiments. This would lead to greater amounts of starting materials and better initial cross linking to cysteine residues. Additionally, this would allow better means of purifying the modified clathrin light chains and improve over all yields for all steps of the reactions.

Another step that may increase data quality is to select a chemical cross linker that can first be used to modify cysteine residues on clathrin light chain and then only cross link a specific residue in a second step. This would allow labeling of clathrin light chain only, and then minimize the possible reactions that could occur with heavy chain. Minimization of reactions are the easiest way to improve data interpretation, which is the most problematic part of these experiments.

3) Blocking assembly of clathrin lattices by clathrin heavy chain point mutants

Goal: To understand molecular interactions occurring during clathrin lattice formation and generate non-assembling clathrin heavy chains.

Though previously Diane Wakeham had made many mutants in charged residues in clathrin heavy chain that were intended to prevent assembly, none seemed to function this way. Often the switching of charged residues had the affect of increasing assembly of clathrin hubs. This is likely because clathrin has evolved to be weakly interacting along a large surface area. Based on my own analysis, I sought to find, instead of charge interactions, small hydrophobic patches that would drive the assembly. I identified one such patch around tryptophan 1340 in bovine clathrin heavy chain. This patch consisted of leucine, phenylalanine, and tryptophan.

This patch is an attractive target for contributing to clathrin assembly due to its position along the leg. If one imagines two interacting clathrin legs as rods attached at a central fulcrum point, these two rods could rotate on that fulcrum point to generate numerous angles of interaction. Similarly the legs of clathrin need to interact at different angles depending on whether the clathrin molecule is in a flat lattice or a curved lattice. The hydrophobic patch around residue 1340 could act as a fulcrum point of interaction, and would be invariable with changing lattice shapes.

To test whether leucine or phenylalanine in this patch contributed significantly to assembly, these residues were mutated to alanine either separately or together. Clathrin hubs containing these mutations were referred to LAFA (double mutant). Unfortunately assembly for these particular mutants was not significantly affected.

To continue this work, the tryptophan should be mutated as well to create the triple mutant and triple mutant hubs tested for assembly. Additionally, mutations to residues other than alanine may be worth investigating.

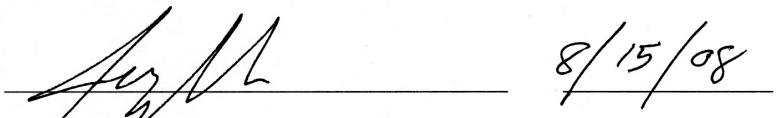
Finally, its seems very unlikely that mutating a couple of hydrophobic interactions will completely block clathrin assembly due to the very large buried surface area in the assembled state. A possible means of creating non-assembling clathrin is implied in chapter 1. If one could mutate clathrin to create a molecule that can only adopt the straight leg form this would likely generate a non-assembling molecule. Mutation that would achieve this may be proline residues in the loop regions of clathrin near the knee or mutating certain small residues to larger bulkier residues.

Publishing Agreement

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

Please sign the following statement:

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



Author Signature Date