MOLECULAR BASIS OF BAR DOMAIN SUPER-FAMILY PROTEINS AND GENETICALLY ENCODED CALCIUM INDICATORS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Qi Wang

January 2011

MOLECULAR BASIS OF BAR DOMAIN SUPER-FAMILY PROTEINS AND GENETICALLY ENCODED CALCIUM INDICATORS

Qi Wang, Ph.D.

Cornell University 2011

Protein domains are the basic functional modules that maintain cell functions at a molecule level. Previous studies have mainly focused on the functions of isolated protein domains. The general objective of this thesis is to understand functions and regulations of multi-domain containing proteins. The study is based on two protein classes: naturally occurred BAR domain-containing proteins and artificially engineered calcium indicators.

BAR domain super-family proteins

BAR (Bin/Amphiphysin/Rvs) domain super-family proteins are peripheral membrane proteins that regulate membrane curvatures during the membrane remodeling events such as endocytosis, vesicular trafficking and cell growth. Via multiple biophysical approaches, I systematically studied BAR domain functions in Sorting Nexin 9, Endophilin and Pacsins at the presence of other protein domains. Two major findings are presented in this thesis. First I show that the diverse membrane sculpture activity of BAR domains is encoded in their unique molecular structures, and is influenced by membrane properties. Second, I show that this function diversity is highly regulated by other protein domains. Some protein domains have synergetic effects and play important roles in regulating cellular membrane remodeling. This work is significant in that it provides the molecular basis for the functional diversity of BAR domains and established the regulatory mechanism of BAR domain mediated-membrane deformation process.

Genetically Encoded Calcium Indicators

Genetically encoded calcium indicator GCAMP is an artificially designed multi-domain containing protein that can be endogenously expressed in cells to monitor calcium signals. The molecular mechanisms of its signal response and fast kinetics are poorly understood. Using fluorescent spectrometry and site-directed mutagenesis, I show that the calcium-dependent brightness of GCAMP is due to the different protonation states of the chromophore. Structural characterization of GCAMP reveals that the calmodulin domain regulates chromophore protonation states via a sophisticated water-mediated hydrogen bond network. This finding provided a general scheme for designing GCAMP-like sensors. Furthermore, I show that distinct electron properties of the protonated and deprotonated chromophore can be applied to design color switchable fluorescent proteins. This finding provides a novel approach to design the ratio metric pH sensors with an improved sensitivity.

BIOGRAPHICAL SKETCH

The author was born in Lanzhou, Gansu Province, China. In 2001, he graduated from the Lanzhou No.1 High School and matriculated at College of Chemistry and Molecular Engineering, Peking University, Beijing. From 2002-2005, he was as a Presidential Undergraduate Research fellow with Dr. Luhua Lai at Center of Theoretical Biology. In 2005 he graduated with honors from Peking University with a B.S. in Chemistry (Physical Chemistry and Applied Chemistry). Subsequently, he was admitted into Cornell University as a graduate student in the field of Biophysics. After one year course preparation and research rotations, the author joined Dr. Holger Sondermann lab to continue his Doctorate training. The author has accepted a position as a Postdoctoral Fellow at the University of California, Berkeley.

这篇论文献给我的奶奶，曹淑珍女士

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my mentor Dr. Holger Sondermann. It is his exceptional mentorship that makes me become an independent scientist. Dr. Sondermann has created an extremely wonderful research environment, in which every student succeeds their training. He is an exemplifier of great educators, as reflected from an old Chinese proverb: Give man a fish and you feed him in a day, teach man to fish and you feed him for a lifetime. After five years training in his lab, I got the most important skills for my future career. Dr. Sondermann is the most influential person in my graduate career. His passion and enthusiasm have overshadowed my failures and frustrations, his long-lasting support and confidence have made me succeed the graduation school.

I am very grateful to have had a unique special committee that constantly oversees my project progress. Dr. Gerald Feigenson is my field Director of Graduate Study and an advisor in membrane biophysics. His expertise in lipid biochemistry plays an important role in my thesis research. Dr. James Sethna is my minor chair and an advisor in computational biology. The dicussions with Dr. Sethna have further developed me analytical thinking ability and quantitative research skills. Dr. Ruth Collins is my cell biology advisor and has provided me great advices to pursue a high-standard, high-quality doctoral training. I have great respect for all the members of my committee and enormous appreciations for their support on my career. Dr. Michael Kotlikoff, although not a member of my special committee, deserves thanks. He is an incredibly supportive collaborator in my fluorescent sensor project. I would also like to express my thanks to Dr. Gary Whittaker and Dr. Richard Cerione for sharing their equipment, reagents and providing me valuable suggestions on my research.

I would also like to thank Dr. Pietro De Camilli of Yale University, Dr Joshua Zimmerberg of NIH and Dr Frank Pollex of UNC School of Medicine for their

insightful suggestions. At Cornell, I would like to thank Dr. William Brown and Dr. Scott Emr for the valuable comments on the research manuscripts.

This thesis can not be completed without the help from many excellent scientists at Cornell University. Dr. Bret Judson provided me technical help on the negative-stained transmission electron microscopy, the major reporting assay used in my research. Dr. Bo Shui shared his screening results and reagents with me, and provided me great convenience to develop the collaborative project. Dr. Marc Antonyak helped me with the basics of mammalian cell lines. Dr. Marcos Navarro, Dr. Kanagalaghatta Rajashankar and Dr. Yang Zhang taught me important skills in X-ray crystallography. Their superior research skills and great patiences must be acknowledged. I would like to thank Dr. Min Wu for doing fluorescent microscopy study on my samples, Dr. Michelle Pirruccello for providing me reagents and Dr. Shawn Ferguson for his contributions in the cell biology related work.

I am in debt with my colleagues at Department of Molecular Medicine and field of Biophysics. Not only did these people provide me enormous help, they have become great friends. Shih-Lin Lynda Goh contributes to the lipid biochemistry assay development and the second part of the Pacsin study; David Cragun contributes to the yeast cell biology assays; Laura Byrnes contributes to the second part of the sensor development study. I must also thank Fred Heberle, Huolin Xin, Ben Machta, Changrui Lu and Nelson Morales-Pennington for their long-lasting interests in my thesis research. They have made my time much more enjoyable in Ithaca. I would also like to thank other graduate students and postdoctoral fellows Menqiao Wang, Sandra Dias, Yading Lin, Bo Li, and Jingwen Zhang for their supports. I wish them best luck in their pursuits.

I have had great help from several very talented undergraduate students and rotation students. Hung-Yi Kristal Kaan made significant headway on Sorting Nexin 9

study; Gary Peng contributes to the first part of the Pacsin study. Justin Torok, Reshma Hooda, Evan Molinelli and Michelle Tong have each made contributions to my projects, for which I am very grateful.

Many thanks go to the staff scientists at synchrotron stations for helping me during the data collection and interpretation. I would like to thank Dr. Richard Gillian of Cornell High Energy Synchrotron Source (CHESS) and Dr. Soenke Seifert of Advanced Photon Source (APS) for their help on Small Angle X-ray Scattering; Dr. Qun Liu of CHESS, Dr. Alexei Soares of Brookhaven National Laboratory for their assistance of X-ray crystal diffractions. The support staff of the Department of Molecular Medicine and the field of biophysics have ensured that things ran smoothly on a day to day basis. I would like to thank Debbie Crane, Cindy Westmiller, Valerie Moore, Blake Werner, Greg Mitchell, Vicki Shaff, Diane Colf and Kathy Dedrick.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

BAR DOMAIN SUPER-FAMILY PROTEINS

Membrane biophysics

One of the evolutionary advantages of eukaryotic cells is the presence of membrane-delimited, cellular organelles. For example, the nucleus sets eukaryotic cells apart from prokaryotic cells. The morphology of membranes is of great variety in nature (Zimmerberg and Kozlov, 2006). Eukaryotic plasma membranes can undergo radical shape transformations when they develop intercellular contacts, or spread and move. Cells like endothelial cells can develop internal tubules for taking up nutrition. One of the most architecturally striking organelle, Endoplasmic reticulum (ER), is composed of the nuclear envelope, sheet-like cisternae, and a polygonal array of tubules connected by three-way junctions (Shibata et al., 2006; Voeltz and Prinz, 2007). This section describes the composition of biomembranes, the phase behavior of model membranes and the cellular components that can influence membrane morphorlogy.

Biomembrane composition

The basic unit of a membrane bilayer is a lipid. The three major classes of membrane lipids are phospholipids, glycolipids and cholesterol. The most abundant biomembrane lipid is phospholipid, which predominately includes phopsphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylserine (PS) and sphingomyelin. PS is the only one among the four that carries one negative net charge. The natural abundance of phospholipids depends on their cellular locations. PC and PE are widely distributed in different cell organelles. PS is usually enriched in inner membrane leaflets and sphingomyelin is mostly found in the exoplasmic leaflet

of the cell membrane (van Meer et al., 2008). The molecular structure of these lipids is shown in Figure 1.1.A. Inositol phospholipid is another type of phospholipids. They exist in very small quantities, but serve as the major signaling lipids that participate in regulating receptor-mediated endocytosis, vesicular trafficking, cell proliferation and growth (De Camili et al., 1998; Fruman et al., 1998; Volpicelli-Daley and De Camili, 2007). Phosphoinositol lipids recruit signaling molecules such as receptor kinases, phosphatases, GTPases and other enzymes that are essential for cells (Di Paolo and De Camilli, 2006; Lemmen, 2007; Cremona and De Camilli, 2001). Cholesterol is highly hydrophobic, with a polar hydroxyl group at one side. In the membrane bilayer, cholesterol prefers to align vertically respect to the membrane bilayers, with the hydroxyl group close to the polar head group of phospholipids. Cholesterol has a major impact on the membrane fluidity and plays an important role in regulating the membrane remodeling event during endocytosis (Subtil et al., 1999; Farsad and De Camilli, 2003; Jia et al., 2006).

Membrane phase behavior

Even a three-component lipid bilayer system may exhibit complicated phase behavior (Zhao et al., 2007). Studies on model membrane systems have indicated that biological membranes may exist in several phases that are different from the degree of order and diffusion rate of the molecules within the bilayer, and the coexistence of two or more phases is possible under certain lipid compositions and temperature (Feigenson, 2007).

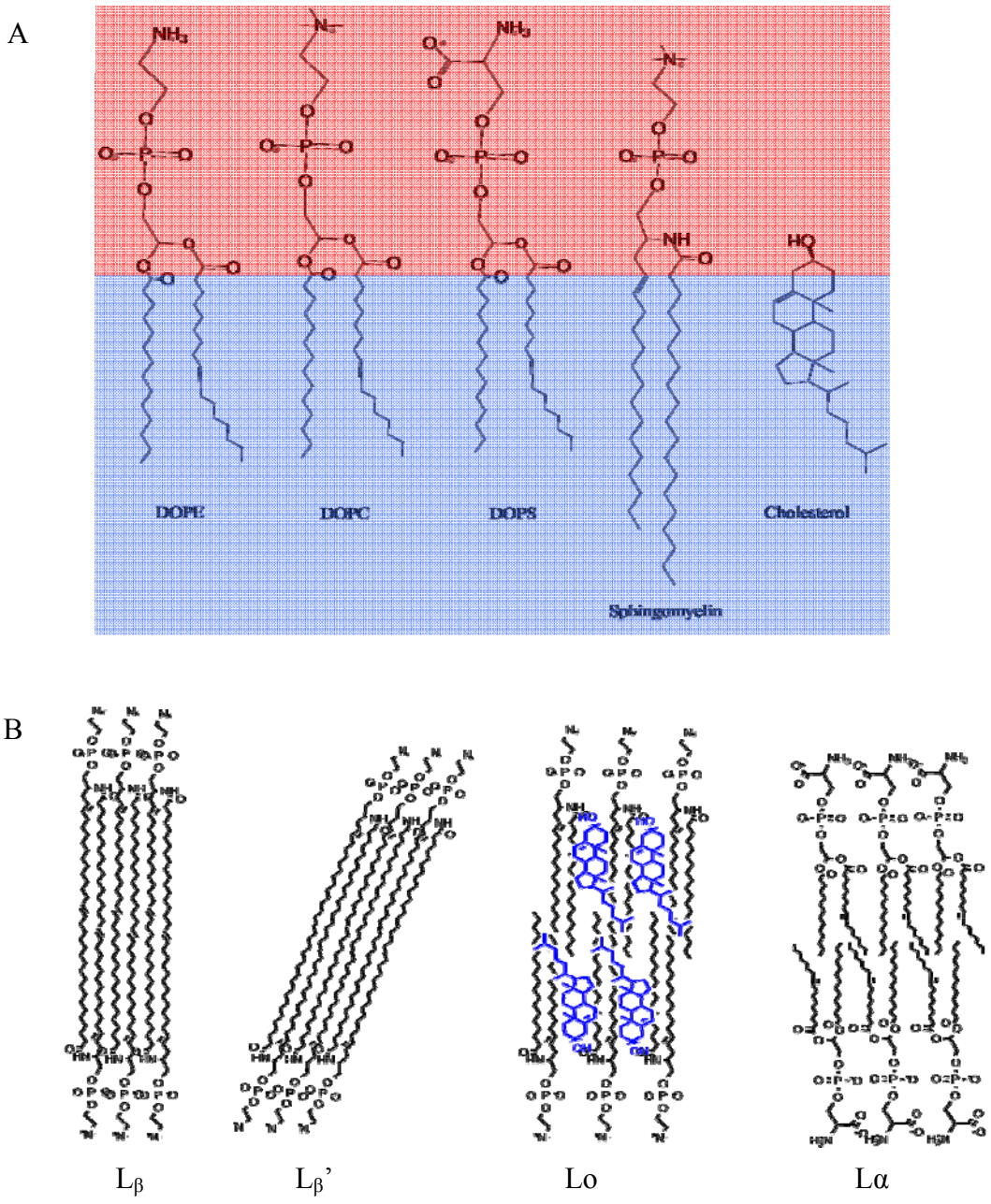


Figure 1.1. Lipid structures and membrane bilayer phases. (A) Chemical structure of DOPE, DOPC, DOPS, Sphingomyelin and Cholesterol. The hydrophobic part is colored in blue and the hydrophilic part is colored in red. (B) Four phases that membrane bilayer can adopt. The cholesterol is shown in blue color.

Liquid-solid phase ($L_{\beta,}$ also known as gel phase) has the high order of a solid and a low translational mobility. Phospholipids with saturated acyl chains (fully

hydrogenated, such as sphingomyelin) tend to pack tightly together and form L_β phase, as shown in Figure 1.1.B. In the L_β phase, the acyl chains can be either perpendicular to the membrane surface (Figure 1.1.B), or significantly tilted (L_β' phase, Figure 1.1.B). Cholesterol can increase the fluidity of gel phase. When increasing the cholesterol concentration in an L_β' phase, the tilted acyl chains may become straightened up. Further increase of cholesterol concentration results in a new phase with high order and the high translational mobility of a liquid. This phase is called the liquid-ordered phase (L_o). Phospholipids with unsaturated acyl chains are not favored in a highly compact conformation, thus they are often found liquid-disordered phase.

Using the phase diagrams determined for three-component model membrane systems, one is able to spectulate about the phase behavior of real biomembrane bilayers. Based on the concentration of the major lipid fractions in biomembranes, liquid phases and gel-like phases can exist, and the coexistences of $L_\alpha + L_o$ or $L_\alpha + L_\beta$ are also possible. Such phase separation will create heterogeneity in the cell membranes, and its function in the living cells has been demonstrated for specific cellcular processes. One of such phase-related phenomenon that is currently studied extensively is the biology of lipid rafts (Simons and van Meer, 1988; Lingwood and Simons, 2010). Lipid rafts are specialized membrane microdomains that compartmentalize cellular signal transduction pathways by locally recruiting and regulating signaling molecules, changing the membrane physical properties and thus regulate cellular events. The direct observation of lipid rafts in the living cells remains challenging, but it has been proposed that lipid rafts are ordered with the lipids being tightly packed but allowed to floate freely in the membrane bilayer, resembling L_o phase bahavior.

Membrane remodeling

Cellular membranes can change morphology in many ways during cell migration, division, vesicle budding, endocytosis and intra-cellular trafficking. In these events, the membrane can locally change its curvature and be remodeled into a new topology. For example, during the endocytosis, protein-coated membrane vesicles with a high-degree of curvature are formed from a relatively flat membrane sheet. Such drastic membrane shape change is usually mediated through sophisticated molecular machineries. This section introduces three different types of molecular machineries that can shape cellular membranes: cytoskeletal elements, lipids and membrane proteins.

Cytoskeletal elements, such as the F-actin filaments and microtubules have long been known to play some role in membrane trafficking, not only by forming the structural scaffold and network over which membrane traffic flows but also by directly deforming membrane bilayers (McMahon and Gallop, 2006). Those cytoskeletal elements can attach to the membrane surface and therefore apply the external pulling or pushing force that results in the bilayer forming tubules. For example, microtubules have been shown to play a role *in vitro* and *in vivo* in the tubular dynamics of ER, the Golgi and endosome tubulation events under certain conditions (Farsad and De Camilli, 2003).

Lipids can induce and stabilize local membrane curvature in multiple ways. Cholesterol, for example, is required for the generation of high-curvature clathrin-coated vesicles *in vivo*. One possible function of cholesterol is to be enriched at the budding site leaflet of the bi-layer and hence to increase the fluidity of this membrane surface (Subtil et al., 1999). Biochemical and enzymatic modification on lipids can produce particular asymmetries that favor certain curvatures. For example, a modified lipid with a relatively large head group and slim acyl chains could favors positive

curvature by adopting a wedge-like geometry in the membrane sheet. Some lipid-modifying enzymes such as PLA2 and PLC have been reported to play important roles in changing membrane structures *in vivo* (de Figueiredo et al., 1998; de Figueiredo et al., 2001).

Integral membrane proteins that entirely penetrate membrane bi-layer, such as the voltage-dependent K+ channel (Mackinnon, 2005; Long et al., 2007) and GPCR receptors (Rosenbaum et al., 2007) can disrupt local membrane curvature given their conical shapes. Another type of integral membrane protein can be partial embedded in the bi-layer structure and perturb the membrane in an asymmetric way, which results in a curvature change around the inserted region. A typical example for this type of proteins is Rtn4, a yeast transmembrane protein which possibly stabilizes the tubular structure of the ER by working as a wedge in into the outer membrane leaflet, increasing the surface area needed to generate the high curvature of a tubular structure (Shibata et al., 2009; Hu et al., 2008; Hu et al., 2009; Voeltz et al., 2006).

Some peripheral membrane proteins can work as scaffolds to change membrane curvature. The coating proteins such as Clathrin, COPI and COPII have been shown to form a rigid exoskeletons that can influence membrane bending by polymerizing into curved structures and therefore constrains the space and volume of the coated membrane vesicles. Another family of peripheral membrane proteins, Dynamins, can bind to inositol lipid and self-assemble into helical rings and spirals. The spiral structure is believed to be the scaffold for the formation of tubular membrane shape (Farsad and De Camilli, 2003).

Clathrin-mediated endocytosis

Clathrin-mediated endocytosis (CME) is the major form for mammalian cells to take up nutrients, and plays an essential role in down-regulating membrane

receptors (Schmid, 1997; Doherty and McMahon, 2009). During CME, membranes undergo drastic morphological changes, with the local curvature decreasing from hundreds of nanometers to tens of nanometers (Figure 1.2.A-E). Numerous factors, including phospholipids and protein machineries, are required to accurately regulate the membrane curvature changes during CME. This section discusses the major players that stabilize membrane bilayers in different curvatures.

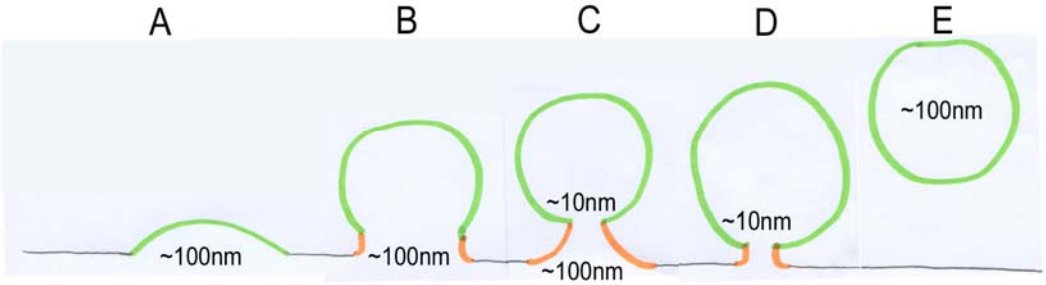


Figure 1.2. Membrane curvature change during clathrin-mediated endocytosis (CME). (A) Low-curvature dimple structure formed at the CME initiation stage. (B) and (C) Hemi-spherical pit connected to membrane tube with dynamic curvatures. (D) Matured spherical membrane vesicles connected to a high-curvature membrane tube. (E) Isolated spherical membrane vesicles. The spherical membrane shape is colored in green; tubular membrane shapes are colored in orange and flat membrane is shown in black. The proteins and protein coats are not shown in the sketch

In the early stage of CME, the membrane curvature changes from a relatively flat state to a gently curved, spherical state (Figure 1.2.A). Several proteins actively modulate the curvature of the membrane at this step. The process initiates when adaptor proteins are localized to the PtdIns-(4,5)-P_2-enriched endocytic site, which further recruits the coating protein clathrin (Kirchhausen 1993; Smith et al., 1998; Owen et al., 2000). Clathrin assembles into a spherical cage on the membrane surface, in which membranes are molded into a vesicular structure, despite the fact that there is no direct interaction between the lipids and clathrin (Schmid, 1997; Kirchhausen, 2000;

Foton et al., 2004). Another important curvature modulator, ENTH domain-containing protein epsin also plays a role in driving membrane bending and promoting dimple formation (Ford et al., 2002). This is achieved by insertion of an amphipathic helix into the outer leaflet of the membrane (Ford et al., 2002; Campelo et al., 2010). It also has also been reported that the F-BAR domain-containing protein FCHo1 and 2 (FCHo1/2) are required for clathrin-coated vesicle (CCV) budding and marking sites of CCV formation (Henne et al., 2010; Henne et al., 2007).

The stalk region between clathrin-coated pits and the plasma membrane is a short, tubular membrane with variable length and curvature (Figure 1.2.B-D). The tubular diameter varies from above 100nm in the early stages to below 20 nm when the pits are ready to undergo fission. This unique structure is usually associated with several proteins including dynamin (Takei et al., 1996; Schmid et al., 1998; Menttlen et al., 2009) and/or BAR domain-containing proteins (Peter et al., 2004; Frost et al., 2009; Camilli, 2008; Henne et al., 2010). Dynamin is a large membrane-associated GTPase that specifically recognize PtdIns-(4, 5)-P_2 via its PH domain (Schmid et al., 1998). Dynamin can self-assemble into spiral-like protein filaments (Hinshaw and Schmid, 1995; Carr and Hinshaw, 1997), and can efficiently generate membrane tubes when incubated with liposomes (Sweitzer and Hinshaw, 1998). Structural studies revealed that dynamin facilitates membrane tubulation by forming a helical protein array on the top of the tube (Hinshaw, 2000; Zhang and Hinshaw, 2001). The mechanism of BAR domain-mediated membrane deformation will be discussed in detail in the next section.

Previous studies have suggested that the tubular membrane structure that is coated by dynamin can undergo fragmentation upon GTP incubation (Sweitzer and Hinshaw, 1998; Stowell et al., 1999; Marks et al., 2001). This implies that the second function of dynamin is to facilitate membrane fission during CME (Figure 1.2.E). The

prevailing model suggests that membrane fission may occur through an intermediate structure which has very high local curvature (Kozlovsky and Kozlov, 2003). The mechanical property of dynamin is coupled with its GTPase activity. Without GTP addition, dynamin deforms membrane bilayers into long, 40 nm wide tubes (Sweitzer and Hinshaw, 1998; Marks et al., 2001). Tube fragmentation is observed by negative-stained electron microscope when the samples are further treated with GTP (Sweitzer and Hinshaw, 1998; Marks et al., 2001). Interestly, the fragmentation ability, but not tubulation activity, disappears when the sample is immediately frozen after preparation, suggesting that the fragmentation activity *in vitro* is also influenced by the sample preparation method (Roux et al., 2006). Several mechanisms have been proposed for dynamin-mediated membrane fission. One proposed model suggests that dynamin helical filaments undergo an extension upon GTP hydrolysis, which pulls the vesicle away from plasma membrane and causes membrane fission (Stowell et al., 1999). The twisting model suggests that the fission occurs due to a rotatory movement of the dynamin helical turns within the dynamin lattice during GTP hydrolysis. The twisting would produce a longitudinal tension that is released by breaking the tubular structure (Roux et al., 2006). Both models link fission activity to a conformational change of dynamin filaments upon GTP hydrolysis. In contrast, the squeezing-and-relaxation model suggests that the membrane fission is a spontaneous process when dynamin is disassembled from highly curved lipid tubules upon GTP hydrolysis (Bashkirov et al., 2009). This mechanism does not require an entire helical filament, but may be efficient with only a few rounds of dynamin rings and therefore is more likely to be predominant mechanism *in vivo*

BAR domain super-family

The BAR domain was first identified as an evolutionary conserved region shared by yeast proteins Rvs161 and Rvs 167 and by Amphiphysin and by Bin 1 in metazoans. A database search based on sequence homology identified a super-family of BAR domain-containing proteins (Peter et al., 2004; Itoh et al., 2005). A large number of these proteins are involved in the cellular membrane remodeling process, such as vesicular trafficking, endocytosis, cell migration and division (Peter et al., 2004; Itoh et al., 2005; Frost et al., 2009).

A unique function of BAR domain proteins is their ability to generate distinct membrane morphologies when interacting with lipid bilayers. These structures include large tubes (30-100 nm), thin tubules (15-20 nm) and small vesicles (35-50 nm). Table 1.1 summarizes the different membrane morphologies that have been reported for known BAR domains. One of the well-characterized morphologies is the large tube. This was first reported for Amphiphysin (Takei et al., 1999) and Endophilin (Fasad et al., 2001), two endocytic proteins that interact with dynamin. The tubulation activity of both proteins is visible *in vivo* upon over-expression of the proteins (Peter et al., 2004; Gallop et al., 2006). These studies established that BAR domain-containing proteins as membrane curvature sensosr and tubulators.

N-BAR, F-BAR and I-BAR

Previous studies have suggested that the membrane sculpture function of BAR domains is intrinsically encoded in their unique molecular structures, as first reported for the Amphiphysin BAR domain (Peter et al., 2004, Figure 1.3A). With more than twenty atomic structures that have been determined for BAR domain proteins, our understanding of the BAR domain-mediated membrane deformation process has advanced significantly. All the BAR domains are dimeric protein modules of a curved

rod-like shape. Each monomer consists of three long alpha helices, which forms a six-helix bundle at the center of the dimer. The charged residues are distributed unevenly on the dimer surface, which produces a positive electrostatic potential on either the concave surface (N-BAR, F-BAR) or the convex surface (I-BAR). These structural properties provide a strong tendency of BAR domains to interact with negatively charged, curved membrane surfaces. Based on their structural elements and sign of their curvature, BAR domains can be further classified into three subgroups: N-BAR domains, F-BAR domains and I-BAR domains.

Table 1.1. Diverse membrane morphologies generated by BAR domain proteins

BAR domain	*in vitro*	*in vivo*	Reference
BAR/N-BAR			
Arfaptin	25-40nm tubes	tubular invaginations	Peter et al., 2004
Amphiphysin	25-50nm tubes, 30-50nm vesicles	tubular invaginations	Farad et al., 1999; Peter et al., 2004
Endophilin	25-40nm tubes,7nm micelle tubes, 36nm pearling tubes	tubular invaginations	Gallop et al., 2005 Masuda et al.,2005 Misuno et al.,2010
SNX9	40nm tubes	tubular invaginations	Pylypenko et al., 2007
nadrin2	25-40nm tubes	tubular invaginations	Peter et al.,2004
centaurin-beta2	25-40nm tubes	tubular invaginations	Peter et al.,2004
oligophrenin1	25-40nm tubes	tubular invaginations	Peter et al.,2004
F-BAR			
CIP4	50-200nm tubes	tubular invaginations	Itoh et al., 2005; Shimata et al.,2007
FBP17	50-200nm tubes	tubular invaginations	Itoh et al., 2005; Shimata et al.,2007
TOCA-1	20nm tubules, 50-100nm tubes	tubular invaginations	Itoh et al., 2005; Henne et al., 2007
PACSIN1	17nm tubules, 50-120nm tubes, 35nm vesicles	tubular invaginations	Itoh et al., 2005; Wang et al., 2009; Zhao et al., 2010
PACSIN2	17nm tubules, 50-120nm tubes, 35nm vesicles	tubular invaginations and microspikes	Wang et al.,2009; Shimata et al., 2010
I-BAR			
srGAP2	150-200nm inward surface	Filopodia	Guerrier et al., 2009
MIM/IRS52	~80nm inward tubes	Filopodia	Mattila et al.,2007

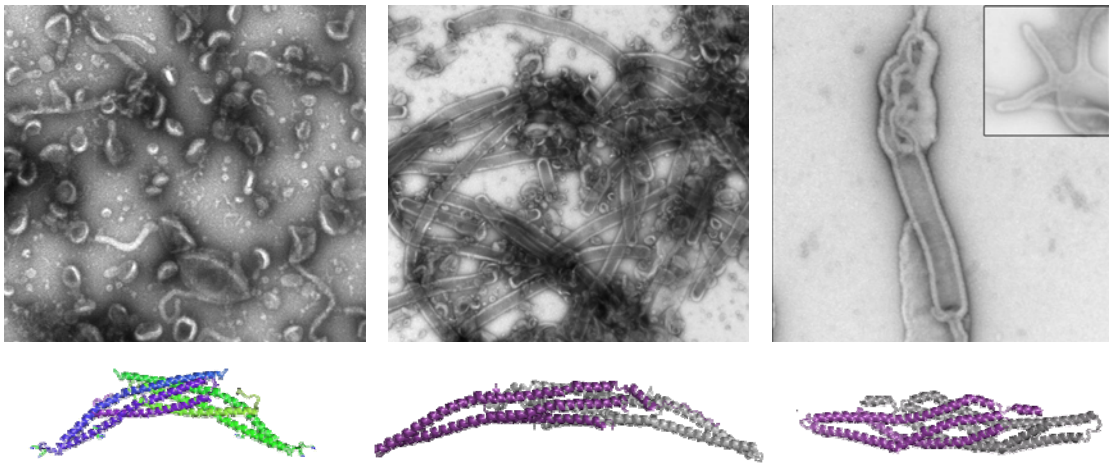


Figure 1.3. Membrane tubulation and ribbon presentation of representative crystal structures in BAR domain super-family proteins. Membrane deformation (top) by BAR domain superfamily proteins is shown by negative-stain electron microscopy. Proteins (10 uM) were incubated with Folch Fraction I lipids. The structures (bottom) of the N-BAR domain of Amphiphysin (Peter et al., 2004), the F-BAR domain of FBP14 (Shimada et al., 2007), and the I-BAR domain of IRSp53 (Mallita et al., 2007) are shown as ribbons. The proteins form tight dimers, and the two monomers are colored differently. The structures were drawn using the program PyMol.

In addition to a curved surface, N-BAR domain-containing proteins employ additional structural units that are essential for the membrane tubulation activity. Endophilin (Gallop et al., 2006; Masuda et al., 2006), Amphiphysin (Peter et al., 2004) and Sorting Nexin 9 (Pylypenko et al., 2007; Wang et al., 2008) are classified as N-BAR domains. Both Amphiphysin and Endophilin contain a ~15-residue peptide at their N-terminus which is essential for their membrane tubulation activity (see below). In Sorting Nexin 9, although there is no such helical structure within the BAR domain, a predicted helical region that is located at the very N-terminus of the PX domain has been shown to be essential for maintaining its function, presumably through the similar mechanism proposed for Endophilin and Amphiphysin (Pylypenko et al., 2007).

The F-BAR domains contain a conserved motif known as FCH domain, followed by the coiled-coil sequences. For this reason, F-BAR domains are also

referred to as EFC domains (Extended FCH domains). Most of the F-BAR domain-containing proteins are adaptor proteins that are involved in cellular trafficking and endocytosis, similar to the BAR domain-containing proteins. F-BAR domains also tubulate membranes (Tsujita et al., 2006; Itoh and De Camilli, 2005). Crystallographic studies revealed that the EFC domain indeed adopts the similar structural fold as BAR domains (Henne et al., 2007; Shimada et al., 2007; Wang et al., 2009), therefore allowing us to classify F-BAR domains as part of the BAR domain superfamily (Frost et al., 2008). In contrast to N-BAR domains, F-BAR domains are more elongated and less curved, with a maximaum diameter of around 220 nm, which is about 80 nm longer than that of an N-BAR domain (Figure 1.3., bottom).

N-BAR and F-BAR domains have their concaved surfaces positively charged; therefore they both stabilize positive membrane curvature. I-BAR domains instead have a positively charged convex surface. This structural feature makes I-BAR domains generate negative membrane curvature (Figure 1.3.C). For example, the I-BAR domain in IRSp53 and MIM (Mattila et al., 2007; Lee et al., 2007) can cause membrane invagination when interacting with bilayers and play a role in forming filopodial protrusions in cells.

BAR domain-containing proteins

BAR domains are often found connected to other functional domains. In general, these multi domain-containing proteins play diverse roles in cells, linking cellular membranes to other components and contribute to the signal transduction pathways. The most frequently found domains include protein interaction domains, lipid interacting domains and domains with catalytic activity (Figure 1.4.).

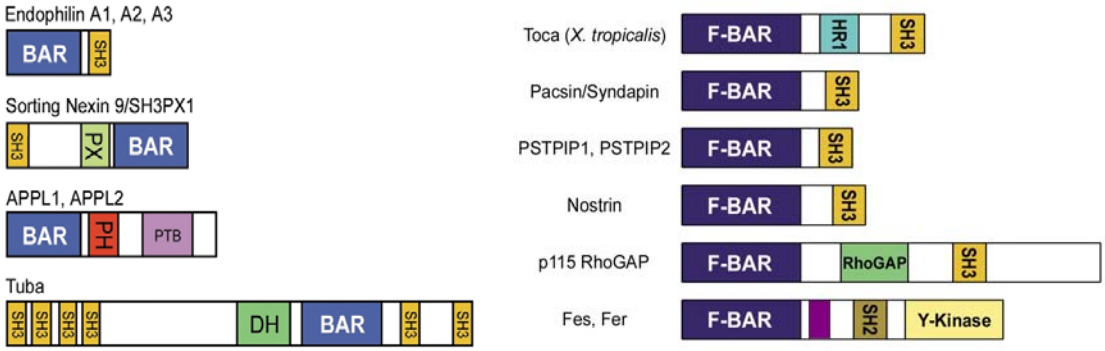


Figure 1.4. Domain organization of BAR domain containing proteins. The domains are drawn to scale considering their polypeptide chain lengths.

Some BAR domain-containing proteins also contain one or more protein interaction domains that can link BAR proteins to other proteins. Quite frequently SH3 or PDZ domain occurs in these proteins. The SH3 domain specifically interacts with proline rich domains (PRDs) via a consensus P-x-x-P motif (where x is an arbitrary hydrophobic residue). The SH3-PRD domain interaction has both adaptor and regulatory functions. The interaction physically recruit these PRD domain-containing proteins onto the membrane surface, therefore regulates membrane trafficking and signal transduction pathways spatially and temporally, as seen in the dynamin-mediated membrane fission process. In some other cases, SH3 domain binding can regulate the catalytic function of target proteins. For example, cytoskeletal regulator N-WASP can be hyper activated upon binding to Pacsin's SH3 domain, which in turn triggers actin polymerization via interaction with the Arp2/3 complex (Padrick et al., 2008; Padrick and Rosen, 2010).

BAR domain-containing proteins may also have additional membrane interacting domains such as PX or PH domain. These domains usually confer membrane localization specificity by introducing binding specificity to PIPs, in addition to the curvature preference through the BAR domain dimer. In some cases,

the membrane interacting unit also contributes to membrane bending and curvature sensing. For example, the PX-BAR unit found in Sorting Nexin 9 has been shown to synergistically interact with membranes. Structural studies indicated that the dual domain unit forms a continuous membrane interacting surface spanning both, the PX and BAR domains (Badour et al., 2007; Pylypenko et al., 2007; Wang et al., 2008).

Many BAR domain proteins are also enzymes that actively participate in chemical modifications in cells. For example, the F-BAR domain-containing protein FES/FER is a non-receptor tyrosine kinase that regulates cell-cell adhesion and mediates signaling from the cell surface to the cytoskeleton via growth factor receptors (Itoh et al., 2010; Smith et al., 2010; Shapovalova et al., 2007; Parsons et al., 2007; Sangarar et al., 2007). Other catalytic domains found in association with BAR or F-BAR domains include GTPase-activating protein (GAP) and Guanine-nucleotide exchange factors (GEF) domains. In some cases, it has been demonstrated that BAR domains confer an auto-inhibitory role on GAP activity of GRAF1 and OPHN1, presumably via a direct intramolecular interaction (Eberth et al., 2009).

Previous studies have concluded that BAR domain-containing proteins possess sophisticated intramolecular regulatory mechanisms that are critical for their functions. Fine-tuning of their functionality can be achieved by ligand binding, allosteric interactions, and/or changes in sub-cellular localization. However, it is still unclear how these intricate processes are achieved mainly due to the lack of structural models. Part of this thesis aims at investigating the regulatory mechanisms of three BAR domain superfamily proteins: Pacsin, Sorting Nexin 9 and Endophilin A1.

Pacsins (also known as Syndapins or FAP52) comprise a conserved family of peripheral membrane proteins in eukaryotes that play a central role in synaptic vesicle recycling and receptor-mediated endocytosis, ultimately controlling the cell surface expression of transmembrane proteins (Anggono et al., 2006; Modregger et al., 2000;

Qualmann et al., 2000; Qualmann et al., 1999). The 3 isoforms of Pacsin in higher eukaryotes differ in their expression pattern, cargo selectivity, and recruitment of other proteins involved in endocytosis. Pacsin 1 is a brain-specific isoform and has been implicated in neuronal transmission and the neuropathology of Huntington's disease, whereas Pacsin 2 and 3 are more widely expressed (Modregger et al., 2000; DiProspero et al., 2004; Modregger et al., 2002). On a molecular level, Pacsins bind to Dynamin and Eps15 homology domain (EHD)-containing proteins, nucleotide-dependent enzymes with membrane remodeling and fission activities, as well as the actin cytoskeleton via neural Wiskott-Aldrich syndrome protein (N-WASP), linking membrane trafficking and cytoskeletal rearrangements (Qualmann et al., 1999; Braun et al., 2005; Kessels et al., 2004). Although Pacsin isoforms bind a specific subset of proteins via the central linker region and their SH3 domain, they all share a highly conserved Fer-CIP4 homology-BAR (F-BAR) domain at their N terminus (Itoh et al., 2005).

Sorting Nexins form a family of peripheral membrane proteins characterized by a conserved PX domain and are involved in intracellular membrane trafficking (Carlton and Cullen, 2005; Worby and Dixon, 2002). Similarly to Endophilin and Amphiphysin, Snx9 (also known as SH3PX1) contains an N-terminal SH3 domain with affinity for dynamin, synaptojanin, and N-WASP, and this domain provides a physical link to vesicle scission and endocytosis (Badour et al., 2007; Ringstad et al., 1997; Shin et al., 2007; Yarar et al., 2007). The SH3 domain also binds to the non-receptor-activated Cdc42-associated kinase (ACK) associating Snx9 with the degradation of epidermal growth factor receptors (Lin et al., 2002; Yeow-Fong et al., 2005). Receptor down regulation requires dimerization of the C-terminal BAR domain and ACK-mediated phosphorylation of Snx9 (Childress et al., 2006). The SH3 domain connects to the C-terminal PX-BAR module via a low-complexity region that

mediates binding to clathrin light chains and the appendage domain of AP-2, linking Snx9 to clathrin-mediated endocytosis (Lundmark and Carlsson, 2002; Lundmark and Carlsson, 2003; Soulet et al., 2005).

Endophilin A1 (also called SH3P4 and SH3GL2) is a peripheral membrane protein that plays multiple roles in Clathrin mediated endocytosis (Ringstad et al., 1999; Schuske et al, 2003; Galli and Haucke, 2004; Wenk and De Camilli, 2004). Similar to Amphiphysin, Endophilin A1 is highly enriched in brains, interacting with dynamin and synaptojanin via its C-terminal SH3 domain (Schmidt et al., 1999; Ringstad et al., 2001; Schuske et al., 2003; de Heuvel et al., 1997; Verstreken et al., 2003). Although the interaction between Endophilin A1 and dynamin is much weaker in solution, inhibition of the interaction between endophilin A1 and dynamin resulted in deeply invaginated clathrin-coated pits in synapses, suggesting its role in synaptic vesicle maturation and membrane remodeling (Simpsion et al., 1999; Schuske et al., 2003; Ringstad et al., 2001). Dynamin and endophilin could form a complex on liposomes, similar to what has been described for dynamin-amphiphysin (Farad et al., 2001). Noticeably, membrane tubules that are coated by endophilin and dynamin cannot undergo vesiculation upon GTP hydrolysis, suggesting an inhibitory role of Endophilin A1 in membrane fission. Endophilin A1 can also deform membranes into long tubular structures on its own, similar to those generated by Amphiphysin (Farad et al., 2001).

Membrane deformation requires a large amount of free energy. This section introduces the fundamental theory of the membrane deformation process using the most simplistic formulas, and describes the prevailing understanding of the BAR domain-mediated membrane tubulation mechanisms.

Protein-mediated membrane deformation

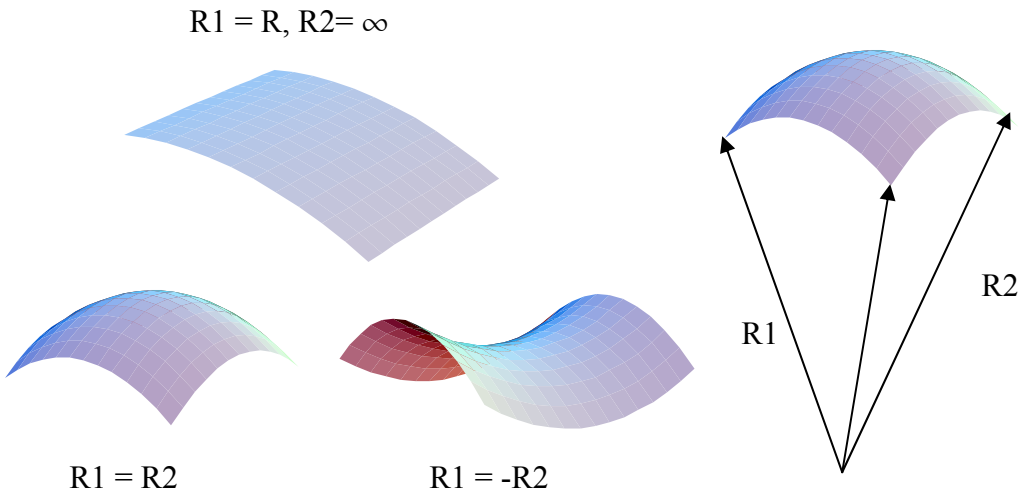


Figure1.5. Geometric description of membrane surfaces. Tubular, spherical and saddle surface sare plotted using Mathmatica. The surfaces are drawn at arbitrary scales. Tubular membrane surfaces have one definite curvature and one zero curvature. A spherical surface usually has two definite curvatures that are identical to each other. The saddle surface has one negative curvature that bends towards the opposite direction as the other positive curvature

Elastic membrane bending theory

A membrane surface can be considered as the integration of infinite small surface units that can be described by the two parameters known as principle curvatures (Kozlov and Zimmberg, 2005): $\left(\frac{1}{R_1},\frac{1}{R_2}\right)$. For an infinitely long tube with radius R, the principle curvature is $\left(\frac{1}{R},0\right)$. For a sphere with radius R, the principle curvature is $\left(\frac{1}{R},\frac{1}{R}\right)$ (Figure 1.5). Although the plasma membrane is assumed to be flat in the following calculations, membrane bilayers are believed to have intrinsic non-zero curvature, which depends on the lipid composition and phase behavior. Tubular, vesicular and flat membranes are frequently observed in biology systems. The saddle surface, in which one of the principle curvatures is negative, also exists

during membrane remodeling processes. For example, the neck region in the vesicle budding process can be described by a saddle surface.

Since in an ideal case membrane bilayers maintain their intrinsic curvature at equilibrium, any attempt to change its curvature consumes energy. According to the Helfrich bending theory, the elastic energy that is required to deform a membrane from its intrinsic curvature R_0 to R_1 can be estimated as (Deuling and Helfrich, 1976):

$$E = \frac{\kappa}{2}\oint dA\left(\frac{1}{R_1} - \frac{1}{R_0}\right)^2$$

κ is the membrane bending moduli, also known as rigidity. It can be experimentally determined through membrane tethering such as thermally driven fluctuation and AFM pulling. It is assumed that the radius of curvature is much larger than the membrane thickness. It should be noted that the bending energy is not the only term that contributes to the free energy of the bending process. Other terms include Gaussian curvature energy and the stretch energy. Within the scope of this discussion, these terms were ignored due to their minor overall contributions.

With the Helfrich theory, the energy cost to generate a tube (diameter 2R, Length L) from a flat membrane is:

$$E_t = \pi\kappa \frac{L}{R}$$

The energy cost per unit area (energy density) will therefore be:

$$\rho_t = \frac{E_t}{2\pi RL} = \frac{\kappa}{2} \times \frac{1}{R^2}$$

For a spherical shape, the energy cost will be:

$$\rho_v = 2\kappa \times \frac{1}{R^2}\Bigg|$$

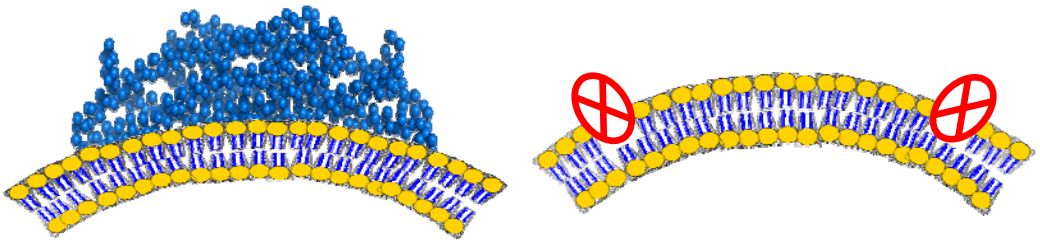


Figure1.6. Cartoon illustration of the protein-mediated membrane deformation mechanism. (A) Scaffolding mechanism. Proteins with positively charged surface interact with membrane bilayers, and restrict its morphology through rigid protein structure elements. (B) Hydrophobic insertion mechanism. Proteins with amphipathic structural motifs partially insert these into the outer leaflet of the membrane bilayer, and generate a highly curved membrane state by increasing the surface area difference between two membrane leaflets.

Scaffolding mechanism

The BAR domain-mediated membrane tubulation process requires collective behavior of multiple proteins. Although BAR domains exist as dimer in solution, they form higher order oligmers when membrane is present (Yarar et al., 2009; Suresh and Edwardson, 2010). High-resolution cryogenic EM studies of the CIP4 F-BAR domain on membrane tubes showed a periodic coating pattern outside the membrane bilayer, presumably formed by the F-BAR domains (Shibata et al., 2007). The most convincing evidence is the EM reconstruction study of FBP17-coated membrane tubes at 19Å resolution (Frost et al., 2008), in which the electron density of the coating region perfectly matches the atomic structure of FBP17. Furthermore, the reconstruction study also suggested that BAR domains modulate the tube size by forming different lattices on the membrane. The lattice formation turns out to be a general process for BAR domain proteins. Similar protein lattice have been also observed for Endophilin BAR domain-generated tubes (Mizuno et al., 2010). These experimental data strongly suggest that BAR domains can drive membrane tubulation

by forming a helical lattice on the membrane surface. This mechanism has been referred to as scaffolding mechanism (Figure 1.6.A).

The scaffolding mechanism suggests that the rigidity of the curved protein and the electrostatic interactions with lipid bilayers are crucial compoments for the membrane tubulation activity. Indeed, mutations that weaken the binding affinity of proteins to the membrane also abolish tubulation activity. (Peter et al., 2004; Masuda et al., 2006; Gallop et al., 2006; Henne et al., 2007; Frost et al., 2008). Furthermore, the protein oligomerization contacts, which maintain the protein scaffold, are also critical for membrane tubulation activity, despite the fact that they are not part of the direct membrane-interacting surface (Frost et al., 2008). Theoretical calculations indicate that the observed protein lattice is sufficient enough to provide bending energy for membrane deformation.

Hydrophobic insertion mechanism

N-BAR domains require additional structural motifs for tubulating membranes *in vitro*. In the case of Endophilin, the first 16 N-terminal residues are critical for its tubulation activity (Gallop et al., 2006; Masuda et al., 2006). Although this region is disordered in solution, the sequence analysis indicates that it has the potential to form an amphipathic helix, which is supported by Circular Dichroism studies (Gallop et al., 2006) and the crystal structure of an endophilin mutant (Masuda et al., 2006). Furthermore, EPR using spin labeled Endophilin has been applied to show that the helix dips into one membrane leaflet about ½ of its length, and therefore can change local membrane curvature through a "wedge" effect (Gallop et al., 2006; Jao et al., 2010; Figure 1.6.B). In addition, the appendage structure, which is located at the central six-helix bundle of endophilin, has also been proposed to interact with membranes in a same manner (Masuda et al., 2006; Gallop et al., 2006). Such motifs

have been also identified in Amphiphysin (Peter et al., 2004) and Sorting Nexin 9 (Pylypenko et al., 2008). The theoretical analysis indicates that such wedging effects on its own can have a significant impact on the membrane curvature (Campelo et al., 2010; Campelo et al., 2008). In part, this may explain the fact that N-BAR domains usually generate more curved membrane tubes than F-BAR domains, although both can form protein coats on the membrane surface.

It is still somewhat unclear if the mechanisms derived from theorical considerations and *in vitro* system can be used to explain the tubular structures that naturally occurred *in vivo*. Unlike *in vitro* assays that are usually performed with excess amount of proteins, the tubular structures that are generated during CME are usually associated with limited number of protein molecules and are dynamic membrane structures.

Overview

Previous studies on isolated BAR domains show significant progress towards understanding its membrane tubulation function. Despite this great advance, BAR domain-mediated membrane deformation, especially the mechanism to generate highly curved membrane structures remains a mystery. Furthermore, the regulatory role of other domains and proteins on BAR domain's function remains elusive. This thesis aims to systematically study BAR domain mediated membrane deformation, and to investigate the role of intra- and inter molecular interaction in regulating BAR domain function. In chapter two, I performed biophysical studies on the membrane deformation process by isolated F-BAR domains. In chapter three and four, I studied how protein interacting domain and lipid interacting domain may have a significant impact on regulating BAR domain's functionality, and *visa versa*.

Chapter two describes the membrane deformation study of the isolated F-BAR domain from Pacsins. Here, we found that Pacsin's F-BAR domain has a diverse membrane deformation activity. The F-BAR domain of Pacsin generates three distinct membrane morphologies under identical experimental conditions. Via biophysical approaches, we explore the molecular basis of this phenomenon, and provide a detailed model for its contribution to membrane remodeling. I further show that this diverse membrane deformation activity is tightly regulated in the full-length Pacsin.

Chapter three studies other factors that influence protein mediated membrane deformation process by Pacsins. I found that the unique atomic structure was a necessary but not sufficient for the deformation diversity. Membrane properties contribute significantly to this process. Most importantly, we found that lipid composition and preparation methods play important roles in regulating this process. Furthermore, we extended our study to full-length Pacsins and all three human isoforms. We found that although Pacsin 1's F-BAR is highly active, this membrane sculpting activity is much inhibited in the context of the full-length protein suggesting an autoinhibitory conformation. Via biophysical analysis, we provide a molecular model for a plausible autoinhibition and activation mechansim. Furthermore, we show that the Pacsin 1 full-length becomes an extraordinary membrane vesiculator when it interacts with the PRD of dynamin via its SH3 domain.

In Chapter four, I studied how BAR domains may influence the function of adjacent domains. In the first part, we investigate the molecular basis of membrane-mediated Endophilin A1 activation, and provide a model for its activation based on structural information. In this study, we show that the membrane bilayer can actively regulate proteins, being an important player in signal transduction pathways. In the second part, we investigate the molecular mechanism of the membrane binding unit of

Sorting Nexin 9, and provide a structural model for the synergistic effect of PX and BAR domains.

GENETICALLY ENCODED CALCIUM INDICATORS

Calcium plays important roles in various cellular signaling events. In cardiac cells, the calcium level is tightly controlled either by cellular calcium binding proteins such as calsequestrin and calreticulin, or by L-type voltage-dependent calcium channels. Studies have shown that reduction of cardiac calsequestrin triggers ventricular arrhythmias in mice (Chopra et al., 2007) and over-expression of calreticulin can cause hypertrophic cardiomyopathy (Mickalak et al., 2004; Chiu et al., 2007). The dysfunction of L-type voltage-dependent calcium channel greatly alters the myocardium contraction and leads to several cardiac diseases (Bodi et al., 2005; Streiessnig, 2005). Therefore, *in vivo* monitoring the cellular calcium flux is of great importance for understanding the pathogenesis of cardiac diseases.

Among several available calcium sensors, GECIs (Genetically Encoded Calcium Indicators) are invaluable tools because they can be expressed in live cells and animals and monitored using fluorescence microscopy. A high-quality optical calcium sensor for *in vivo* studies should satisfy several criteria. It should be well expressed producing stable protein in the mammalian cell lines. The sensor signal must be calcium sensitive, specific and detectable. Finally, the calcium response rate should be fast enough to monitor high-frequency calcium flux in live cells. Over the past few years, several generations of GECIs have been designed to meet those criteria, and substantial progress has been made toward high-quality optical sensors (Baird et al., 1999; Griesbeck et al., 2001; Nakai et al., 2001; Nagai et al., 2001; Ohkura et al., 2005). The recently developed sensor GCaMP2 can capture the calcium flux transition from diastole to systole states in the spontaneously beating heart of an

anaesthetized and ventilated mouse with strong and robust signal intensity (Tallini et al., 2006). Despite the success, the remaining challenges are clear. The reported GCaMP2 has a transition rate of the order of 10-100 ms. This transition rate is not fast enough to monitor high-frequency physiological process such as calcium flux in myocardium contraction and in neuronal cells. This greatly limits GCaMP2 to be used in studying several important cardiac pathogenic scenarios. The thesis is designed to understand the working mechanism of GCaMP2 and to apply the knowledge to develop next generation GECIs.

Green Fluorescent Protein

Green Fluorescent Protein (GFP) was first purified from *Aequorea Victoria,* a jelly fish that lives at the west coast of America. In *Aequorea Victoria*, green emission occurs when GFP receives blue light from aequorin, a protein that glows upon Calcium binding (Shimomura, 1995). Recombinant GFP can be expressed in bacterial cells and can be purified to crystallization-quality purity. Purified GFP glows with a bright green color under ambient light. The fluorescence intensity is influenced by intrinsic factors such as amino acid composition and external factors such as pH, temperature and pressure (Tsien, 1998).

At molecular level, a fluorescent protein can be divided into chromophore unit and regulatory unit. Chromophore is the minimal structural unit that is responsible for its fluorescence, such as Coelenterazine in Aequorin. The regulatory unit is the peptide region that interacts and regulates the chromophore fluorescence. Unlike Aequorin, in which the chromophore is held by noncovalent bonds (Head et al., 2000), GFP's chromophore is covalently linked to the protein and is formed from three consecutive amino acids Ser-Tyr-Gly at the central region of the protein (Ormo et al., 1996). Wild-type GFP possesses an open-end beta barrel structure, consisting of eleven anti-

parallel β-sheets with one alpha helix running through the center (Yang et al., 1996; Brejc et al., 1997; Figure 1.7.A). The chromophore is located at the center of the beta-barrel and is connected to the helical structure. The inner barrel environment is largely hydrophobic, with some charged patches that are important for chromophore maturation and stability.

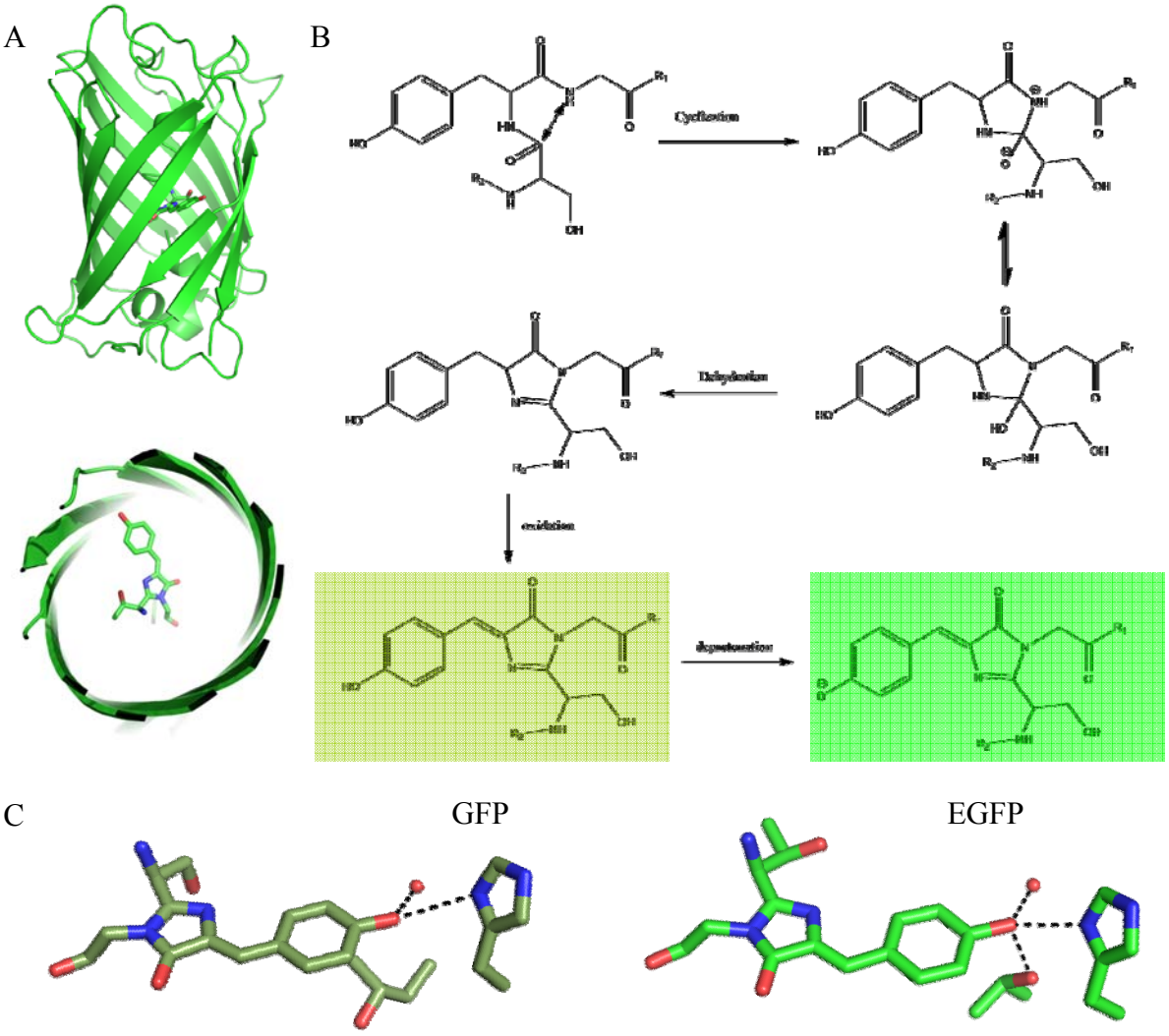


Figure 1.7. Structure of Green Fluorescent Protein. (A) Overall crystal structure of Green Fluorescent Protein. (PDB: 1EMA. the 11-strain β barrel is shown as cartoon and the chromophore is highlighted as sticks. (B) Proposed maturation mechanism for the GFP chromophore. The colored box indicates the fluorescence of the corresponding structures. (C) Chromophore coordination of EGFP and GFP. The dashed lines represent hydrogen bonds. Threnine 203 adopts different conformations in the two structures.

Chromophore Structure

The GFP chromophore contains 4-(*p*-hydroxybenzylidene) imidazolid-5-one as core structure, which was first proposed based on biochemical studies on synthetic small molecules (Shimomura, 1979), and later proven by the X-ray crystal structure of GFP (Ormo et al., 1996; Yang et al., 1996; Brejc et al., 1997). In the crystal structure, two rings are conjugated with a double bond, forming an extended coplanar π system in the beta-barrel. Interestly, the synthetic chromophore only glows at extremely low temperature, since the planar conformation cannot be maintained in solution due to thermal fluctuations. This evidence suggests that the GFP backbone contributes significantly to the stabilization of such a planar system (Figure 1.7.A-B).

Chromophore Maturation

The prevailing model suggests that GFP chromophore is formed through a post-translational cyclization process that is catalyzed by the rest of peptide chain in the protein (Heim et al., 1994; Cubit et al., 1995; Figure 1.7.B). The cyclization of a five-membered ring structure is thought to be the first step towards chromophore maturation. This is achieved by an unusual turn structure adopted by the triple residues, in which the Nitrogen of Glycine attacks the C=O moeity of the Tyrosine. Interestingly, such cyclization reaction can be extremely slow in bulk solution, suggesting that the unique GFP structure catalyzes this reaction, presumably by reducign the conformational freedom of tri-peptide. A second critical step for chromophore maturation is the formation of the imdazole ring, most likely involving exogenous oxygen that serves as an oxidant.

Deprotonation

Although high resolution crystal structures contributed significantly to our understandings of the GFP chemical structure, they can not immediately tell the

protonation state of each residue since the scattering power of hydrogens is extremely low. Therefore, the determination of the chromophore protonation states usually requires other, additional evidence such as spectral properties and theoretical calculations. Wild-type GFP has two absorbance maxima. The major peak is at 398 nm and the minor peak is around 475 nm (Tsien, 1998). This indicates that the chromophore must exist in two different chemical states. Further evidence indicated that the dual absorbance can be explained by two different protonation states of the GFP chromophore. A solvent exposed chromophore has single absorbance at 398 nm at acidic pH, which can be assigned to a neutral chromophore. In contrast, at high pH when the protein remains folded, the 475 nm peak dominates the spectrum. Therefore, 475 nm is likely due to a deprotonated state. Interestly, the emission spectra at both absorbance maxima are very similar, with peaks at around 508 nm. This phenomenon can be explained by the photo-isomorization model. Theoretical studies have indicated that the phenolic oxygen becomes vey acidic when in its excited state. Upon excitation, the neutral specie will loose the proton and become the anionic species (Figure 1.7.B-C).

Enhanced GFP

A single mutation, S65T, significantly changes GFP performance *in vitro*, resulting in a much brighter fluorescent marker known as the prototype of Enhanced Green Fluorescent Protein (EGFP). Noticeably, EGFP only has one single absorbance peak at 480 nm, and the absorbance at 395 nm is negligible. This suggests that the chromophore is most likely trapped in the deprotonated state. This type of GFP is therefore classified as phenolate anion according to Roger Tsien's classification system. In the crystal structure, the phenolic oxygen interacts with three other residues and water molecules, adopting a sp^3 configuration (Figure 1.7.C). This further proves

that the chromophore is indeed deprotonated. Interestingly, Threnine 203 that interacts with the phenolic oxygen can adopt a second side chain conformation in the wild-type GFP with the hydroxyl group far away from the phenolate. In wild-type GFP, the chromophore interacts only with one residue (Histine 148), which agrees with the assumption that it is protonated. The detailed crystallographic analysis indicates that roughly 75% of the electron density around Threnine 203 is in the conformation that is far way from interacting with the phenolic oxygen. The other 25% is in a similarconformation to what has been reported for EGFP, which is consistent with the fact that the chromophore is in a protonation/deprotonation equilibrium in wild-type GFP (Figure 1.7.C). The mechanism of the S65T mutant spectral change can be rationalized at atomic level based on available crystal structures (Ormo et al., 1996; Tsien et al., 1998). In the S65T mutant, the oxygen of the Threnine forms a hydrogen bond with the main chain carbonyl oxygen of Val 61, which is not observed in the wild-type protein. Therefore, Threnine can only be a hydrogen bond acceptor in the interaction with Glutamic acid 222. This suggests that Glutamic acid 222 must be in a protonated state in the S65T mutant. In contrast to wild-type GFP, Glutamic acid 222 has a dual role of being a hydrogen donor or an acceptor. The protonated Glutamic acid 222 further alters the electron density at the position of the Threnine 205 residue, which changes the position of water molecule that interacts with the chromophore.

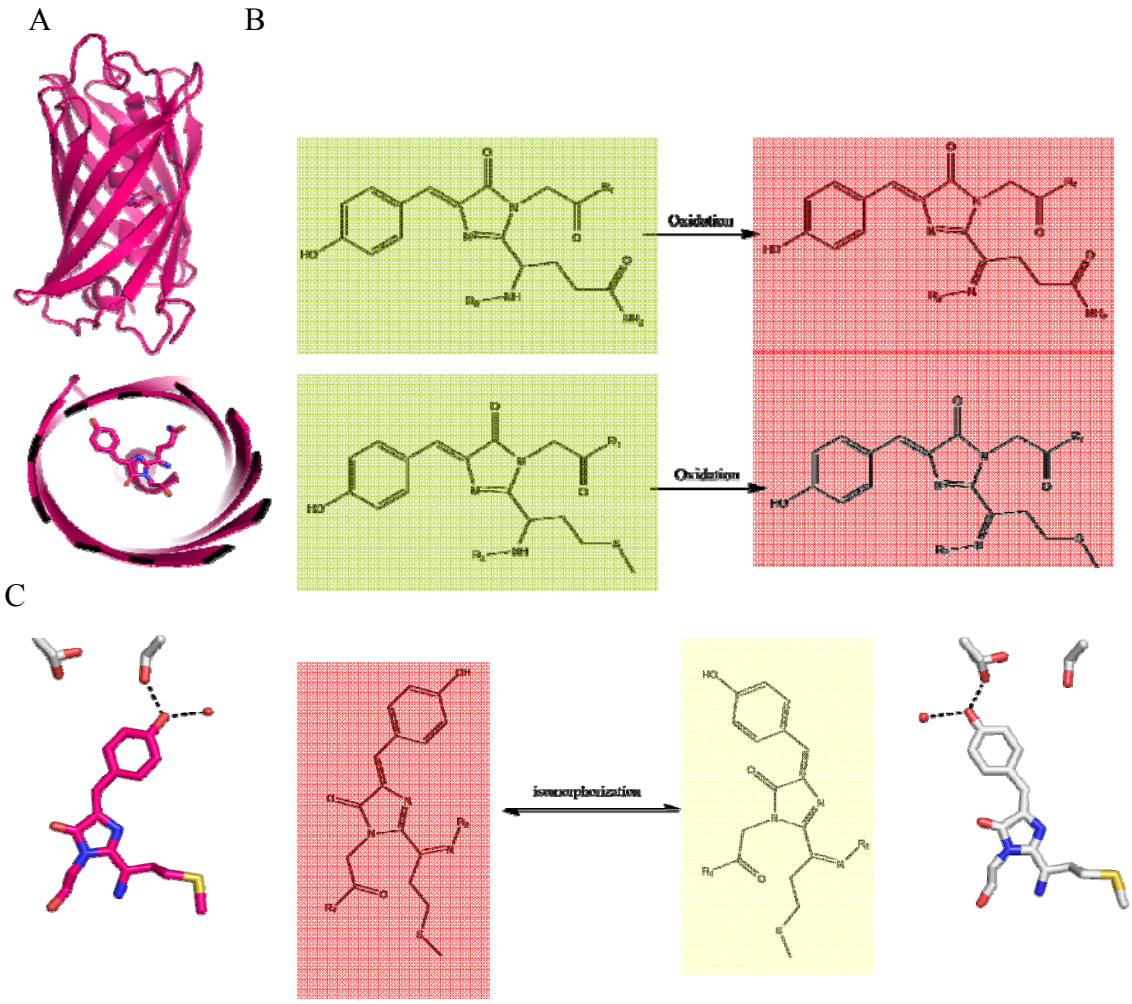


Figure 1.8. Structures of Red Fluorescent Proteins. (A) Crystal structure of Red Fluorescent Protein DsRed. (PDB: 1G7K. the 11-strain β barrel is as cartoon with the chromophore highlighted as sticks. (B) Proposed maturation mechanism for the DsRed chromophore (top). The colored boxes indicate the fluorescence of the corresponding structures. The hypothetic maturation model for mKate (bottom) is proposed based on the DsRed studies. (C) *cis-trans* isomorhporization of the mKate chromophore. The cis-(red) chromophore is stabilized by S143. The trans-(pale) chromophore is stabilized by S158. The hydrogen bonds are shown as dashed lines.

Red Fluorescent Protein

Fluorescent proteins of other colors can be derived from green fluorescent protein. For example, by evolutionary mutagenesis starting with the GFP sequence, a series of fluorescent proteins has been obtained, with colors ranging from blue to yellow (Heim et al., 1994; Tsien, 1998; Nattan et al., 2005). This advance enables the

simultaneous tracking of different proteins in cells and the development of sensors that are based on Forster resonance energy transfer. Despite great success, the development of red-emission fluorescence sensors from GFP proves challenging from the evolutionary mutagenesis approach. Fluorescent proteins of orange and red emission colors, especially these that emit above 650 nm have great *in vivo* cell imaging advantages (Ben et al., 2006; Tsien, 2003) since the tissue is more transparent to the light between 650-1100 nm. DsRed and eqFP578 are two naturally occured red fluorescent proteins that have provided novel opportunities for the development of far-red fluorescent proteins.

DsRed

DsRed was initially extracted from the coral Discosoma, one of the oldest stand-bys in the reef aquarium hobby (Baird et al., 2000; Gross et al., 2000). It has an orange-red fluorescence with an emission maximum at 583 nm. Although the sequence identity between GFP and DsRed is below 30%, DsRed adopts a structural fold that is very similar to the GFP molecule, consisting of 11-stranded β-barrel with one central helix and α-helical caps on the barrel's ends (Yarbrough et al., 2001; Figure 1.8.A). The chromophore is formed by three consecutive residues Gln-Tyr-Gly via an auto-cylization reaction similar to that observed in GFP, indicating that two proteins may share the same evolutionary origin. The core chemical structure of the DsRed Chromophore also contains a p-hydroxybenzylideneimidazolidinone, which further extends the pi conjugation by linking with the acylimine structure that is formed by the Gln alpha-carbon and nitrogen (Figure 1.8.B). Such extended conjugation system is essential for the red-shift in DsRed compared to GFP (Gross et al., 2000; Yarbrough et al., 2000). Indeed, the mutant K83R which lacks the acylimine bond emits green fluorescence instead of red color (Gross et al., 2000). This strongly

suggests that the green fluorescent chromophore is an intermediate structure that can be converted to the red fluorescent chromophore upon oxidation. Several inner barrel residues have great impact on the chromphore's optical properties, and through evolutionary mutagenesis, a series of red-zone fluorescent proteins, known as the mFruit series, have been developed for biomedical use (Nattan et al., 2005). mPlum, the DsRed derivative with the longest emission wavelength, has emission maxima at 650 nm, which is generally considered as far-red fluorescent protein (Wang et al., 2004).

eqFP578

Another widely used red fluorescent protein, mKate, is derived from eqFP578, a red fluorescent protein cloned from the sea anemone *Entacmaea quadricolor* (Shcherbo et al., 2007). mKate is a bright red fluorescent protein with excitation/emission peaks at 552/578 nm, and following directed, evolutionary mutagenesis, substantial improvements have been made with the best available version being mKate2 (Shcherbo et al., 2009). Similar to DsRed, mKate also adopts a 11-strain β-barrel conformation, with the chromophore containing the p-hydroxybenzylideneimidazolidinone and acrylimine moeities (Figure 1.8.B). Crystallographic studies revealed that the mKate chromophore can undergo cis-trans isomorphorization, considering the Tyrosine 64 phenolic ring with respect to the Cα-N bond. This process depends on the external solution pH (Pletnev et al., 2008). The cis-trans isomorphorization is spatially constrained in several other red fluorescent proteins since the cavity around the chromophore may not be large enough to adopt the intermediate structures.

Spectral studies indicated that the trans-conformation that is observed at low pH has extremely low fluorescence (Pletnev et al., 2008). The cis-conformation is

much brighter and is the major origin of its bright red emission at physiological pH. In the crystal structure, the residues that interact with the two conformations are largely identical, with the exception for Serine143 and Serine 158 that form hydrogen bonds with *cis*- and *trans*-chromophore, respectively (Figure 1.8.C). Arginine 197 also adopts a slightly different conformation in low and high pH conditions, presumably due to the different protonation states resulting from external pH difference. Arginine 197 may form cation-π interactions with the chromophore and is proposed to stabilize its anionic state (Lin et al., 2009). The mutation S158A that disrupts the Serine 158-chromophore interaction destabilizes the *trans*-conformation and increases the fluorescent intensity by roughly 3-fold (Pletnev et al., 2008; Shcherbo et al., 2009). Mutation S158C has similar effect on protein's fluorescence, but with much simplified photo-bleaching kinetics and enhanced photostability (Lin et al., 2009). Directed, evolutionary mutagensis further indentify M41G, S61C and Y197F that cause the bathometric shift of mKate, resulting in a far-red fluorescent protein with an excitation peak at 600 nm (Lin et al., 2009).

Fluorescent proteins with convertible colors

Fluorescent proteins (FPs) that can switch between distinct colors have made significant contributions to biomedical imaging technologies and fundamental molecular cell biology (Andresen et al., 2008; Adam et al., 2010). The most popular convertible fluorescent proteins are the ones that can be activated by photons. Photoactivable fluorescent proteins are great selective optical markers. They are widely applied to detect diffusion kinetics, protein trafficking, reaction turnover rates and super-resolution imaging (Patterson and Lippincott-Schwartz, 2002; Nienhaus et al., 2006; Subach et al., 2010; Kruhlak et al., 2009). This section introduces the

molecular mechanisms of irreversible photo-conversion and reversible photo-conversion for photo-activable proteins.

Kaede-like proteins

Kaede was the first widely used photoactivatable fluorescent sensor (Ryoko et al., 2002). The chromophore is composed of tri-peptide Histidine 62-Tyrosine 63-Glycine 64 that undergoes GFP-like autocyclization process to form 4-(p-hydroxybenzylidene)-5-imidazolinone. Before photoconversion, Kaede absorbs at 508nm and emits at 518nm. When exposed to 390 nm light, Kaede showed an emission maximum at 582 upon 480 excitation. The same chromophore is also present in the photoactivable protein EosFP (Wiedenmann et al., 2004) and its variant IrisFP (Adam et al., 2008). The crystallographic studies of EosFPs have shown that the bond between the amide nitrogen and the alpha carbon at Histidine 62 is cleavage via a formal β-elimination reaction, and a double bond between Histidine 62-Cα and -Cβ is formed to further extend the π-conjugation system (Figure 1.9.). This results in a new chromophore that emits red fluorescence (Nienhaus et al., 2005; Adam et al., 2008; Adam et al., 2009). In Kaede, EosFP and IrisFP, the activated chromophore adopts a *trans* double bond configuration between His^{62}-C_α and His^{62}-C_β. The KikGR variant possesses the C = C double bond in a *cis* configuration, presumably through and E1 elimination reaction (Tsutsui et al., 2009).

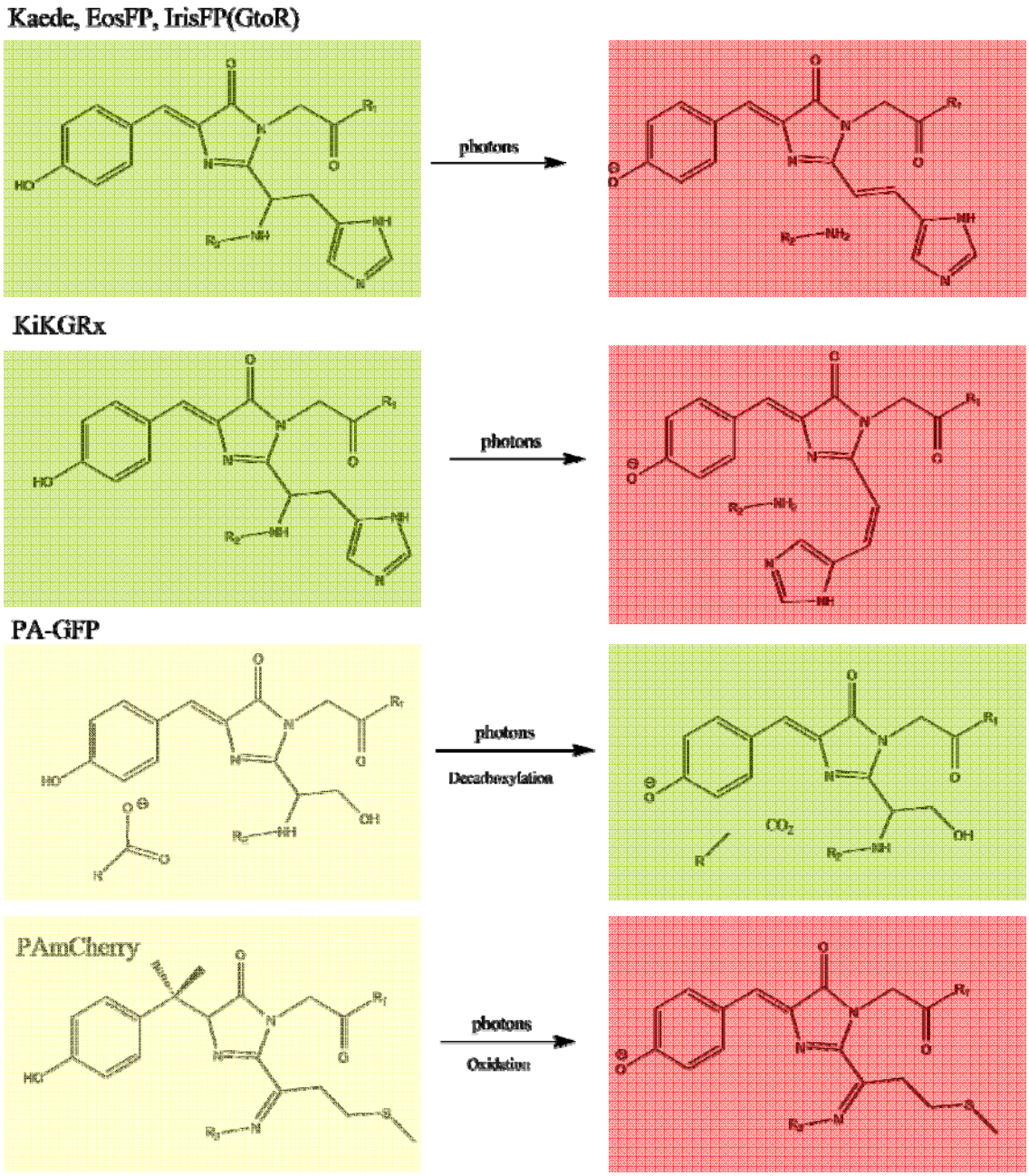


Figure 1.9. Photoactivation mechanisms for irreversible photo-activatable fluorescent proteins. Three types of chemical modifications are observed. Kaede-like proteins: an extra double bond forms between His^{62}-C_α and $-C_\beta$; PA-GFP: Glutamic acid 222 loses the carboxyl group; PAmCherry: an extra double bond forms between Tyr^{63}-C_α and $-C_\beta$.

PAGFP

Photo-activable GFP (PAGFP) and its derivatives use a different mechanism for their photoactivation. This class of PAFPs switches color by transitioning from a protonated to the deprotonated state. Before photoactivation, PAGFP contains a matured, typical GFP chromophore that only absorbs at 400nm, indicating that the majority of the chromophore is in the protonated state. Upon photoactivation, residue Glu222 undergoes a decarboxylation process and changes the hydrogen bond network around the chromophore (Hendreson et al., 2009). This rearrangement results in a new chromophore state which absorbs at 500nm, consistent with a depronated GFP chromorphore (Figure 1.9).

PAmCherry

Photo-activatable mCherry (PAmCherry) absorbs at 400 nm but is barely fluorescent. Illumination with 405 nm violet light results in a red-colored chromophore with absorbance and emission peaks at 564 and 595 nm, respectively. The photoactivation process requires oxygen, suggesting that an oxidation reaction is involved in the activation process. In the crystal structure, the off-state of PAmCherry has its Tyrosine^{63}-C_β in an sp^3 configuration, suggesting that the double bond between Tyr 63 C_α and C_β is not formed (Subach et al., 2009). As a result, the phenol ring is not in the same plane as the imdazole ring. In the activated on-state, the double bond matured and the chromophore emits red light due to an extended electron conjugation system (Figure 1.9).

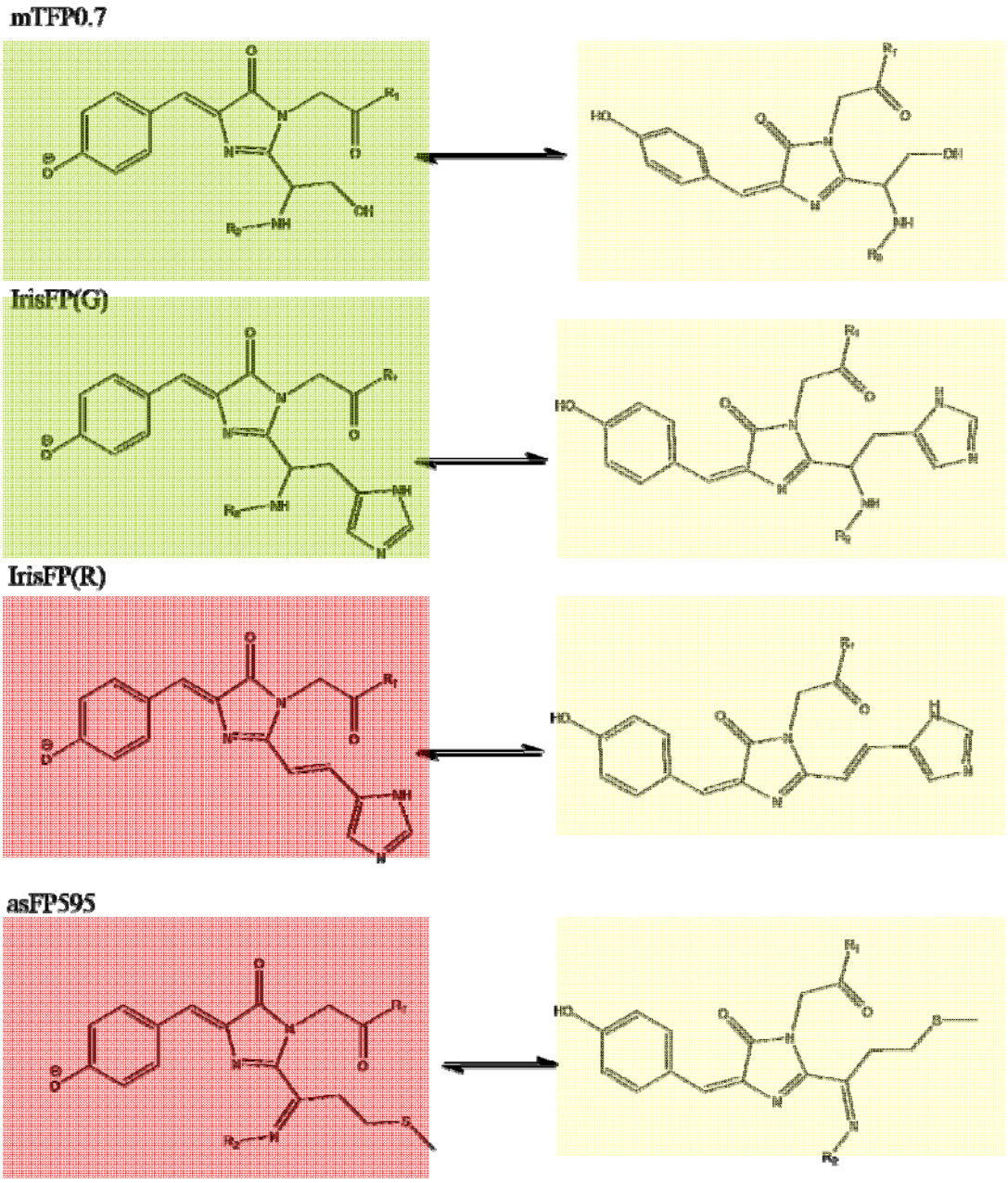


Figure 1.10. Photoactivation mechanisms for reversible PA fluorescent proteins. Four photoactivation reactions share two common features: first, photons induce chromophore isomorphorization, and second, the chromophore protonation status changes during the photoactivation process.

mTFP0.7

mTFP0.7 was discovered as an intermediate in the process of identifying monomeric teal FP (mTFP1). mTFP0.7 has a GFP-like Chromophore and emits brightly fluorescent upon 453 nm excitation, but quickly loose the fluorescence upon continuously stimulation. The fluorescence can completely recover at room temperature within several minutes. Crystallographic studies suggest that the reversible conversion is due to a photo-induced cis-trans isomorphorization process. In the fluorescent state, the chromophore adopts an almost co-planar cis-conformation, presumably in the deprotonated anionic state (Henderson et al., 2007). In the photo-bleached, dark state, the chromophore adopts a non-planar and flexible *trans* conformation with fewer hydrogen-bond partners to stabilize the hydroxyl group in the anionic form (Figure 1.10.). The planarity has been considered as an important factor for the quantum yield and fluorescence intensity. Thus, loss of planarity and the hydrogen bonds that stabilize the anionic hydroxyl group resulted in the dark state.

asFP595

asFP595 was originally cloned from the sea anemone *Anemonia sulcata*. The wild-type FP595 barely emits fluorescences. Upon exposure to green light, asFP595 is converted into a red-fluorescent form, which can be quickly quenched by 450 nm blue light. The "on-off" cycle can be repeated many times with slow switching kinetics. Different from DsRed-like red-fluorescent proteins, asFP595 has a keto group conjugated to the imdazole ring, which results from the hydration of the acrylimine bond between alpha carbon and nitrogen of residue Met63 (Tretyakova et al., 2007). In the off state, as observed in its A143S mutant, around 90% of the chromophore adopts a trans-conformation. In contrast, the on-state chromophore is almost 100% in the cis conformation, which is further stabilized by an additional hydrogen bond of the hydroxyphenyl ring to S143 (Andresen et al., 2005; Figure 1.10.). Noticeably,

theoretical calculation suggested that the chromophore protonation state also plays a role in controlling photoswitching of asFP595 (Schafer et al., 2008; Schafer et al., 2007; Grigorenko et al., 2006). Dronpa (Andresen et al., 2007) and eqFP611 (Nienhaus et al., 2008) adopt very similar photoactivation mechanisms. Another photoconvertable fluorescent protein that may involve a chromphore isomerization process is mGrape. Freshly extracted mGrape2 shows a blue absorption peak at 470 nm and a green emission peak at 545 nm. After exposure to ambient light, the 470 nm absorbance peak disappears and a 605 nm absorbance peak appears (Lin et al., 2009). It is hypothesized that the 470 nm absorbance peak represents a protonated DsRed-like chromophore, and that photoactivation occurs via isomerization of the chromophore phenolate initiated by absorption of blue light. The detailed mechanism remains elusive.

IrisFP

IrisFP is the eosFP variant that can undergo two different reversible photo-conversion events upon activation (Adam et al., 2008; Adam et al., 2009). Illumination of green-colored IrisFP via a 488 nm light leads to a non-fluorescent protein with spectral properties resembling a neutral GFP chromophore, which was confirmed by the crystal structure of non-fluorescent IrisFP. The protonation is achieved via a chromophore isomerization process, with the neutral form being stabilized by residue Serine142. The green fluorescent state can then be recovered by illuminating with 405 nm light. Chromophore isomerization also occurs when red IrisFP is illuminated with 532 nm light, which generates a second non-fluorescent neutral chromophore (Figure 1.10).

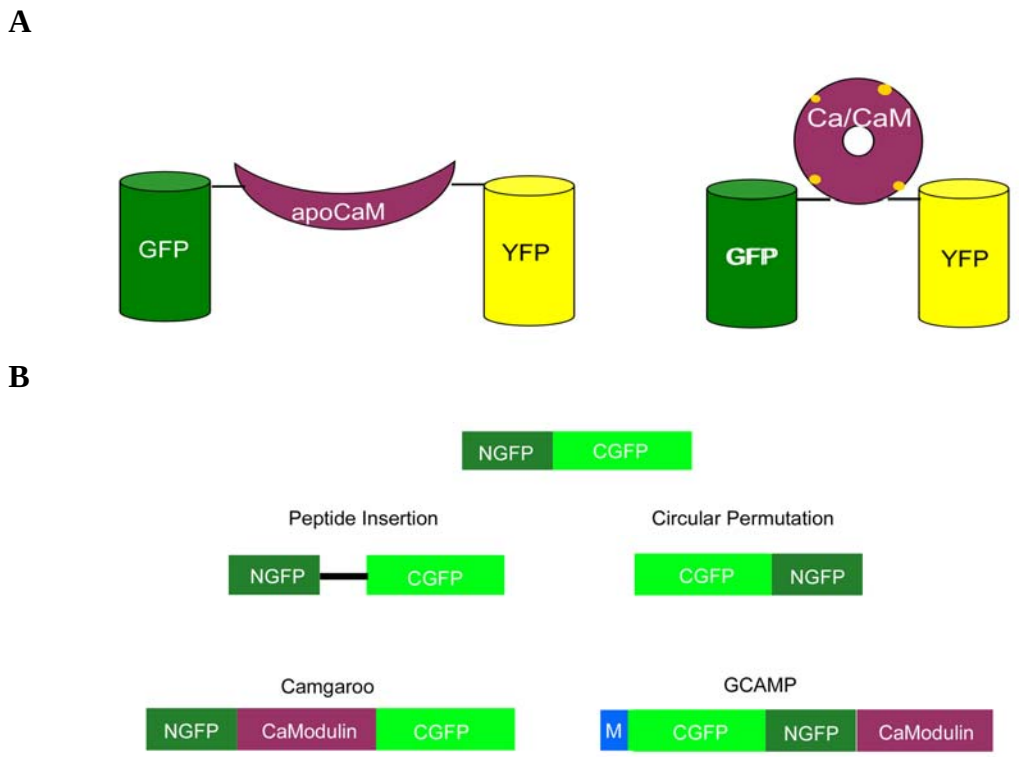


Figure 1.11. Genetically encoded calcium sensors. (A) FRET-based ,Type I GECI. FRET donor (GFP) and acceptor (RFP) are connected via Calmodulin. The Calcium (orange dot) introduces a compaction of calmodulin and increases the FRET signal. (B) Single-FP based, Type II GECI. Two successful strategies, using peptide insertion or circular permutation, are illustrated.

Genetically encoded calcium indicators

Type1 GECIs

Type I GECIs were designed based on Forster resonance energy transfer (FRET) that can correlate the calcium-induced conformational change in calmodulin with a change in fluorescent signal. Two GFP-like fluorescent proteins that serve as FRET donor and acceptor are flanking a calmodulin within one polypeptide chain. Upon calcium binding, the EF hand of calmodulin changes its conformation, thus bringing together the donor and acceptor close enough for an increase in FRET signal (Figure 1.11.A). The effective distance for a detectable FRET signal is around 100 Å (Jares-Erijman and Jovin, 2003). Cameleons were the first sensors that utilize this

design principle (Miyawaki et al., 1997). The further optimization of this sensor includes yellow camelons (Miyawaki et al., 1999; Miyawaki and Tsien, 2000; Griesbeck et al., 2001) with enhanced pH stability and red cameleons with improved FRET efficiency (Mizuno et al., 2001).

Type II GECIs

Some GFP variants are very sensitive to local changes in their structure. In these cases, the fluorescence can be altered through protein-protein interaction involving the beta-barrel of GFP. Type II GECIs utilize this property and fuse one fluorescent protein module with other protein domains that can change conformation upon calcium binding (Figure 1.11.B). The most widely used Type II GECIs are Camgaroos and GCaMPs. In Camgaroos, calmodulin is flanked by the half-sites of split-EYFP (Baird et al., 1999). Although the mechanism of Camgaroo remains elusive, it is general considered that the binding of calcium brings the two half-sites of EYFP close enough to form an intact β-barrel that emits high fluorescence. Improved Camgaroos are particularly useful to detect calcium signals in mitochondria and the ER (Griesbeck et al., 2001). GCaMPs fuse Calmodulin to the C-terminus of a circular permutated (cp) EGFP, and an M13 helix to the N-terminus so that the interaction between Calmodulin and M13 upon calcium binding may have an impact on cpEGFP's fluorescence signal (Nakai et al., 2001; Nagai et al., 2001). Optimized GCaMPs have unexpectedly high brightness and signal strength, great thermal stability, rapid kinetics and favorable *K*D for calcium (Ohkura et al., 2005; Tallini et al. 2006a). These features provide GCaMPs great advantages in imaging physiological events in intact cells (Ji et al., 2004; Tallini et al., 2006b). The most up-to-date sensor, GCaMP3, can be used to directly detect neuronal activity when expressed in the motor cortex of the mouse (Tian et al., 2009).

Overview

Despite the significant progress in the development of GCaMPs, the challenges are clear. The reported GCaMPs undergo activation transitions at a rate in the order of 10-100 ms. Such transition rate is not fast enough to monitor high frequency, physiological processes such as calcium flux in the myocardium upon contraction or in neurons. Furthermore, the gap in our understanding of the molecular mechanism employed by these sensors hampers the optimization and development of other GCaMP-like sensors. The goal of this thesis is to understand various working mechanism of GECIs and apply the knowledge to develop next generation genetically encoded calcium sensors.

In chapter five, I introduce the biophysical studies on the GCaMP2. A major contribution of this chapter is the high resolution crystal structure of calcium-bound GCAMP2 and the low-resolution solution conformation of apo-GCaMP2. The structural characterization serves as basis for our mechanistic study. The combination of the two structural models unveiled the thus far enigmatic molecular mechanism of GCAMP illumination, and thus provided a first clue how this sensor works. Based on the structural studies, several strategies that may improve its performance are discussed. As a next step and great challenge, we attempted to design a red fluorescent GCaMP-like sensor based on the knowledge we gain from our current studies.

In chapter six, I will introduce the biophysical studies on GmKate, a fluorescent protein that has dual green and red emission. At physiological pH, GmKate absorbs at 455 nm and emit green fluorescence (525 nm). In the high pH conditions, Gmkate instead absorbs at 600 nm and emit far-red fluorescence (648 nm) Based on the high resolution crystal structures of GmKate in green state and red state, I proposed a proton-mediated color transition mechanism. In the current color

transition model, the green emission is from a protonated, cis-chromophore and the far-red emission is from a deprotonated cis-chromophore. This work is significant in that it identified a novel color transition mechanism that shed light on the next generation GECI design. A side product of this chapter is a novel pH sensor with fast, reversible transition kenetics.

Chapter seven is the last chapter of this thesis. It briefly summaries the major conclusions of this thesis and discusses the potential future directions for both research projects mentioned above. Several important problems which have not been addressed in this thesis will also be introduced.

REFERENCES

Adam V., Carpentier P., Violot S., Lelimousin M., Darnault C., Nienhaus G.U., Bourgeois D. (2009) Structural basis of X-ray-induced transient photobleaching in a photoactivatable green fluorescent protein. J. Am. Chem. Soc. 131(50):18063-5.

Alto N.M., et al., (2007) The type III effector EspF coordinates membrane trafficking by the spatiotemporal activation of two eukaryotic signaling pathways. J. Cell Biol. 178(7):1265-78.

Amat P., Granucci G., Buda F., Persico M., Tozzini V. (2006) The chromophore of asFP595: a theoretical study. J. Phys. Chem. B. 110(18):9348-53.

Ando R., Hama H., Yamamoto-Hino M., Mizuno H., Miyawaki A. (2002) An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein Proc. Natl. Acad. Sci. USA 99: 12651–12656.

Andresen M., Wahl M.C., Stiel A.C., Gräter F., Schäfer L.V., Trowitzsch S., Weber G., Eggeling C., Grubmüller H., Hell S.W., Jakobs S. (2005) Structure and mechanism of the reversible photoswitch of a fluorescent protein. Proc. Natl. Acad. Sc.i USA 102 (37):13070-4.

Baird G.S., Zacharias D.A., Tsien R.Y. (2000) Biochemistry, mutagenesis, and oligomerization of dsRed, a red fluorescent protein from coral. Proc. Natl. Acad. Sci., 97: 11984-11989.

Bashkirov P.V., Akimov S.A., Evseev A.I., Schmid S.L., Zimmerberg J., Frolov V.A.(2008) GTPase cycle of dynamin is coupled to membrane squeeze and release, leading to spontaneous fission. Cell. 26;135(7):1276-86.

Bravaya K.B., Bochenkova A.V., Granovsky A.A., Savitsky A.P., Nemukhin AV. (2008) Modeling photoabsorption of the asFP595 chromophore. J. Phys. Chem. A. 112 (37):8804-10.

Brejc K., Sixma T.K., Kitts P.A., Kain S.R., Tsien R.Y., Ormö M., Remington S.J. (1997) Structural basis for dual excitation and photoisomerization of the Aequorea victoria green fluorescent protein. Proc. Natl. Acad. Sci. 94: 2306-2311.

Campelo F., Fabrikant G., McMahon H.T., Kozlov M.M. (2010) Modeling membrane shaping by proteins: focus on EHD2 and N-BAR domains. FEBS Lett. 584(9):1830-9.

Campelo F., McMahon H.T., Kozlov M.M.(2008) The hydrophobic insertion mechanism of membrane curvature generation by proteins. Biophys. J. 95(5):2325-39.

Carr J.F., Hinshaw J.E.(1997) Dynamin assembles into spirals under physiological salt conditions upon the addition of GDP and gamma-phosphate analogues. J. Biol. Chem. 272(44):28030-5.

Chen Y.J., Zhang P., Egelman E.H., Hinshaw J.E. (2004) The stalk region of dynamin drives the constriction of dynamin tubes. Nat. Struct. Mol. Biol. 11(6):574-5.

Chernomordik L.V., Kozlov M.M. (2008) Mechanics of membrane fusion. Nat. Struct. Mol. Biol. 5(7):675-83.

Cremona O., De Camilli P..(2001) Phosphoinositides in membrane traffic at the synapse. J. Cell. Sci. 114(Pt 6):1041-52.

Cubitt A.B., Heim R., Adams S.R., Boyd A.E., Gross L.A., Tsien R.Y. (1995) Understanding, improving and using green fluorescent proteins. TIBS 20, 448−455.

Danino D., Moon K.H., Hinshaw J.E. (2004) Rapid constriction of lipid bilayers by the mechanochemical enzyme dynamin. J. Struct. Biol. 147(3):259-67.

Danino D., Hinshaw J.E. (2001) Dynamin family of mechanoenzymes. Curr. Opin. Cell. Biol. 13(4):454-60.

De Camilli P., Emr S.D., McPherson P.S., Novick P. (1996) Phosphoinositides as regulators in membrane traffic. Science. 271(5255):1533-9.

Deuling H.J., Helfrich W. (1976) The curvature elasticity of fluid membranes: A catalog of vesicle shapes. J. Phys.Chem. 121(4):12-16

Di Paolo G., De Camilli P. (2006) Phosphoinositides in cell regulation and membrane dynamics. Nature. 443(7112):651-7.

Doherty G.J., McMahon H.T. (2009) Mechanisms of endocytosis. Annu. Rev. Biochem. 78:857-902.

Donnelly M., Fedeles F., Wirstam M., Siegbahn P.E.M., Zimmer M. (2001) J. Am. Chem. Soc. 123, 4679−4686.

Eberth A., Lundmark R., Gremer L., Dvorsky R., Koessmeier K.T., McMahon H.T., Ahmadian M.R. (2009) A clathrin-independent tubulovesicular endocytic pathway dependent on the BAR-domain protein GRAF1, A BAR domain-mediated autoinhibitory mechanism for RhoGAPs of the GRAF family. Biochem. J. 417, 371-377.

Farsad K., Ringstad N., Takei K., Floyd S.R., Rose K. and De Camilli P. (2001) Generation of high curvature membranes mediated by direct endophilin bilayer interactions. J. Cell Biol. 155: 193–200

Fotin A., Cheng Y., Sliz P., Grigorieff N., Harrison S.C., Kirchhausen T., Walz T. (2004) Molecular model for a complete clathrin lattice from electron cryomicroscopy. Nature. 432(7017):573-9.

Frost A., De Camilli P., Unger V.M. (2007) F-BAR proteins join the BAR family fold. Structure. 15(7):751-3.

Frost A, Unger V.M., De Camilli P. (2009). The BAR domain superfamily: membrane-molding macromolecules. Cell. 137(2):191-6.

.

Frost A., Perera R., Roux A., Spasov K., Destaing O., Egelman E.H., De Camilli P., Unger V.M. (2008). Structural basis of membrane invagination by F-BAR domains. Cell. 2008 132(5):807-17.

Fruman D.A., Meyers R.E., Cantley L.C. (1998) Phosphoinositide kinases. Annu. Rev. Biochem. 67, 481–507

Gallusser A., Kirchhausen T. (1993) The beta 1 and beta 2 subunits of the AP complexes are the clathrin coat assembly components. EMBO J. 12:5237–44

Giepmans B.N.G., Adams S.R., Ellisman M.H., Tsien R.Y. (2006) The fluorescent toolbox for assessing protein location and function. Science. 312: 217-224.

Grigorenko B., Savitsky A., Topol I., Burt S., Nemukhin A. (2006) Ground-state structures and vertical excitations for the kindling fluorescent protein asFP595. J. Phys. Chem. B. 110(37):18635-40.

Gross L.A., Baird G.S., Hoffman R.C., Baldridge K.K., Tsien, R.Y. (2000) The structure of the chromophore within DsRed, a red fluorescent protein from coral. Proc. Natl. Acad. Sci., 97: 11990-11995.

Guerrier S., Coutinho-Budd J., Sassa T., Gresset A., Jordan N.V., Chen K., Jin W.L., Frost A., Polleux F. (2009) The F-BAR domain of srGAP2 induces membrane protrusions required for neuronal migration and morphogenesis. Cell. 138(5):990-1004.

Gurskaya N.G., Verkhusha V.V., Shcheglov A.S., Staroverov D.B., Chepurnykh T.V., Fradkov A.F., Lukyanov S., Lukyanov K.A. (2006) Engineering of a monomeric green-to-red photoactivatable fluorescent protein induced by blue light. Nat. Biotechnol. 24(4):461-5.

Head J.F., Inouye S., Teranishi K., Shimomura O.(2000) The crystal structure of the photoprotein aequorin at 2.3 A resolution. Nature. 405(6784):372-6.

Heikal A.A., Hess S.T., Baird G.S., Tsien R.Y., Webb W.W. (2000) Molecular Spectroscopy and Dynamics of Intrinsically Fluorescent Proteins: Coral Red (DsRed) and Yellow (Citrine). Proc. Natl. Acad. Sci., 97: 11996-12001

Heim R., Prasher D.C., Tsien R.Y. (1994) Wavelength mutations and post-translational autooxidation of green fluorescent protein. Proc. Natl. Acad. Sci. U.S.A. 91: 12501−12504.

Henne W.M., Boucrot E., Meinecke M., Evergren E., Vallis Y., Mittal R., McMahon HT. (2010) FCHo proteins are nucleators of clathrin-mediated endocytosis. Science. 328(5983):1281-4.

Hill E., Van der Kaay J., Downes C.P., Smythe E. (2001) The role of dynamin and its binding partners in coated pit invagination and scission. J. Cell Biol. 152: 309–323.

Hinshaw J.E., Schmid S.L. (1995) Dynamin self-assembles into rings suggesting a mechanism for coated vesicle budding. Nature. 374(6518):190-2.

Hinshaw J.E. (1999) Dynamin spirals. Curr. Opin. Struct. Biol. 9(2):260-7.

Hinshaw J.E. (2000) Dynamin and its role in membrane fission. Annu. Rev. Cell Dev. Biol. 16:483-519.

Hu J., Shibata Y., Zhu P.P., Voss C., Rismanchi N., Prinz W.A., Rapoport T.A., Blackstone C. (2009) A class of dynamin-like GTPases involved in the generation of the tubular ER network. Cell. 138(3):549-61.

Hu J., Shibata Y., Voss C., Shemesh T., Li Z., Coughlin M., Kozlov M.M., Rapoport T.A., Prinz W.A. (2008) Membrane proteins of the endoplasmic reticulum induce high-curvature tubules. Science. 319(5867):1247-50.

Itoh T., Erdmann K.S., Roux A., Habermann B., Werner H., De Camilli P. (2006) Dynamin and the actin cytoskeleton cooperatively regulate plasma membrane invagination by BAR and F-BAR proteins. Dev. Cell. 9(6):791-804.

Itoh T., Hasegawa J., Tsujita K., Kanaho Y., Takenawa T. (2009) The tyrosine kinase Fer is a downstream target of the PLD-PA pathway that regulates cell migration. Sci. Signal. 8;2(87):ra52.

Jao C.C., Hegde B.G., Gallop J.L., Hegde P.B., McMahon H.T., Haworth I.S., Langen R. (2010) Roles of amphipathic helices and the BAR domain of endophilin in membrane curvature generation. J. Biol. Chem. 2010 Apr 23. Paper in press

Jares-Erijman E.A., Jovin T.M. (2008) FRET imaging. Nat. Biotechnol. 21: 1387-95

Jia J.Y., Lamer S., Schumann M., Schmidt M.R., Krause E., Haucke V. (1996) Quantitative Proteomics Analysis of Detergent-resistant Membranes from Chemical Synapses: Evidence for Cholesterol as Spatial Organizer of Synaptic Vesicle Cycling Mol. Cell. Proteomics 5(11): 2060 - 2071.

Kirchhausen T. (1993) Coated pits and coated vesicles – sorting it all out. Curr. Opin. Struct. Biol. 3:182–88

Kirchhausen T. (2002) Clathrin adaptors really adapt. Cell. 109(4):413-6

Kozlov M.M. (2010) Biophysics: Joint effort bends membrane. Nature. 463 (7280):439-40.

Lemmon M.A. (2007) Pleckstrin homology (PH) domains and phosphoinositides. Biochem Soc Symp. (74):81-93.

Lingwood D., Simons K. (2010). Lipid rafts as a membrane-organizing principle. Science. 327(5961):46-50.

Loerke D., Mettlen M., Yarar D., Jaqaman K., Jaqaman H., Danuser G., Schmid S.L. (2009) Cargo and dynamin regulate clathrin-coated pit maturation.PLoS Biol. 17; 7(3):e57.

Lundmark R., Doherty G.J., Howes M.T., Cortese K., Vallis Y., Parton R.G. and McMahon H.T. (2008) The GTPase-activating protein GRAF1 regulates the

CLIC/GEEC endocytic pathway. Curr. Biol. 18, 1802-1808.

Marks B., Stowell M.H., Vallis Y., Mills I.G., Gibson A., Hopkins C.R., McMahon H.T. (2001). GTPase activity of dynamin and resulting conformation change are essential for endocytosis. Nature. 410(6825):231-5.

Mattila P.K., Pykäläinen A., Saarikangas J., Paavilainen V.O., Vihinen H., Jokitalo E., Lappalainen P. (2007) Missing-in-metastasis and IRSp53 deform PI(4,5)P2-rich membranes by an inverse BAR domain-like mechanism. J. Cell Biol. 176(7):953-64.

McMahon H.T., Kozlov M.M., Martens S. (2010) Membrane curvature in synaptic vesicle fusion and beyond. Cell. 140(5):601-5.

Mettlen M., Pucadyil T., Ramachandran R., Schmid S.L. (2009) Dissecting dynamin's role in clathrin-mediated endocytosis. Biochem. Soc. Trans. 37(Pt 5):1022-6.

Merzlyak E.M. et al., Bright far-red fluorescent protein for whole-body imaging. 2007 Nat. Methods. 4(9):741-6.

Miyawaki A., Llopis J., Heim R., McCaffery J.M., Adams J.A., Ikurak M., Tsien R.Y. (1997). Fluorescent indicators for Ca2+ based on green fluorescent proteins and Calmodulin. Nature. 388 (6645): 882–7

Miyawaki A., Tsien R.Y. (2000) Monitoring protein conformations and interactions by fluorescence resonance energy transfer between mutants of green fluorescent protein. Methods Enzymol. 327: 472-500

Mizuno N., Jao C.C., Langen R., Steven A.C. (2010) Multiple modes of endophilin-mediated conversion of lipid vesicles into coated tubes: implications for synaptic endocytosis. J. Biol. Chem. 2010 May 18. Paper in press.

Modregger J., Ritter B., Witter B., Paulsson M., Plomann M. (2000) All three PACSIN isoforms bind to endocytic proteins and inhibit endocytosis. J. Cell. Sci. 113:4511-4521.

Modregger J., DiProspero N.A., Charles V., Tagle D.A., Plomann M. (2002) PACSIN 1 interacts with huntingtin and is absent from synaptic varicosities in presymptomatic Huntington's disease brains. Hum. Mol. Genet. 11:2547-2558.

Nienhaus G.U. et al. (2006) Photoconvertible fluorescent protein EosFP: biophysical properties and cell biology applications. Photochem. Photobiol. 2006 Mar-Apr;82(2):351-8.

Nienhaus K., Nar H., Heilker R., Wiedenmann J., Nienhaus G.U. (2008) Trans-cis isomerization is responsible for the red-shifted fluorescence in variants of the red fluorescent protein eqFP611. J. Am. Chem. Soc. 130(38):12578-9.

Nienhaus K., Nienhaus G.U., Wiedenmann J., Nar H. (2005) Structural basis for photo-induced protein cleavage and green-to-red conversion of fluorescent protein EosFP. Proc. Natl. Acad. Sci. U. S. A. 102(26):9156-9.

Ormö M., Cubitt A.B., Kallio K., Gross L.A., Tsien R.Y., Remington S.J. (1996) Crystal structure of Aequorea victoria green fluorescent protein. Science. 273: 1392-1395.

Padrick S.B., Cheng H.C., Ismail A.M., Panchal S.C., Doolittle L.K., Kim S., Skehan B.M., Umetani J., Brautigam C.A., Leong J.M., Rosen M.K. (2008) Hierarchical regulation of WASP/WAVE proteins. Mol. Cell. 32(3):426-438

Padrick S.B., Rosen M.K. (2010) Physical mechanisms of signal integration by WASP family proteins. Annu. Rev. Biochem. 2010;79:707-35.

Parsons S.A., Mewburn J.D., Truesdell P., Greer P.A. (2007) The Fps/Fes kinase regulates leucocyte recruitment and extravasation during inflammation. Immunology 122(4):542-50.

Peter B.J., Kent H.M., Mills I.G., Vallis Y., Butler P.J., Evans P.R., McMahon H.T. (2004) BAR domains as sensors of membrane curvature: the amphiphysin BAR structure. Science. 303(5657):495-9.

Payne G.S. (1990). Genetic analysis of clathrin function in yeast. J. Membr. Biol. 116:93–105

Pletnev S., Shcherbo D., Chudakov D.M., Pletneva N., Merzlyak E.M., Wlodawer A., Dauter Z., Pletnev V. (2008) A crystallographic study of bright far-red fluorescent protein mKate reveals pH-induced cis-trans isomerization of the chromophore. J. Biol. Chem. 24; 283(43):28980-7

Pucadyil T.J., Schmid S.L. (2009) Conserved functions of membrane active GTPases in coated vesicle formation. Science. 325(5945):1217-20.

Quillin M.L., Anstrom D.M., Shu X., O'Leary S., Kallio K., Chudakov D.M., Remington S.J. (2005) Kindling fluorescent protein from Anemonia sulcata: dark-state structure at 1.38 A resolution. Biochemistry. 44(15):5774-87.

Qualmann B., Kelly R.B. (2000) Syndapin isoforms participate in receptor-mediated endocytosis and actin organization. J. Cell. Biol. 148(5):1047-1062.

Qualmann B., Roos J., DiGregorio P.J., Kelly R.B. (1999) Syndapin I, a synaptic dynamin-binding protein that associates with the neural Wiskott-Aldrich syndrome protein. Mol. Biol. Cell. 10:501-513.

Ramachandran R., Pucadyil T.J., Liu Y.W., Acharya S., Leonard M., Lukiyanchuk V., Schmid S.L. (2009) Membrane insertion of the pleckstrin homology domain variable loop 1 is critical for dynamin-catalyzed vesicle scission. Mol. Biol. Cell. 20(22):4630-9.

Ringstad N., Gad H., Löw P., Di Paolo G., Brodin L., Shupliakov O., De Camilli P. (1999) Endophilin/SH3p4 is required for the transition from early to late stages in clathrin-mediated synaptic vesicle endocytosis. Neuron 24(1):143-54

Ringstad N., Nemoto Y., De Camilli P. (1997) The SH3p4/Sh3p8/SH3p13 protein family: binding partners for synaptojanin and dynamin via a Grb2-like Src homology 3 domain. Proc. Natl. Acad. Sci. USA 94: 8569–8574

Ringstad N., Nemoto Y., De Camilli P. (2001), Differential expression of endophilin-1 and endophilin-2 dimers at central nervous system synapses. J. Biol. Chem. 276: 40424–40430

Roux A., Uyhazi K., Frost A., De Camilli P. (2006) GTP-dependent twisting of dynamin implicates constriction and tension in membrane fission, Nature 441: 528–531.

Schäfer L.V., Groenhof G., Boggio-Pasqua M., Robb M.A., Grubmüller H. (2008) Chromophore protonation state controls photoswitching of the fluoroprotein asFP595. PLoS Comput. Biol. 4(3):e1000034.

Schäfer L.V, Groenhof G., Klingen A.R., Ullmann G.M., Boggio-Pasqua M., Robb M.A., Grubmüller H.(2007) Photoswitching of the fluorescent protein asFP595: mechanism, proton pathways, and absorption spectra. Angew. Chem. Int. Ed. Engl. 46(4):530-6.

Schmid S.L. (1997) Clathrin-coated vesicle formation and protein sorting: an integrated process. Annu. Rev. Biochem. 66:511-48.

Schmid S.L., McNiven M.A., De Camilli P. (1998) Dynamin and its partners: a progress report. Curr. Opin. Cell. Biol. 10(4):504-12.

Schmidt A., Wolde M., Thiele C., Fest W., Kratzin H., Podtelejnikov A.V., Witke W., Huttner W.B., Söling H.D. (1999) , Endophilin-1 mediates synaptic vesicle formation by transfer of arachidonate to lysophosphatidic acid. Nature 401:133–141

Schuske K.R., Richmond J.E., Matthies D.S., Davis W.S., Runz S., Rube D.A., van der Bliek A.M., Jorgensen E.M. (2003) Endophilin is required for synaptic vesicle endocytosis by localizing synaptojanin. Neuron 40: 749–762

Shaner N.C., Campbell R.E., Steinbach P.A., Giepmans B.N., Palmer A.E., Tsien R.Y. (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nat. Biotechnol. 22, 1567–1572

Shaner N.C., Steinbach P.A., Tsien R.Y. (2005) A guide to choosing fluorescent proteins, Nature Methods 2: 905 – 909

Shapovalova Z., Tabunshchyk K., Greer PA. (2007) The Fer tyrosine kinase regulates an axon retraction response to Semaphorin 3A in dorsal root ganglion neurons.BMC Dev. Biol. 30;7:133.

Shcherbo D. et al. (2009) Far-red fluorescent tags for protein imaging in living tissues. Biochem. J. 418(3):567-74.

Shibata Y., Hu J., Kozlov M.M., Rapoport T.A. (2009) Mechanisms shaping the membranes of cellular organelles. Annu. Rev. Cell. Dev. Biol. 2009;25:329-54.

Shih W., Gallusser A., Kirchhausen T. (1995) A clathrin-binding site in the hinge of the beta 2 chain of mammalian AP-2 complexes. J. Biol. Chem. 270:31083–90

Shimada A., Takano K., Shirouzu M., Hanawa-Suetsugu K., Terada T., Toyooka K., Umehara T., Yamamoto M., Yokoyama S., Suetsugu S. (2010) Mapping of the basic amino-acid residues responsible for tubulation and cellular protrusion by the EFC/F-BAR domain of pacsin2/Syndapin II. FEBS Lett. 19;584(6):1111-8.

Shimomura O. (1979) Structure of the chromophore of Aequorea green fluorescent protein. FEBS Lett.1979, 104, 220−222

Shimomura O. (1995). "A short story of aequorin.". Biol. Bull. 189 (1): 1–5.

Simpson F., Hussain N.K., Qualmann B., Kelly R.B., Kay B.K., McPherson P.S., Schmid S.L. (1999) SH3-domain-containing proteins function at distinct steps in clathrin-coated vesicle formation. Nat. Cell. Biol. 1: 119–124

Smith J.A., Samayawardhena L.A., Craig A.W. (2010) Fps/Fes protein-tyrosine kinase regulates mast cell adhesion and migration downstream of Kit and beta1 integrin receptors. Cell. Signal. 22(3):427-36.

Soulet F., Yarar D., Leonard M., Schmid S.L. (2005) SNX9 regulates dynamin assembly and is required for efficient clathrin-mediated endocytosis. Mol. Biol. Cell. 16(4):2058-67.

Stepp J.D., Pellicena-Palle A., Hamilton S., Kirchhausen T., Lemmon S.K. (1995) A late Golgi sorting function for Saccharomyces cerevisiae Apm1p, but not for Apm2p, a second yeast clathrin AP medium chain-related protein. Mol. Biol. Cell 6:41–58

Stiel A.C., Trowitzsch S., Weber G., Andresen M., Eggeling C., Hell S.W., Jakobs S, Wahl M.C. (2007) 1.8 Å bright-state structure of the reversibly switchable fluorescent protein Dronpa guides the generation of fast switching variants. Biochem. J. 402(1):35-42.

Stowell M.H., Marks B., Wigge P., McMahon H.T. (1999) Nucleotide-dependent conformational changes in dynamin: evidence for a mechanochemical molecular spring. Nat. Cell Biol. 1:27–32

Subtil A., Gaidarov I., Kobylarz K., Lampson M.A., Keen J.H., McGraw T.E. (1999) Acute cholesterol depletion inhibits clathrin-coated pit budding. Proc. Natl. Acad. Sci. USA. 8;96(12):6775-80.

Suresh S., Edwardson J.M. (2010) The endophilin N-BAR domain perturbs the structure of lipid bilayers. Biochemistry. 2010 Jun 8. Paper in press

Sweitzer SM, Hinshaw JE.(1998) Dynamin undergoes a GTP-dependent conformational change causing vesiculation. Cell. 93(6):1021-9.

Takei K., McPherson P.S., Schmid S.L., De Camilli P. (1995) Tubular membrane invaginations coated by dynamin rings are induced by GTPγS in nerve terminals. Nature 374:186–190.

Takei K., Slepnev V.I., Haucke V., De Camilli P.(1999) Functional partnership between amphiphysin and dynamin in clathrin-mediated endocytosis.Nat. Cell. Biol. 1(1):33-9.
Tsien R.Y. (1998) Annu. Rev. Biochem.1998, 67, 510−544

van Meer G., Voelker D.R., Feigenson G.W. (2008) Membrane lipids: where they are and how they behave. Nat. Rev. Mol. Cell Biol. 9(2):112-124.

Volpicelli-Daley L., De Camilli P. (2007) Phosphoinositides' link to neurodegeneration. Nat. Med. 13: 784-786

Wang L., Jackson W.C., Steinbach P.A., Tsien R.Y. (2004) Evolution of new nonantibody proteins via iterative somatic hypermutation. Proc. Natl. Acad. Sci. USA 101, 16745-16749 (2004).

Warnock D.E., Hinshaw J.E., Schmid S.L. (1996) Dynamin self-assembly stimulates its GTPase activity. J. Biol. Chem. 271(37):22310-4.

Wenk M.R., De Camilli P. (2004) Protein–lipid interactions and phosphoinositide metabolism in membrane traffic: insights from vesicle recycling in nerve terminals. Proc. Natl. Acad. Sci. USA 101: 8262–8269

Wiedenmann J., Ivanchenko S., Oswald F., Schmitt F., Röcker C., Salih A., Spindler K.D., Nienhaus G.U. (2004) EosFP, a fluorescent marker protein with UV-inducible green-to-red fluorescence conversion. Proc. Natl. Acad. Sci. USA. 101(45):15905-10.

Wigge P., Vallis Y., McMahon H.T. (1997) Inhibition of receptor-mediated endocytosis by the amphiphysin SH3 domain of special interest . Curr. Biol. 7: 554–560.

Yampolsky I.V., Remington S.J., Martynov V.I., Potapov V.K., Lukyanov S., Lukyanov K.A. (2005) Synthesis and properties of the chromophore of the asFP595 chromoprotein from Anemonia sulcata. Biochemistry. 44(15):5788-93.

Yang F., Moss L.G., Phillips G.N. (1996) Structure and fluorescence mechanism of GFP. Nature Biotech. 14, 1246-1251

Yarar D., Surka M.C., Leonard M.C., Schmid S.L. (2007) SNX9 activities are regulated by multiple phosphoinositides through both PX and BAR domains. Traffic 12(1):133-46..

Yarar D., Waterman-Storer C.M., Schmid S.L. (2007) SNX9 couples actin assembly to phosphoinositide signals and is required for membrane remodeling during endocytosis. Dev. Cell. 13(1):43-56.

Yarbrough D., Wachter R.M., Kallio K., Matz M.V., Remington S.J. (2001) Refined crystal structure of DsRed, a red fluorescent protein from coral, at 2.0-Å resolution. Proc. Natl. Acad. Sci. USA. 98(2):462-7.

Zhang P., Hinshaw J.E. (2010) Three-dimensional reconstruction of dynamin in the constricted state. Nat. Cell. Biol. 3(10):922-6.

Zhao J., Wu J., Heberle F.A., Mills T.T., Klawitter P., Huang G., Costanza G., Feigenson G.W. (2007). Phase studies of model biomembranes: complex behavior of DSPC/DOPC/cholesterol. Biochim. Biophys. Acta.1768(11):2764-76.

Zhao J., Wu J., Shao H., Kong F., Jain N., Hunt G., Feigenson G. (2007). Phase studies of model biomembranes: macroscopic coexistence of Lalpha+Lbeta, with light-induced coexistence of Lα+Lo Phases. Biochim. Biophys. Acta. 1768(11):2777-86.

Zimmerberg J., Kozlov M. (2006) How proteins produce cellular membrane curvature. Nat. Rev. Mol. Cell. Biol. 7:9-19.

Zimmerberg J., McLaughlin S. (2004) Membrane curvature: how BAR domains bend bilayers. Curr. Biol. 14:R250-252.

CHAPTER 2

MOLECULAR MECHANISM OF MEMBRANE TUBULATION AND RESTRICTION BY THE F-BAR PROTEIN PACSINS

ABSTRACT

Peripheral membrane proteins of the Bin/amphiphysin/Rvs (BAR) and Fer-CIP4 homology-BAR (F-BAR) family participate in cellular membrane trafficking and have been shown to generate membrane tubules. The degree of membrane bending appears to be encoded in the structure and immanent curvature of the particular protein domains, with BAR and F-BAR domains inducing high- and low-curvature tubules, respectively. In addition, oligomerization and the formation of ordered arrays influences tubule stabilization. Here, the F-BAR domain-containing protein Pacsin was found to possess a unique activity, creating small tubules and tubule constrictions, in addition to the wide tubules characteristic for this subfamily. Based on crystal structures of the F-BAR domain of Pacsin and mutagenesis studies, vesiculation could be linked to the presence of unique structural features distinguishing it from other F-BAR proteins. Tubulation was suppressed in the context of the full-length protein suggesting that Pacsin is autoinhibited in solution. The regulated deformation of membranes and promotion of tubule constrictions by Pacsin suggests a more versatile function of these proteins in vesiculation and endocytosis beyond their role as scaffold proteins.

INTRODUCTION

Pacsins (also known as Syndapins or FAP52) comprise a conserved family of peripheral membrane proteins in eukaryotes that play a central role in synaptic vesicle recycling and receptor-mediated endocytosis, ultimately controlling the cell surface expression of transmembrane proteins (Anggono et al., 2006; Modregger et al., 2000; Qualmann and Kelly, 2000; Qualmann et al., 1999). The three isoforms of Pacsin in higher eukaryotes differ in their expression pattern, cargo selectivity, and recruitment of other proteins involved in endocytosis. Pacsin 1 is a brain-specific isoform, and has been implicated in neuronal transmission and the neuropathology of Huntington's disease, whereas Pacsin 2 and 3 are more widely expressed (Modregger et al., 2000; DiProspero et al., 2004; Modregger et al., 2002). On a molecular level, Pacsins bind to Dynamin and Eps15 homology domain (EHD)-containing proteins, nucleotide-dependent enzymes with membrane remodeling and fission activities, as well as the actin cytoskeleton via N-WASP, linking membrane trafficking and cytoskeletal rearrangements (Qualmann et al., 1999; Braun et al., 2005; Kessels et al., 2004). While Pacsin isoforms bind a specific subset of proteins via the central linker region and their SH3 domain, they all share a highly conserved F-BAR domain at their N-terminus (Itoh et al., 2005).

F-BAR domains constitute a subfamily of BAR domains, protein modules that stabilize or induce membrane curvature (Itoh et al., 2005; Farsad et al., 2001; Peter et al., 2004; Takei et al., 1999; Tsujita et al., 2006). In contrast to N-BAR domains that prefer highly curved membrane tubules, F-BAR domains usually stabilize membrane structures with larger diameter (Itoh et al., 2005; Henne et al., 2007; Shimada et al., 2007). Such curvature preference can be attributed in part to the structures of the particular proteins (Peter et al., 2004; Henne et al., 2007; Shimada et al., 2007; Weissenhorn et al., 2005). Both BAR and F-BAR domains consist of a three helix

bundle and form extended crescent-shaped dimers. The concave surface of the arc presents an overall positive electrostatic potential that allows for interactions with the plasma membrane. To some degree, the intrinsic curvature of the dimer correlates with the preference for a particular membrane tubule or vesicle size (Itoh et al., 2005; Henne et al., 2007; Shimada et al., 2007). N-terminal, amphipathic helices in N-BAR domains contribute to the driving force required to deform membranes, an energetically expensive process (Gallop et al., 2006; Masuda et al., 2006; Zimmerberg and Kozlov, 2006). The amphipathic helix in F-BAR domain-containing proteins appears to be dispensable for their tubulation activity (Henne et al., 2007). Instead, F-BAR dimers can form ordered arrays on membranes, predominantly driven by electrostatic and intermolecular interactions (Shimada et al., 2007; Frost et al., 2008).

BAR and F-BAR domain-containing proteins encoded in eukaryotic genomes cluster into distinct branches, but little is known with regard to the functional relevance of such complexity (Itoh et al., 2005). Elucidating the morphogenic potential and molecular mechanism of these proteins will provide insight into their role in membrane trafficking and endocytosis. Here we show that Pacsin proteins stabilize tubules of both large and narrow diameter but also facilitate membrane constriction *in vitro*, suggesting unanticipated diversity in the family of F-BAR domain-containing proteins. Structural studies elucidate the molecular mechanism and unique motifs in Pacsin that contribute to its distinct membrane remodeling potential.

MATERIALS AND METHODS

Protein expression and purification

The coding regions corresponding to full-length human Pacsin 1 and 2, Pacsin $1^{F\text{-}BAR}$ (residues 1-325), Pacsin $1^{\Delta SH3}$ (residues 1-381) and Pacsin $2^{F\text{-}BAR}$ (residues 1-324) were amplified by standard PCR and cloned into a modified pET28a expression

plasmid (Novagen) yielding N-terminally hexahistidine-tagged SUMO fusion proteins. The hexahistidine-tagged SUMO-moiety was cleavable using the protease Ulp-1 from *S. cerevisiae*.

Native and selenomethionine-derivatized proteins were overexpressed in *Escherichia coli* cells BL21 (DE3) (Novagen) or in T7 Crystal Express (NEB), respectively. For the expression of native proteins, cells were grown in Terrific Broth (TB) media supplemented with 50 mg/l kanamycin at 37°C. Selenomethionine-derivatized proteins were produced in cells grown in M9 minimal media supplemented with 50 mg/l kanamycin and 40 mg/l selenomethionine. At a cell density corresponding to an absorbance of 0.8 at 600 nm, the temperature was reduced to 18°C, and protein production was induced with 1 mM IPTG. After 16 hours, cells were harvested by centrifugation, resuspended in NiNTA buffer A (25 mM Tris-HCl, pH 8.2, 500 mM NaCl and 20 mM imidazole), and flash-frozen in liquid nitrogen.

After cell lysis by sonication and removal of cell debris by centrifugation, clear lysates were loaded onto a NiNTA columns (HisTrap; GE Healthcare) equilibrated in NiNTA buffer A. The resin was washed with 20 column volumes of NiNTA buffer A, and proteins were eluted on a single step of NiNTA buffer A supplemented with 500 mM imidazole. Proteins were buffer exchanged into desalting buffer (25 mM Tris-HCl, pH 7.4, 400 mM NaCl), and affinity tags were removed by incubation with the protease Ulp-1 at 4°C overnight. Cleaved proteins were collected in the flow-through during NiNTA affinity chromatography, and were subjected to size exclusion chromatography on a Superdex 200 column (GE Healthcare) equilibrated in gel filtration buffer (25 mM Tris-HCl, pH 7.4, 400 mM NaCl). Proteins were concentrated on a Centricon ultrafiltration device (10 kDa cutoff; Millipore) to a final concentration of ~0.5-1 mM (~25-50 mg/ml). Protein aliquots were frozen in liquid nitrogen and stored at -80°C.

Crystallization of Pacsin$^{F\text{-}BAR}$

Crystals were obtained by hanging drop vapor diffusion by mixing equal volumes of protein (10 mg/ml) and reservoir solution followed by incubation at 20°C. Initial crystals were obtained by using full-length Pacsin 1 and 2, that partially proteolysed during crystallization. For Pacsin 1$^{F\text{-}BAR}$, the reservoir solution contained 0.1 M Hepes, pH 7.0, 18% PEG3350 and 1% Tryptone. Diamond-shaped crystals appeared within 3 days with typical dimensions of 0.20 mm x 0.15 mm x 0.05 mm. For cryo-protection, crystals were soaked in reservoir solution supplemented with 18% glycerol. For Pacsin 2$^{F\text{-}BAR}$, the reservoir solution contained 20% PEG-MME 5000 and 0.1 M Bis-Tris, pH 6.5. Crystals appeared as thin plates after one week with typical dimensions of 0.15 mm x 0.15 mm x 0.02 mm. Crystals were cryo-protected by soaking in the crystallization solutions supplemented with 20% glycerol. Crystals grown with the isolated F-BAR domain of Pacsin 1 appeared after 2 days by using a reservoir solution containing 0.2 M potassium sodium tartrate tetrahydrate and 20% PEG3350. Cryo-preserved crystals were flash-frozen and stored in liquid nitrogen. Data was collected on frozen crystals at 100 K.

Data reduction was carried out with the software package XDS (Kabsch et al., 1993). Experimental phases for Pacsin 1$^{F\text{-}BAR}$ crystals were obtained from Single Anomalous Diffraction (SAD) experiments on crystals grown from selenomethionine-derivatived proteins. By using the software package SHELX (Sheldrick et al., 2008), 40 (out of 44) heavy atom positions could be determined, and were further refined by using the program MLPHARE. Solvent flattening was carried out by using the program RESOLVE (12). The structure of Pacsin 2$^{F\text{-}BAR}$ was determined by Molecular Replacement by using the software package PHENIX (Adams et al., 2002) with the F-BAR domain of Pacsin 1 as the search model. Refinement in PHENIX (Adams et al., 2002) and COOT (Emsley and Cowtan, 2004) yielded the final models. Data

collection and refinement statistics are summarized in Table 2.1. Illustrations were made in Pymol (DeLano Scientific).

Analytical ultracentrifugation

Sedimentation equilibrium experiments were performed using a XL-I analytical ultracentrifuge (Beckman Coulter) equipped with an AN-60 Ti rotor. Experiments were done in the 12 mm optical path. Proteins (5-20 μM) were diluted in assay buffer (25 mM Tris-HCl, pH 7.0, 100 mM NaCl, and 1 mM TCEP) and were analyzed at a centrifugation speed of 40,000 rpm. Data collection was carried out at 280 nm, followed by data analysis using the program SedFit (version 11.0).

Small angle X-ray scattering (SAXS) and SAXS-based shape reconstruction

SAXS experiments were carried out at the Advanced Photon Source (APS, BESSRC-CAT, beamline 12-IDC, Argonne, IL) at the electron energy of 12 KeV and 4°C. Prior to data collection, samples were centrifuged for 10 min at 13,200 xg. Scattering data were collected in triplicates at protein concentrations between 0.5-4 mg/ml. Exposure times (usually 2 seconds) yielding good signal-to-noise but minimal radiation damage was determined experimentally. Background scattering was collected from low salt gel filtration buffer (25 mM Tris-HCl, pH 7.4, 150 mM NaCl), and scattering data was background-corrected, averaged, scaled, and analyzed using the programs IGOR PRO (version 5.04B) , GNOM, and CRYSOL (Svergun 1992, Svergun et al.,1995). Only data with $S_{max}*R_g<1.3$, computed from Guinier plots at low angle regions, were considered for further analysis. Kratky plots, experimental molecular weight determinations, and the normalized scattering intensity were used to assess the folded-state of the proteins and overall data quality. Distance distribution functions were determined using the program GNOM (Svergun et al.,1995).

Ab initio free atom modeling was performed by using the program GASBOR22 with the distance distribution function as the fitting target (Svergun et al.,2001). Proteins were found to be dimeric under the experimental conditions as confirmed by sedimentation equilibrium analytical ultracentrifugation and multi-angle light scattering (data not shown), and P2 symmetry was applied during modeling. Forty independent simulations were carried out for each protein. Solutions were superimposed, averaged, and filtered using the program DAMAVER (Volkov and Svergun, 2003). Data statistics are summarized in Table 2.2.

Liposome preparation

Folch fraction 1 lipids (Sigma Aldrich) were dissolved in chloroform and stored at -20°C. For liposome preparations, chloroform was evaporated under a nitrogen atmosphere for 15 min, and the samples were subjected to vacuum desiccation for 2 hrs at 60 mtorr. Dry films were hydrated in low salt buffer (25 mM Tris-HCl, pH 7.4 and 50 mM NaCl) to a concentration of 10 mg/ml.

To prepare large multi-lamellar vesicles (LMVs), hydrated liposomes were sonicated for 15 min followed by 8 freeze-thaw cycles using liquid nitrogen. For most applications, liposomes were diluted to a concentration of 2 mg/ml in low salt buffer and were incubated in a heat block at 30°C for 1 hr before exposure to protein solutions. This method yields LMVs with an average diameter of 300-600 nm, and the occasional occurrence of larger vesicles (≥1 μm). For the preparation of large unilamellar vesicles (LUVs), hydrated liposomes were subjected to 6 freeze-thaw cycles followed by extrusion through filters with pore sizes ranging from 50 nm to 400 nm. LUV preparations were diluted to a concentration of 2 mg/ml, and used immediately after extrusion.

Liposome co-pelleting assay

Liposome samples (1 mg/ml) were incubated in the presence or absence of protein (10 μM) in 40 μl low salt buffer (corresponding to a 1:1 lipid-protein ratio) for 20 min at room temperature. Samples were centrifuged in a Optima MAX-E ultracentrifuge (Beckman) equipped with a TLA-100 rotor at 87,000 rpm at 4°C for 45 min. Pellets and supernatants were separated, pellets were resuspended in 40 μl low salt buffer, and both fractions were analyzed by SDS-PAGE and Coomassie Blue staining.

Negative Staining Electron Microscopy (EM)

Liposome samples (2 mg/ml) were incubated in the presence or absence of protein (0.2-1 mg/ml or 4-20 μM) in low salt buffer for 1 min at 25°C, if not indicated otherwise. The sample (8 μl) was applied to a Carbon-formvar-coated copper grid (EMS) and incubated for 2 min. Excess liquid was carefully removed by using a wet absorbent tissue (Kimwipes; Kimberly Clark). The grids were incubated three times with 2% filtered uranyl acetate solution for 5 sec and air-dried. Negative staining was performed at 25°C (except for samples prepared at 4°C, which were stained in the cold room). Membrane morphologies were examined on a FEI Morgagni Transmission Electron Microscope and a FEI Tecnai G2 F20 TEM/STEM with the electron energy set to 80 kV and 200 kV, respectively. Representative images were taken on an AMT camera with a direct magnification of 11kx-36kx. Distance and size measurements were taken by using the program ImageJ.

RESULTS AND DISCUSSIONS

Pacsin-mediated membrane deformation.

The F-BAR domain of human Pacsin 1 and 2 were overexpressed in *E. coli* and purified to homogeneity. Light scattering and analytical ultracentrifugation

revealed a predominantly dimeric state of the isolated domain (data not shown and Figure. 2.1.A). Using the purified proteins, we analyzed their morphogenic potential in a membrane tubulation assay using negative stain electron microscopy. Large multi-lamellar vesicles (LMVs) generated from isolated brain lipids (Folch fraction I) were incubated with the F-BAR domain of Pacsin 1. Unlike the homogeneous and distinctive tubule morphology induced by other BAR and other F-BAR proteins (Endophilin, Sorting nexin 9 and Toca-1), the F-BAR domain of Pacsin 1 generated a broad spectrum of membrane structures (Figure. 2.1 and 2.2.A).

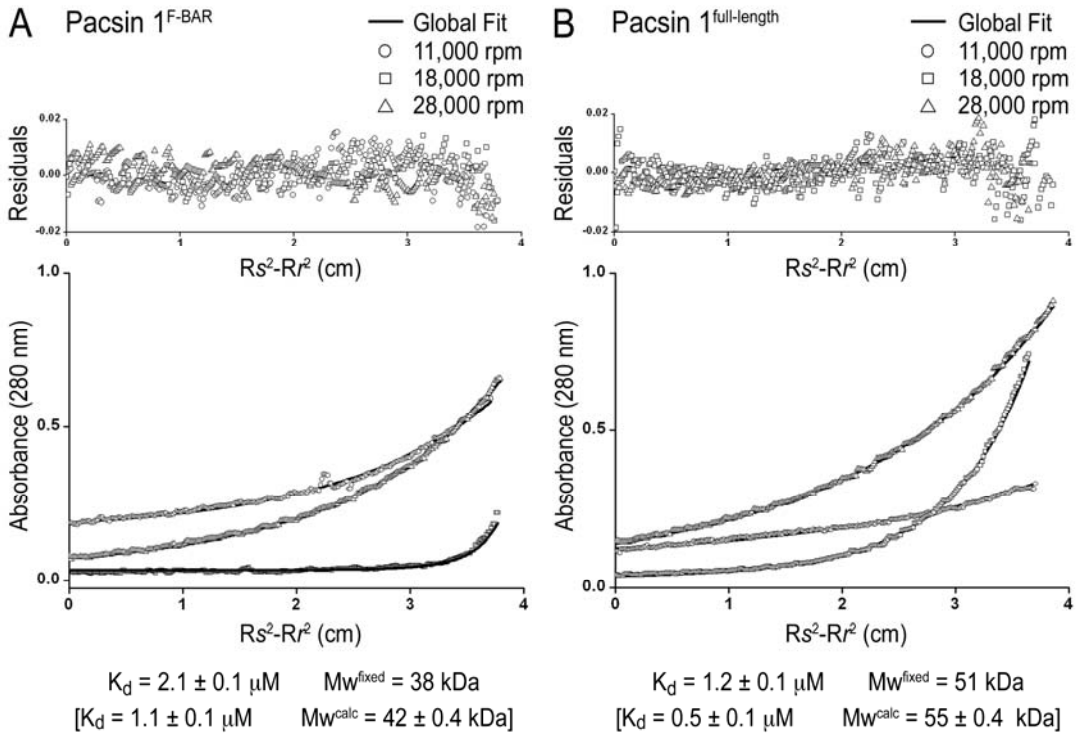


Figure 2.1. Sedimentation equilibrium analytical ultracentrifugation (AUC) of human Pacsin 1. (A) Pacsin $1^{F\text{-}BAR}$. Purified Pacsin $1^{F\text{-}BAR}$ was subjected to sedimentation equilibrium experiments. Fitting of the data to a monomer-dimer model provided the dissociation constant for F-BAR domain dimers of Pacsin 1. Molecular weight determination based on the AUC data is in good agreement with the value calculated based on the protein sequence, and the dissociation constants from calculations using the theoretical or experimentally determined molecular weight are in good agreement. (B) Full-length Pacsin 1. Pacsin $1^{full\ length}$ was analyzed as described in (A).

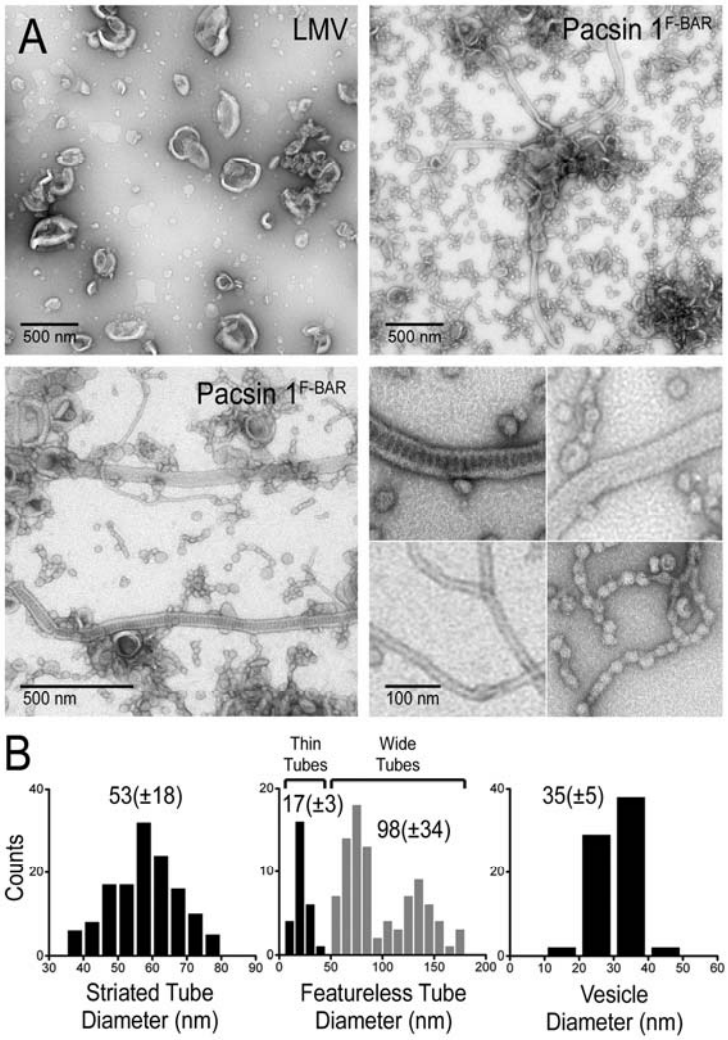


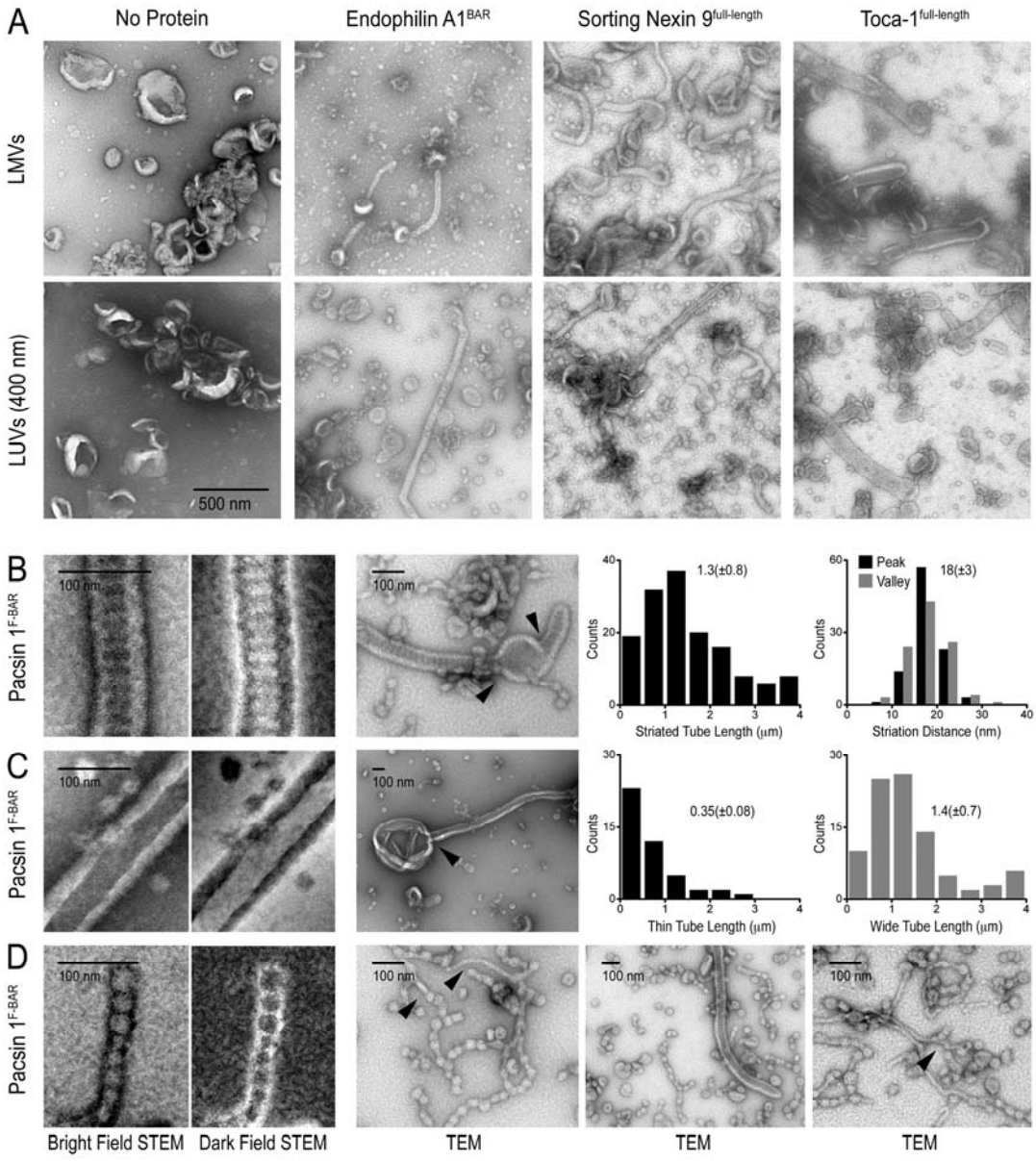
Figure 2.2. Pacsin-mediated liposome tubulation.(A) Negative stain electron microscopy (EM). Large multi-lamellar vesicles (LMVs) from Folch fraction I lipids were incubated with the isolated F-BAR domain of human Pacsin 1 (residues 1-325). The bottom-right panel shows higher magnifications of the four main membrane morphotypes induced by Pacsin.(B) Statistical analysis. Diameters of the different morphotypes shown in (A) were quantified by determining their size on at least three independently prepared EM grids. Diameters (especially for wider tubules) may appear larger due to flattening effects during negative staining.

We observed two classes of tubules with shallow curvature, striated and featureless, with a mean diameter of 53±18 nm and 98±34 nm, respectively, and average length of 1.3-1.4 μM (Figure 2.2). Given the size of LMVs, the observed

tubules were likely to originate from a single vesicle, not requiring fusion events. Occasionally, tubes were associated with vesicular structures (Figures 2.3.B and 2.3.C, arrows). Featureless tubules were comparable to structures observed with other F-BAR proteins and showed a wider distribution with regard to diameter, both across an individual tubule as well as between tubules, which in part may be attributed to flattening effects during negative staining (Figure 2.2.B, middle panel) (Itoh et al., 2005). The striated tubules induced by Pacsin had an invariable diameter and regular staining pattern suggesting that these structures were more rigid and ordered (Figure 2.2.B, left panel; Figure 2.3.B). Striation has been reported for Dynamin assemblies on membranes, as well as for the BAR domain-containing proteins endophilin and amphiphysin, but the staining pattern generated by Pacsin was distinct in that its repeat units were perpendicular to the long tube axis (Farsad et al., 2001; Takei et al., 1995; Takei et al., 1999). A similar pattern has been observed on 40 nm-tubules formed with Sorting Nexin 9 (Yarar et al., 2008), but the significance and molecular basis for such patterns awaits further elucidation.

In contrast to the morphologies observed with the majority of F-BAR domain-containing proteins, the F-BAR domain of Pacsin also induced narrow tubules with high curvature (d=17±3 nm) (Figure 2.2.A and 2.2.B, middle panel). Such structures were comparable to the ones observed with Endophilin or Amphiphysin, and slightly narrower than those obtained with the F-BAR domain of FCHo2 (Henne et al., 2007; Peter et al., 2004). With both, Pacsin and FCHo2, narrow tubules occurred at the same conditions as broader tubes. Variations in the higher-order assemblies of F-BAR domains on the membranes might contribute to the stabilization of different degrees of curvature (Henne et al., 2007; Frost et al., 2008). In addition, narrower and broader tubes may originate from liposomes with different properties such as in bilayer tension, membrane stiffness or intrinsic curvature.

Figure 2.3 Membrane tubulation by Endophilin, Sorting Nexin 9, Toca-1 and Pacsin 1.(A) Endophilin, Sorting Nexin 9 and Toca-1. Either large multi-lamellar vesicles or liposomes extruded through a filter with a pore size of 400 nm (both made from Folch fraction I lipids) were used in tubulation experiments and analyzed by negative-stain electron microscopy. The scale bar is 500 nm and applies to all images in this figure. Liposomes were incubated with the BAR domain of human Endophilin A1, human full-length Sorting Nexin 9, or *Xenopus tropicalis* full-length Toca-1, spotted onto EM grids, stained with uranyl acetate, and analyzed by electron microscopy. (B) Striated tubules induced by the F-BAR domain of human Pacsin 1. The left panel shows high-resolution images from negative stain scanning transmission electron microscopy (STEM) taken in bright and dark field mode. The middle panel shows a transmission electron microscopy (TEM) image of a striated tube associated with a large vesicular structure (arrow). The bar diagrams (right panels) summarize the quantification of tubule length and striation distance based on at least three independent experiments. (C) Featureless tubes induced by the F-BAR domain of human Pacsin 1. Bright and dark field STEM images of a smooth, narrow tube are shown (left panel). The TEM image in the middle panel shows a featureless tube that is connected to a vesicular structure (arrow). Bar diagrams (right panels) summarize quantifications of tube length for narrow and wide featureless tubules from at least three independent experiments. (D) Membrane bulges and narrow tubules induced by the F-BAR domain of human Pacsin 1. Bright and dark field STEM images of vesicular structures induced by Pacsin are shown. TEM images show further examples of Pacsin-induced vesicular structures that are occasionally connected to narrow tubules (arrows).


A
No Protein
Endophilin A1BAR
Sorting Nexin 9full-length
Toca-1full-length
LMVs
LUVs (400 nm)
500 nm
B
Pacsin 1F-BAR
100 nm
1.3(±0.8)
Counts
Striated Tube Length (μm)
Peak
Valley
18(±3)
Striation Distance (nm)
C
Pacsin 1F-BAR
0.35(±0.08)
Thin Tube Length (μm)
1.4(±0.7)
Wide Tube Length (μm)
D
Pacsin 1F-BAR
Bright Field STEM
Dark Field STEM
TEM
TEM
TEM

Most strikingly though, the background of the electron micrographs was dominated by small vesicle structures (d=35±5 nm) that often were connected and clustered together resembling beads on a string (Figure. 2.2 and 2.3.D). These structures occurred at protein concentrations at which tubulation was observed suggesting that this activity was not due to excess F-BAR domains, unlike the vesiculation activity described for samples containing an excess of BAR domains (Peter et al., 2004). Both, narrow tubules and tube constrictions were absent in the LMV preparations in the absence of protein, and significantly less abundant in samples containing other BAR and F-BAR domain, indicating a Pacsin-specific effect (Figure 2.3.A-B). In contrast to isolated vesicles with irregular diameters found in LMV preparations, the Pacsin-induced morphologies were regular and significantly more abundant. The tight clustering of vesicular structures and their occasional association with narrow tubes suggested a pathway where constrictions emerge on highly curved tube structures (Figure 2.2.A and 2.3.D).

Comparable results regarding the occurrence of the various membrane morphologies were obtained by using large vesicles that were generated by extrusion through a 400 nm filter (Figure. 2.4.A). Usage of smaller vesicles with diameters between 50 and 200 nm abolished the formation of wider tubules but vesiculation remained apparent. Formation of tubules and vesicles was rapid, being visible and fully developed after short incubation times (Figure. 2.4.B). Incubation at 4°C as opposed to 25°C or 37°C prevented tubules with large diameter to form, and an increased abundance of high-curvature tubules was apparent suggesting that bilayer deformability and lipid order may play a role for the membrane sculpting potential (Figure. 2.4.B). In addition, Pacsin-mediated vesiculation was observed with inner-leaflet lipid mixtures comprising phosphatidylserine, phosphatidylcholine, and phosphatidylethanolamine instead of brain lipid extracts (data not shown).

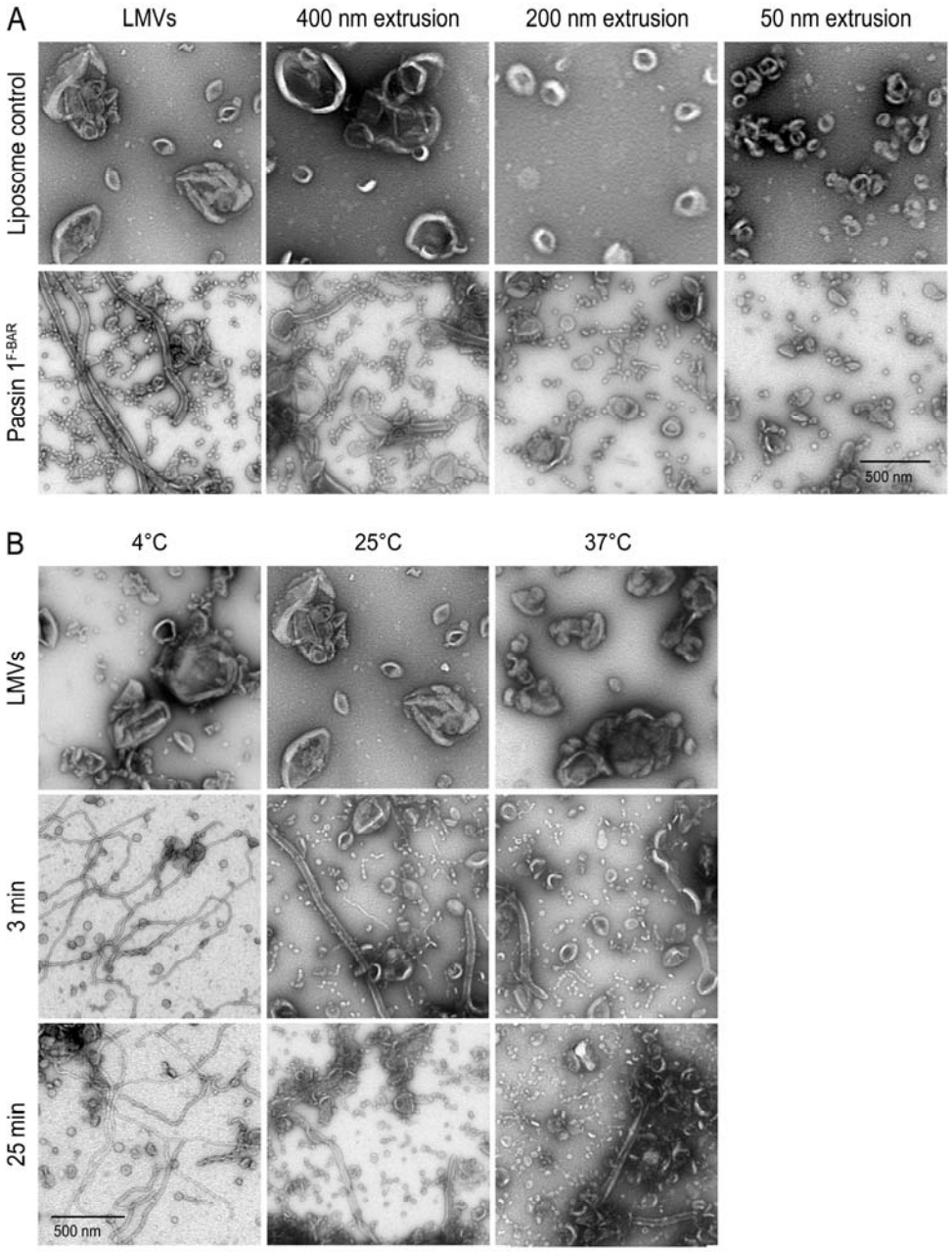


Figure 2.4. Dependence of Pacsin-mediated tubulation on liposomes size, temperature and time. (A) Liposome size dependence. Folch fraction I LMVs, or lipids extruded through filters with pore sizes of 400 nm, 200 nm, or 50 nm were incubated with the F-BAR domain of human Pacsin 1. Samples were spotted onto EM grids and stained with uranyl acetate. Grids were analyzed by electron microscopy. The scale bar is 500 nm and applies to all images in this figure. (B) Time course of tubulation at 4°C, 20°C and 37°C. Folch fraction I LMVs were incubated in the presence of Pacsin $1^{F\text{-}BAR}$ at the indicated temperatures, and samples were spotted onto EM grids and stained after 3 min and 25 min, followed by electron microscopic analysis.

Structures of the F-BAR domain of Pacsin 1 and 2.

In order to elucidate the molecular basis for the generation of Pacsin-specific membrane morphologies, we determined the crystal structures of the F-BAR domain of Pacsin 1 and 2. Screens using the full-length proteins yielded crystals of the F-BAR domains due to limited proteolysis over the course of the crystallization. The structure of the F-BAR domain from human Pacsin 1 was solved by single anomalous dispersion (SAD) phasing, using data collected on a crystal grown from selenomethionine-derivatized protein. The model was refined to a maximum resolution of 2.9 Å in the space group $P2_1$ (Figure. 2.5.A). Final refinement steps were aided by determining the structure of the isolated F-BAR domain (residues 1-325 of human Pacsin 1) at 2.8 Å in space group C2 using SAD phasing. The structure of the F-BAR domain of Pacsin 2 was solved by molecular replacement in space group $p2_12_12_1$, and was refined to a maximum resolution of 2.8 Å. The F-BAR domains of Pacsin 1 and 2 are very similar (r.m.s. deviation of 1.4 Å between all main chain atoms), and resemble the structure of F-BAR domains from CIP4, FBP17, and FCHo2 regarding overall fold and dimeric state, with a central six-helix bundle and distal tips composed of a helix-turn-helix motif (Figure. 2.5.A and 2.6.A) (Henne et al.,2007; Shimada et al., 2007). All F-BAR domains display a rather shallow concave surface that is enriched in positively charged residues that form the major membrane interaction interface (Figure. 2.5.A).

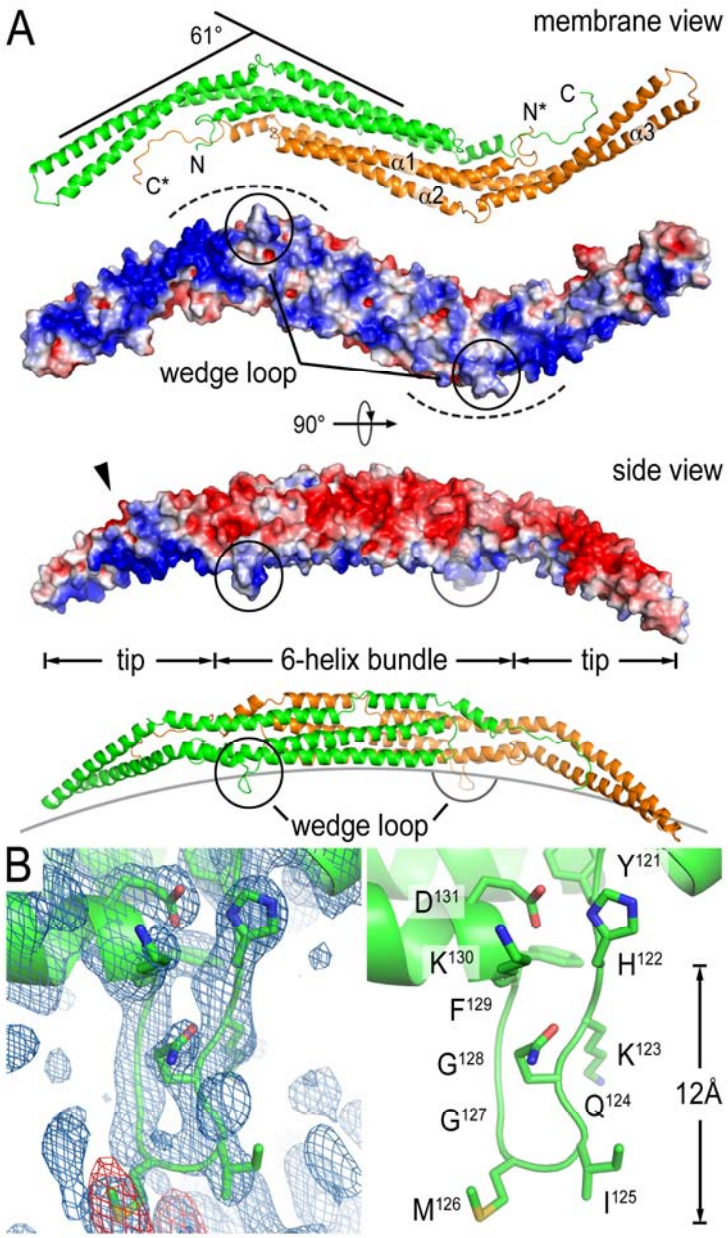


Figure 2.5. Structure of Pacsin's F-BAR domain. (A) Ribbon and surface presentation of the F-BAR domain dimer of Pacsin 1. Two perpendicular views are shown. Only two protomers are shown and are colored in green and orange, respectively. The electrostatic potential of the F-BAR domain was mapped onto its molecular surface, with red representing negative potential and blue representing positive potential (–4 to +4 k_BT). Electrostatic potentials were calculated by using the program APBS (Baker et al., 2001).(B) Close-up view of the wedge loop. The blue and red maps in the left panel show a (2Fo-Fc) omit map (contoured at 1.6σ) and an anomalous map (contoured at 3σ) around the wedge loop, respectively. The data collected on a selenomethionine-derivatized Pacsin 1 crystal (space group $P2_1$) was used. The right panel shows the same view with the wedge loop residues presented as sticks.

Table 2.1. Data collection and refinement statistics.

	Pacsin 1 (full-length)[1]	**Pacsin 1 (1-325)**	**Pacsin 2 (full-length)[1]**
Data collection[2,3]			
X-ray source	CHESS A1	CHESS A1	APS 24-ID-C
Wavelength (Å)	0.9793	0.9792	0.9745
Space group	$P2_1$	C2	$P2_12_12_1$
Unit cell parameters			
a, b, c (Å)	163.6, 32.2, 181.2	160.1, 85.5, 88.4	31.3, 88.4, 357.7
α, β, γ (°)	90.0, 111.2, 90.0	90.0, 117.6, 90.0	90.0, 90.0, 90.0
Resolution range (Å)	48.8 – 2.9 (3.06 – 2.9)	48.1 – 2.8 (2.94 – 2.8)	44.7 – 2.8 (2.95 – 2.8)
No. of reflections			
Total	199712 (27607)	196113 (27145)	104714 (10216)
Unique	76038 (11372)	51664 (7480)	25179 (3092)
Completeness (%)	97.7 (91.4)	97.5 (87.3)	95.4 (84.1)
Redundancy	2.6 (2.4)	3.8 (3.6)	4.2 (3.3)
$I/\sigma(I)$	9.9 (2.1)	10.1 (2.5)	14.2 (2.8)
R_{meas} (%)	9.3 (50.1)	11.2 (64.0)	9.5 (53.0)
Refinement[4]			
R_{work} / R_{free} (%)	28.5 / 30.9	21.5 / 26.0	22.4 / 27.4
r.m.s. deviations			
Bond length (Å)	0.002	0.006	0.003
Bond angles (°)	0.609	0.998	0.672
No. of atoms			
Protein	9531	4450	4717
Water	20	60	36
Average B-factors (Å^2)			
Protein	86.9	66.1	78.8
Water	70.4	54.5	55.1
Ramachandram (%)[5]			
Favored	96.5	97.3	96.8
Allowed	3.3	2.7	3.2
Generously allowed	0.2	0.0	0.0
Disallowed	0.0	0.0	0.0

Distinct features of the F-BAR domain of Pacsin.

A striking difference to previously determined structures of F-BAR proteins is the extreme degree of lateral kinking at the tip regions relative to the central dimerization region in Pacsin. Whereas the Toca-related F-BAR domains (from CIP4 and FBP17) are rather straight molecules and the one from FCHo2 has moderately kinked tips Henne et al., 2007; Shimata et al., 2007), Pacsin's tips are bent away from its central body in a ~61° angle, giving the dimeric molecule a pronounced twisted S-

shape (Figure. 2.5.A and 2.6.A). Although the curvature of the concave surface remains shallow similar to other F-BAR domains, it adds a second degree of curvature to the Pacsin F-BAR dimer. Interaction with the membrane will require that both principle curvatures of the protein will be satisfied indicating that Pacsin may have the tendency to bind to or stabilize membranes with higher curvature or vesicular structure. In contrast, other BAR and F-BAR domain-containing proteins may predominantly stabilize radial curvature, e.g. that of a tube (see below).

In this conformation, positively charged residues that mediate interactions with the negatively charged lipid head groups are concentrated at the concave surface but are arranged in a slight pitch lining the bending point and the tip regions, which results in the display of a twisted, non-flat surface with positive electrostatic potential (Figure. 2.5.A). The pitch and presence of both curvatures suggests a mechanism by which Pacsin not only tubulates bilayers but also may introduce or stabilize vesicular structures and potentially saddle points. Such regions of negative curvature occur during fission at the bud-neck and the plasma membrane-neck junctions during vesicle fission, and the interconnected, Pacsin-induced vesicles observed *in vitro* resemble such topologies (Figure. 2.5.A and 2.4.D).

Another feature that distinguishes Pacsin's F-BAR domain from that of other F-BAR domain-containing proteins is a unique, 8-residues-long insertion in helix 2 of Pacsin that forms a flexible loop in the crystal structures. The loops are located at the end of the central six-helix bundle close to the bending points of the tips, and protrude from the membrane interaction surface (Figure. 2.3.B and 2.5.B). At the base of the loop, there are two bulky hydrophobic residues, isoleucine^{125}-methionine^{126} (I^{125}-M^{126}) or a double methione$^{123\text{-}124}$ (M^{123}-M^{124}) motif in Pacsin 1 and 2, respectively.

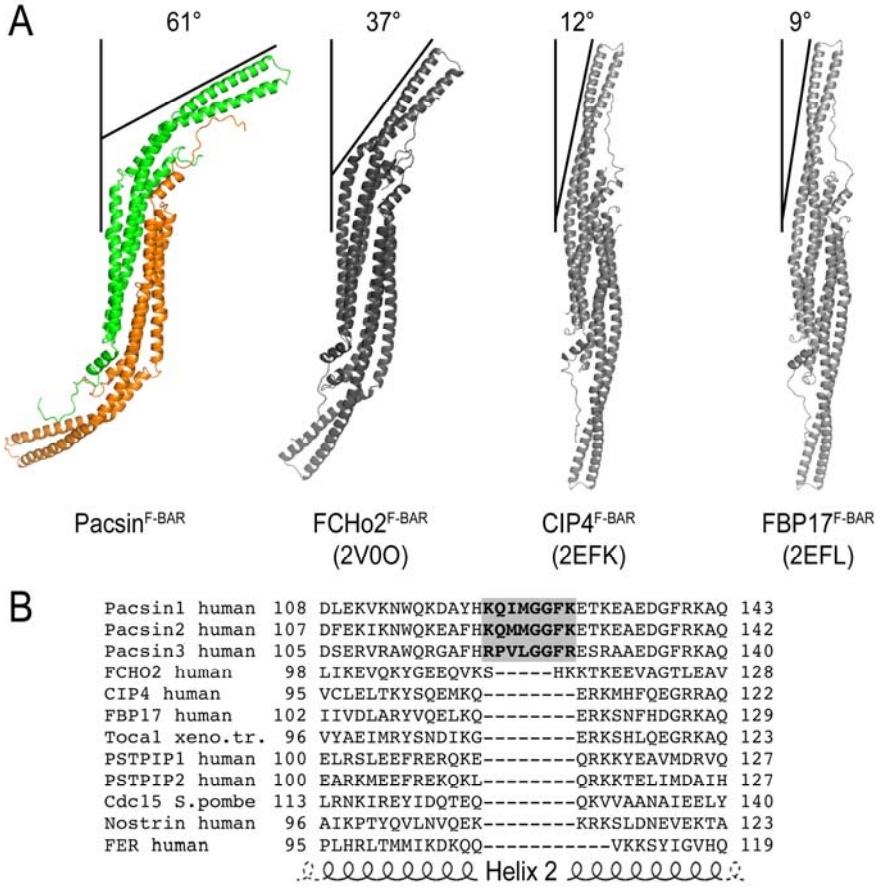


Figure. 2.6.. Structure of the F-BAR domain of Pacsin and comparison with other F-BAR domains.(A) Structurcal comparison of F-BAR domains. The structures of F-BAR domain dimers from Pacsin 1, FCHo2, CIP4, and FBP17 are shown (2, 3). The degree of tip bending (lateral curvature) was quantified by determining the angle between the long axis of the central 6-helix bundle and the long axis of the helical tip bundle.(B) Conservation of the wedge loop. A multiple sequence alignment prepared in Clustal W (19) is shown, and agrees with data presented previously (20). The wedge loop motif is shaded in grey.

The loop is flanked by lysines on either side (arginines in Pacsin 3), and contains a double glycine motif next to the central hydrophobic residues. Given the nature, position, and protrusion depth of the loop (~12 Å), we proposed a function equivalent to an amphipathic helix that would intercalate into one bilayer leaflet and would work as a wedge for membrane bending. We will refer to this motif as a wedge

loop. Similar motifs are of functional importance in proteins involved in fusion and fission.

Wild-type protein and site-directed mutants were analyzed with regard to their tubulation activity and membrane binding affinity. Protein binding to liposomes was assessed in a pelleting assay by detection of protein in the pellet or supernatant fraction after high-speed centrifugation. Replacing either isoleucine or methionine in the wedge loop of Pacsin 1's F-BAR domain with a negatively charged residue ($I^{125}E$ or $M^{126}E$) decreased the overall affinity of the protein for membranes accompanied by the loss of its membrane sculpting potential (Figure 2.7.A-C). In contrast, introduction of positively charged or tryptophan residues at those positions ($I^{125}K$, $M^{126}K$, $I^{125}W$, or $M^{126}W$) maintained near wild-type membrane affinity of the mutant proteins (insets in Figure. 2.7.D-E). While tryptophan is hydrophobic, its indole nitrogen may prevent it from deeply inserting into the membrane. In all cases except for the mutation $I^{125}W$, vesiculation was largely suppressed (Figure. 2.7.D-E). Instead, predominantly wide tubes ($M^{126}K$) or thin tubules ($I^{125}K$, $M^{126}W$), similar to morphologies that have been observed at low temperature, formed when LMVs were incubated with the mutant proteins. In samples prepared with proteins containing a tryptophan instead of the wedge loop isoleucine ($I^{125}W$), vesiculation dominated coinciding with a loss of tubular structures (Figure. 2.7.E). The results are consistent with a model in which the methionine functions as a wedge that acts close to the acyl chains of a membrane leaflet and drives vesiculation. The adjacent isoleucine serves a supporting role but may be closer to the lipid head group-acyl chain junction, explaining the bilateral

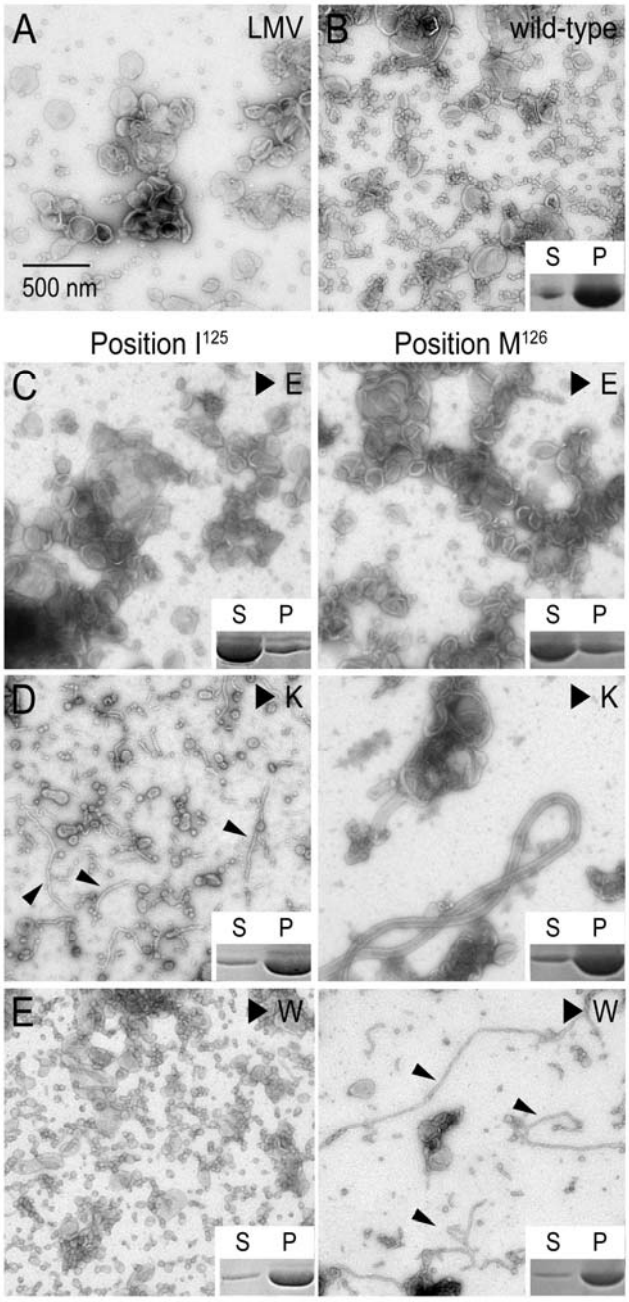


Figure 2.7. Mutational analysis of the wedge loop motif. A) Negative-stained micrograph of LMVs. The scale bar is 500 nm and applies to all panels in this figure. (B)-(E) Tubulation activity and membrane binding of wild-type and mutant Pacsin $1^{F\text{-}BAR}$. LMVs were incubated with wild-type Pacsin $1^{F\text{-}BAR}$ (B) or proteins with glutamate (C), lysine (D), or tryptophan (E) substitutions at positions 125 and 126 located at the center of the wedge loop. Small arrows point to thin tubular structures. Insets show results from vesicle co-pelleting assays assessing the binding of proteins to the liposomes. Supernatant and pellet fractions were analyzed by SDS-PAGE followed by Coomassie staining.

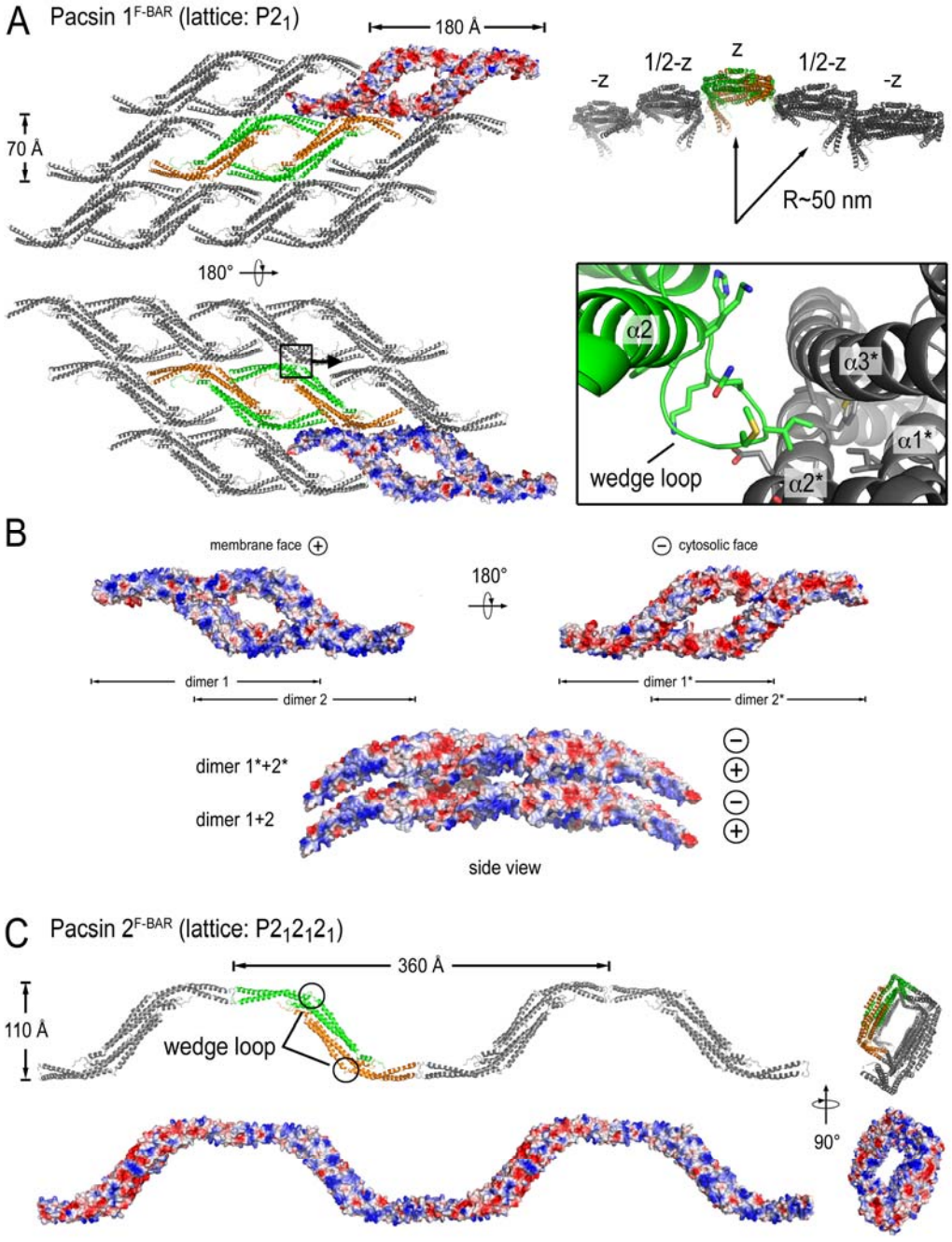
sensitivity to tryptophan or lysine mutations at that site. Decreased salt-sensitivity of the wild-type protein compared to the lysine mutants ($I^{125}K$, $M^{126}K$) in binding and tubulation assays support a wedge loop insertion mechanism (Figure 2.7).

Taken together, the structural and mutational analyses identified the wedge loop as a motif that contributes to membrane affinity and that determines Pacsin's unique vesiculation activity. Also, our data suggest that vesicles originated from tubules, which can form in the absence of loop insertions.

Higher-order arrangements in crystals of the F-BAR domain of Pacsin.

Previous studies revealed that the F-BAR domains of CIP4 and FBP17 could form ordered arrays that cover the membrane tube, and that distinct membrane curvatures can be attributed to different lattice arrangements (Shimata et al., 2007; Frost et al., 2008). Both, tip-to-tip and lateral intermolecular interactions are responsible for the assembly of protein lattices (Frost et al., 2008). Given the laterally curved state and distinct structural motifs of Pacsin, membrane engagement, lattice properties and spacing will be different than those of the more straight F-BAR domains studied thus far. In our crystallographic studies of Pacsin 1 and 2, the F-BAR domains pack in unique, primitive lattices that may provide insight into the Pacsin-specific membrane sculpting activities. Cryo-electron microscopic studies revealed that the F-BAR domains of CIP4 and FBP17 could form ordered helical arrays that cover the membrane tube, and that distinct membrane curvatures can be attributed to different

Figure 2.8. Crystallographic lattices of human Pacsin 1 and 2. (A) Pacsin 1. Crystals grown from full-length protein belong to space group $P2_1$. Two views, separated by a 180° rotation along the x-axis, of a sheet of F-BAR domains are shown (left). The molecules from one asymmetric unit are shown in orange and green, another asymmetric unit is shown in surface presentation. The electrostatic potential of the F-BAR domains were calculated in Pymol and mapped onto the molecular surface (red: negative; blue: positive). The top-right panel shows the same sheet in the z-direction. The inset illustrates a crystal lattice contact between a wedge loop of one chain and the 3-helix bundle of an adjacent molecule. (B) Stacking of Pacsin F-BAR sheets in the $P2_1$ crystal lattice of human Pacsin 1. Two asymmetric units are shown comprising two F-BAR domain dimers each. The electrostatic potential of the F-BAR domains was calculated in Pymol and mapped onto the molecular surface (red: negative; blue: positive). Orthogonal views are shown. The minus and plus signs indicate the anisotropic charge distribution of the dimers that is the basis for the stacking of the proteins in the crystal lattice.(C) Pacsin 2. Crystals grown from full-length protein belong to space group $P2_12_12_1$. Two perpendicular views are shown as ribbon (top) or surface (bottom) presentations. Molecules in one asymmetric unit are colored in green and orange. The surfaces are colored according to the electrostatic potential of the F-BAR domains, as described above.


A
Pacsin 1F-BAR (lattice: P21)
180 Å
70 Å
180°
-z
1/2-z
z
1/2-z
-z
R~50 nm
α2
α3*
α1*
α2*
wedge loop
B
membrane face ⊕
⊖ cytosolic face
180°
dimer 1
dimer 2
dimer 1*
dimer 2*
dimer 1*+2*
dimer 1+2
side view
C
Pacsin 2F-BAR (lattice: P21212121)
360 Å
110 Å
wedge loop
90°

lattice arrangements (Frost et al., 2008; Shimada et al., 2007). Both, tip-to-tip and lateral intermolecular interactions are responsible for the assembly of protein lattices (Frost et al., 2008). So far, isolated F-BAR domains, including that of Pacsin 1, crystallized in C-centered monoclinic lattices (space group C2), and provided only limited insight into potentially relevant arrays of F-BAR domain dimers (Shimada et al., 2007, Henne et al., 2007; this study). In our crystallographic studies of Pacsin 1 and 2, for which solutions of full-length proteins were used initially, the F-BAR domains pack in unique, primitive lattices (space groups $P2_1$ and $P2_12_12_1$, respectively) that provide a basis for the rationalization of Pacsin-specific membrane sculpting activities. Given the laterally curved state and distinct structural motifs of Pacsin, membrane engagement, lattice properties and spacing will be different than those of the more straight F-BAR domains studied thus far.

Indeed, we observed sheets of F-BAR domains in the crystals of Pacsin 1 mediated by interactions of the tips and the six-helix bundle (Figure. 2.8.A). The individual sheets in the crystal stacked via their anisotropic electrostatic potential (Figure. 2.8.B). Considering a single sheet of F-BAR domains from the crystal lattice, the residues of the protomers that contribute to the positive electrostatic potential of the F-BAR domain form a continuous surface for engagement with the membrane. In addition, a curvature of such a protein sheet is apparent in the z-direction that would be consistent with the curvature of the tubules with larger diameter (~50 nm; Figure. 2.8.A, top-right panel). The pair-wise interactions between F-BAR domain dimers are rather weak, and shape, charge and hydrophobic complementarity appear to dominate as driving forces for assembling this sheet. Notably, the wedge loops of some of the protomers within one sheet (and other crystal forms) are systematically tucked away between adjacent F-BAR domain dimers, and would not be available for insertion into the bilayer (Figure. 2.8.A, inset). Although it is not clear whether this particular

arrangement is relevant in the context of Pacsin-mediated membrane deformation, sequestration of the wedge loop in higher order assemblies provides an attractive mechanism for the modulation of Pacsin-induced membrane morphologies, which is in agreement with our mutagenesis data.

In crystals of the F-BAR domain of Pacsin 2, alternative lattice contacts were observed (Figure. 2.8.C). The F-BAR domains form spiral-like assemblies held together by tip-to-tip interactions. In such an oligomer, the membrane interacting surface would line the inside of the spiral, with the wedge loops available for membrane insertion. Occurrence of the spirals can be attributed in part to Pacsin's large degree of intrinsic lateral curvature. We speculate that these specific properties, in addition to high protein packing density on the membrane, could facilitate formation of narrow tubules (see below). This is consistent with the apparent trend that the moderately kinked F-BAR domain of FCHo2 can stabilize both narrower and wider membrane tubules, whereas the almost straight domains of CIP4 and FBP17 prefer large-diameter tubules (Shimada et al., 2007, Henne et al., 2007)

The spacing between repeating units in the spiral formed by Pacsin 2 is ~35 nm, similar in size to that of the membrane bulges or vesicles generated by Pacsin (Figure. 2.2 and 2.8.C). While such an oligomeric assembly may contribute to the formation of narrow tubes without requiring the insertion of the wedge loop in one bilayer leaflet, the wedge loop was critical for constriction of the tubules. The regular spacing of the wedge loops (and the regions of negative and positive curvatures within a F-BAR domain dimer) may be responsible for vesiculation activity, consistent with our mutational analysis.

In general, probing the relevance of these lattice contacts by targeted mutagenesis is complicated by the fact that single pair-wise interactions are rather weak, and are likely to be stabilized and driven by the presence of a membrane

support (Dommersnes and Fournier, 2002). Given the cooperative and interdependent events driving tubulation, including membrane-dependent oligomerization, alternative binding modes, and insertion mechanisms, it is challenging to discriminate between mutations that affect lattice formation directly by disrupting protein-protein interactions, or indirectly by decreasing the affinity for the membrane. It is likely that mutations that affect protein oligomerization also affect membrane affinity, and vice versa.

The two different types of higher-order arrangements observed with Pacsin 1 and 2, respectively, suggest a loose packing mode with lower density of F-BAR domains (Pacsin 1, space group $P2_1$; Figure. 2.8.A-B) that may be responsible for the stabilization of wide tubes, and a tight spiral-like assembly (Pacsin 2, space group $P2_12_12_1$; Figure. 2.8.C) that at high protein density may induce narrow tubule formation and possibly vesiculation (see below).

Energy estimations for protein-mediated membrane deformation

The energy cost to generate a tube of diameter 2R and Length L from a flat membrane can be described as:

$$E_t = \pi\kappa\frac{L}{R}$$

The energy cost per unit area (energy density) will therefore be:

$$\rho_t = \frac{E_t}{2\pi RL} = \frac{\kappa}{2}\times\frac{1}{R^2}$$

where κ is the membrane bending rigidity, generally estimated as 20 k_BT for a plasma membrane (Reynwar et al., 2007; Hu et al., 2008; Zimmerberg and McLaughlin, 2004). Accordingly, the energy densities (per nm^2) to generate large (R=30 nm) or small tubules (R=9 nm) are:

$$\rho_t(30) = 10k_BT\times\frac{1}{30^2} = 0.011k_BT \qquad \text{or} \qquad \rho_t(9) = 10k_BT\times\frac{1}{9^2} = 0.125k_BT$$

Similarly, the energy density to generate a spherical vesicle with a radius R=15 nm from a flat membrane can be estimated as:

$$\rho_v(15) = 2\kappa \times \frac{1}{R^2}\bigg|_{R=15} = 0.18 k_B T$$

The interaction between proteins and the membrane and its associated binding energy is residue-specific and sensitive to membrane rigidity. Roughly, each positive residue contributes 1 k_BT to the system, as determined for flexible poly-lysine peptides experimentally (Murray et al., 2002). However, due to the geometry constraints in the folded proteins and the variety of protein residues and lipid types, we use 0.5 k_BT as a rough estimation (Zimmerberg and McLaughlin, 2004; Zimmerberg and Kozlov, 2006).

The crystal structure of the CIP4 F-BAR dimer revealed that potentially ~40 positively charged residues along the concave surface and tip regions of one F-BAR domain dimer contribute to a protein-membrane interface. Electron microscopic reconstructions of CIP4-stabilized tubules revealed that only ~22 out of these 40 residues interacted with the membrane simultaneously (Frost et al., 2008). Based on these considerations, we estimated that one F-BAR domain dimer could contribute ~10 k_BT across an area of 75 nm^2. If we assume complete coverage of a membrane surface by F-BAR domains, the estimated energy density donated by the protein coat would be ~0.133 k_BT. Lattice formation and cooperativity may contribute to the energy donated by protein-membrane interactions but were omitted in these calculations for simplicity and due to limitations in estimating the extent of their contributions.

The energy cost to generate a tube with diameter 2R and Length L from a membrane with spontaneous curvature R_0 is:

$$E_t = \pi\kappa L(\frac{1}{R} - \frac{2}{R_0})$$

The energy density (per nm^2) can be estimated as:

$$\rho_t = \frac{E_t}{2\pi RL} = \frac{\kappa}{2} \times \frac{1}{R}\left(\frac{1}{R} - \frac{2}{R_0}\right) = \rho_t^{\,0}\left(1 - \frac{2R}{R_0}\right)$$

where $\rho_t^{\,0}$ is the energy density assuming a flat membrane sheet. If we consider a spontaneous curvature with R_0=300 nm, energetic estimations for the generation of wide tubules with R=30nm, differ by 20% compared to calculations assuming a flat sheet as starting point. For narrow tubules with R=10 nm, we overestimate the energy requirements by only 8%. The trend that low density is sufficient to generate wide tubes and higher density is required for generating highly curved tubules is valid, even if some degree of spontaneous membrane curvature is assumed.

Geometric considerations for protein-mediated membrane deformation

Total or Gaussian membrane curvature can be quantified by two parameters known as principle curvatures (Zimmerberg and Kozlov, 2006). For an infinite long tube of radius R, the principle curvature is $\left(\frac{1}{R}, 0\right)$. The curvature of a sphere of radius R can be described by $\left(\frac{1}{R}, \frac{1}{R}\right)$. In this calculation it is assumed that the direction of the principle curvature vector is identical to that of the reference axis. A rotation of the reference axis by an angle θ would change the description of the principle curvatures of a tube in the new reference frame to $\left(\frac{\cos\theta}{R}, \frac{\sin\theta}{R}\right)$, while the description of a sphere remains unchanged.

Comparing the geometric properties of a membrane tube with those of F-BAR domains allows us to rationalize the differences we observed in the tubulation assays, and reveals how structural variations translate into distinct membrane curvature preferences. The curvature of a F-BAR domain dimer with little or no lateral curvature such as that of CIP4 or FBP17 can be described as $\left(\frac{1}{R_0}, 0\right)$, with R_0 being the radius that can be accommodated by the convex surface of the dimeric domain. Considering that this F-BAR domain interacts with a membrane tube with the total curvature $\left(\frac{\cos\theta}{R}, \frac{\sin\theta}{R}\right)$, with θ being the angle between the long axis of the F-BAR domain

and the long axis of the tube, we can make the following prediction: An optimal geometrical fit would be achieved at

$$\frac{\cos\theta}{R} = \frac{1}{R_0} \quad \text{and} \quad \theta = \cos^{-1}\frac{R}{R_0}.$$

This relationship predicts a good geometrical fit when θ is considerably small, equivalent to R being close to R_0. In such a scenario, the second term, $\frac{\sin\theta}{R}$, would be close to zero (equivalent to an infinitely long tube), and both principle curvatures of the tube would be satisfied. The long axis of the F-BAR domain dimer would be oriented vertically to the long axis of the tube. If R is smaller but close to R_0, the second term would be dependent on θ, and the F-BAR domain would have to rotate slightly to compensate for the mismatch. If R is significantly smaller than R_0, a rotation of the F-BAR domain would result in a suboptimal fit, which could attenuate its membrane binding affinity.

The considerations above also indicate that straight F-BAR (and BAR) domains with only one principle curvature would be less suitable for stabilizing narrow tubules with radii R that are smaller than the radius R_0 imposed by the protein. In this case, adjusting the angle θ would not improve the fit, and only one principle curvature would be satisfied while the second principle curvature would deviate significantly from zero. Hence, the interactions of the protein dimers with the membrane would be suboptimal.

The total curvature of the F-BAR domain dimer of Pacsin contains two non-zero terms and can be described as $\left(\frac{1}{R_0}, \frac{1}{R_1}\right)$. Introducing a second principle curvature term, $\frac{1}{R_1}$, can compensate for the deviation of $\frac{\sin\theta}{R}$ in a regime when the tube radius R is smaller than R_0. An ideal binding mode can be achieved when R is close to R_1, which would satisfy both curvature terms. Based on these estimations, we hypothesize that the serpentine shape of Pacsin's F-BAR domain dimer may preferably stabilize thin tubes (as well as wide tubes).

Similar to the energetic estimations, this description is valid for an isolated F-BAR domain dimer, and the formation of ordered lattices is likely contribute to the tuning of curvature preference, as has been described for the F-BAR domains of CIP4 and FBP17 (Frost et al., 2008). A model for a helical packing mode of Pascin's F-BAR domain on a narrow tubule is shown in Figure 2.9.E. Lateral interactions between F-BAR domain dimers via their central 6-helix bundles have been observed in the crystal lattice of Pacsin 2, supporting such a model.

Calculations estimating the energy density requirements and geometrical constraints for Pacsin-mediated membrane deformation provided a quantitative assessment of the factors that contribute to the various liposome morphologies (see Supplementary Information for a detailed description of the calculations). The energy required to bend a membrane is in part dependent on the degree of curvature. While the stabilization of wide tubes (d=60 nm) requires only a small energy density (~0.011 k_BT/nm^2), the generation of narrow tubules (d=18 nm) or vesicles (d=30 nm) is energetically expensive (~0.125 k_BT/nm^2 or ~0.18 k_BT/nm^2, respectively) (Figure 2.9.A). Considering the electrostatic interactions between proteins and the membrane, a single F-BAR domain could contribute ~0.133 k_BT/nm^2 at most. Based on these estimations, the generation of wide tubes may only call for 10-20% of membrane coverage by protein. In contrast, the generation of thin tubules would require densely packed F-BAR domains on the membrane. The narrow tubules observed in the tubulation assays may be the highest degree of curvature that can be achieved by F-BAR domains. In addition, vesiculation cannot be rationalized solely by protein packing and is likely to involve other mechanisms such as the insertion of a wedge to increase local curvature (Zimmerberg and Kozlov, 2006; Campelo et al., 2008).

Pacsin's unique S-shaped conformation is likely to play a pivotal role in determining the total (Gaussian) membrane curvature preferred by this F-BAR domain.

According to our estimations, rather straight F-BAR domains with mainly one-dimensional intrinsic curvature may preferably orient on a membrane tube (close to) vertical to the long tube axis and appear optimized to maintain tubes with radii close to the radius imposed by the curvature of a F-BAR dimer (Figure 2.9.C). In contrast, the F-BAR domain of Pacsin constitutes a scaffold that is optimized for the stabilization of a high degree of membrane curvature due to its two-dimensional curvature profile. The Pacsin fold may also accommodate vesicular structures better than the F-BAR domains of CIP4, FBP17, and FCHo2 (data not shown).

Suppressed activities in full-length Pacsin

Experiments so far were carried out with the isolated F-BAR domain. Its domain boundaries were determined based on limited proteolysis and the available crystal structures. We next determined how the central linker and C-terminal SH3 domain affect Pacsin's membrane sculpting activity. When the linker (Pacsin $1^{\Delta SH3}$) or linker-SH3 module (Pacsin $1^{full\text{-}length}$) was included in the construct, Pacsin's membrane deformation activity was significantly impaired suggesting an autoinhibitory role of these units in Pacsin (Figure 2.9.A). In the case of the protein containing the linker region, full activity comparable to that of the isolated F-BAR domain could be achieved at a higher protein concentration used in the tubulation assays. In contrast, the autoinhibition in the full-length protein including the SH3 domain appeared stronger and could not be compensated for by high protein concentration (Figure 2.7.B).

In addition, we analyzed the curvature preference of the different constructs. Folch liposomes were extruded using filters with pore sizes raging from 50-400 nm, and protein binding was determined by using a vesicle pelleting assay (Figure 2.9.B). The isolated F-BAR domain showed no curvature preferences, binding equally well to

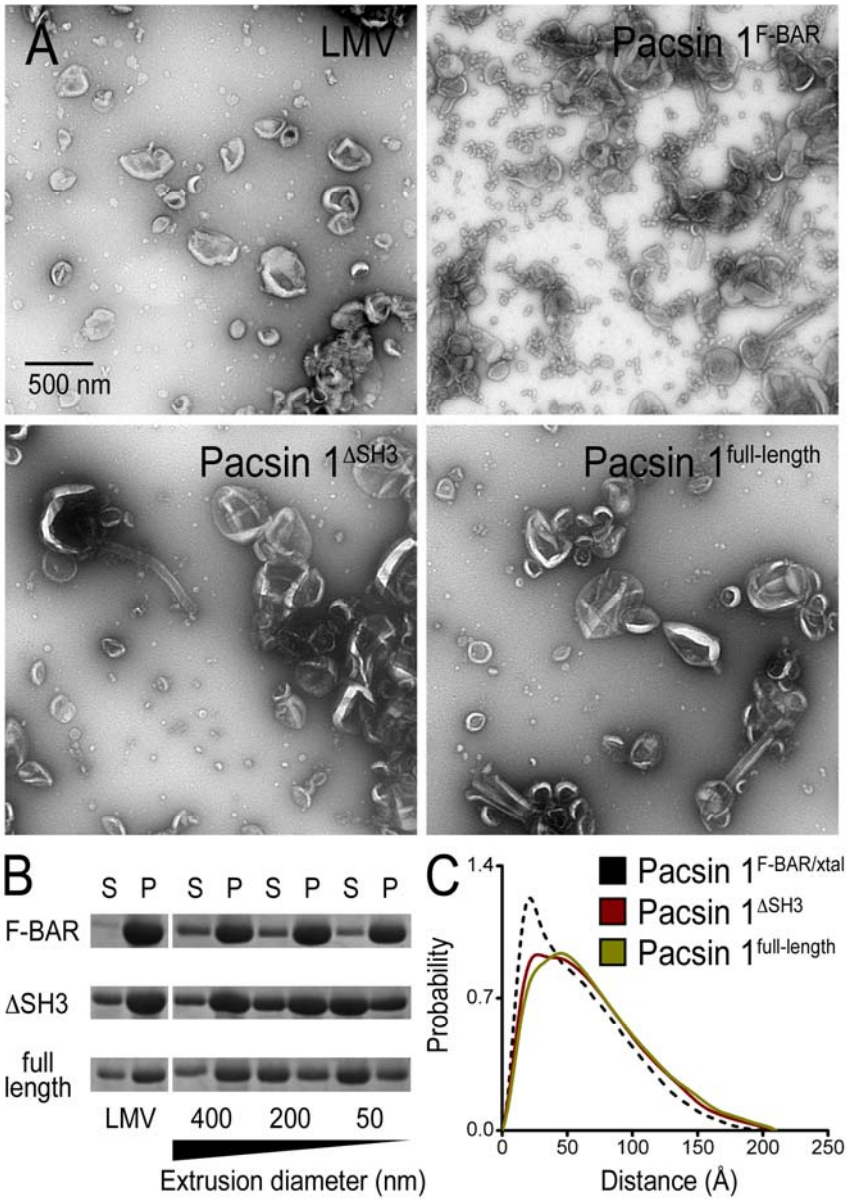
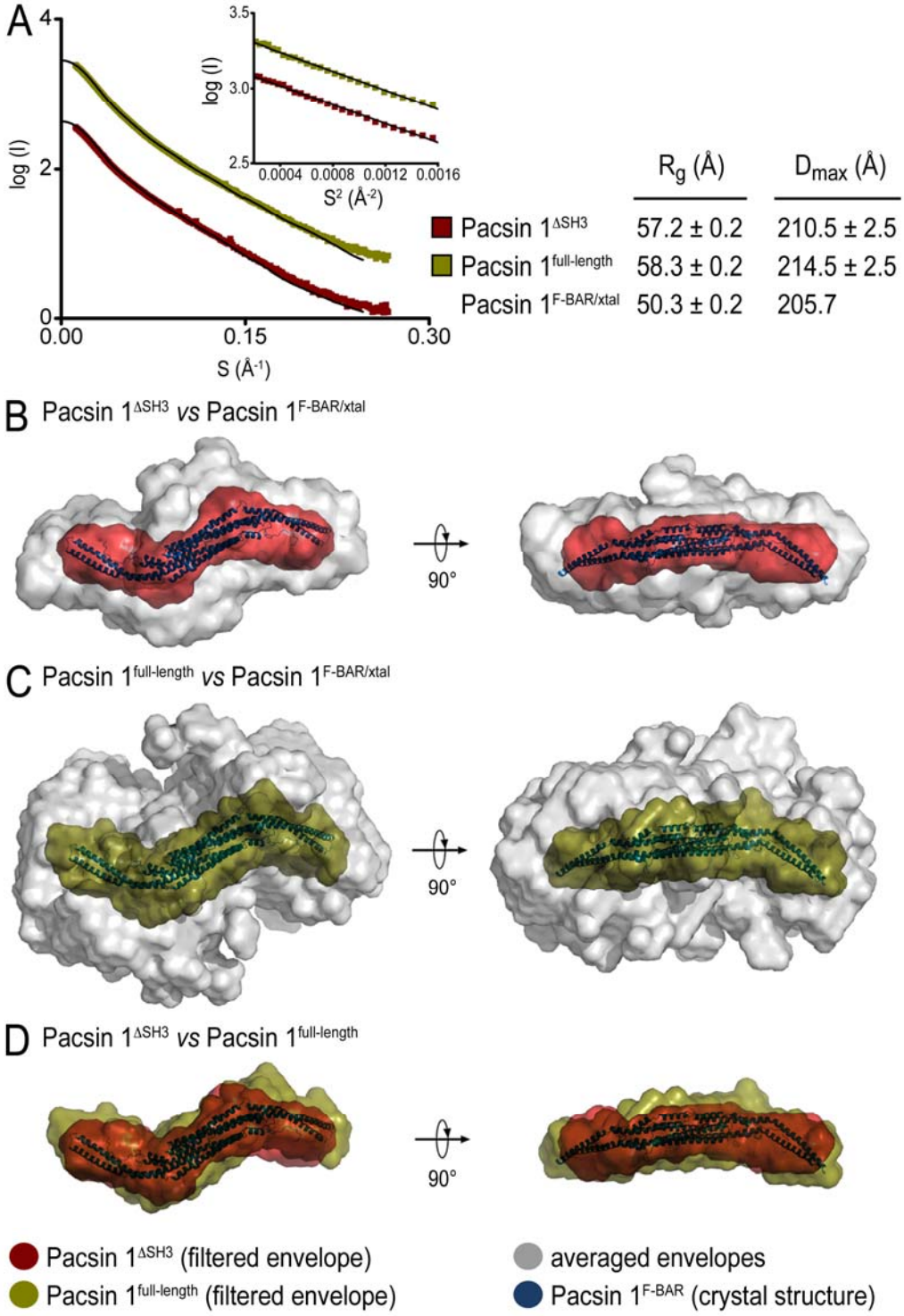


Figure. 2.9. Autoinhibition of human Pacsin 1. (A) Negative stain electron microscopy. LMVs were incubated with Pacsin $1^{F\text{-}BAR}$, Pacsin $1^{\Delta SH3}$, or Pacsin $1^{full\text{-}length}$. Samples were analyzed by negative stain EM. The scale bar is 500 nm and applies to all micrographs. (B) Membrane binding of Pacsin to liposomes of different size. LMVs or LMVs that were extruded through filters with pore sizes of 400, 200, or 50 nm were incubated with the three different Pacsin 1 construct. Samples were subjected to high-speed centrifugation. Pellet and supernatant fraction were analyzed on a Coomassie-stained SDS-PAGE. (C) Solution conformation of Pacsin $1^{\Delta SH3}$ and Pacsin $1^{full\text{-}length}$. Small angle X-ray scattering data was collected on homogeneous solutions of Pacsin $1^{\Delta SH3}$ and Pacsin $1^{full\text{-}length}$. Based on the scattering profile in solution, the distance distribution functions, P(R), were computed and compared to the P(R) function calculated from the crystal structure of Pacsin $1^{F\text{-}BAR}$.

vesicles of all sizes and comparable to LMVs. In contrast, the full-length protein bound preferentially to LMVs and liposomes extruded through a 400 nm filter. Pacsin $1^{\Delta SH3}$ yielded intermediate results. It still bound fairly well to liposomes with an average diameter of 200 nm, but binding to smaller vesicles was weaker. For the F-BAR domain of FCHo2, curvature-insensitive binding has been linked to the presence of an intact N-terminal amphipathic helix (Henne et al., 2007). In Pacsin, domains located at the C-terminus of the protein altered its curvature preference and overall membrane deformation potential.

The dissociation constant for the F-BAR domain and full-length dimer in solution were comparable as determined by analytical ultracentrifugation (2.1 μM and 1.2 μM, respectively), indicating that different dimerization propensities or additional intermolecular contacts via the additional domains are unlikely to contribute to the attenuated activity of full-length Pacsin (Figure 2.1.A). Small angle X-ray scattering (SAXS) studies revealed that the overall size (maximum diameter and radius of gyration) of the F-BAR domain dimer was comparable to the dimensions of the full-length protein, suggesting that the SH3 domains were positioned in close proximity to the F-BAR domains in solution (Figure 2.9.C). Modeling of SAXS data has been used successfully to determine the position of the SH3 domain of Endophilin in the context of the full-length protein and to reconstruct the PX-BAR domain module of Sorting nexin 9 (Wang et al., 2008). Yet, a similar approach indicated that the SH3 domains of Pacsin may be linked flexibly to a rigid F-BAR domain dimer in solution, since their position relative to the F-BAR domain could not be determined unambiguously (Figure 2.8). The shape of the F-BAR domain in full-length Pacsin modeled from the SAXS data was comparable to the structure obtained by X-ray crystallography,

Figure 2.10. Modeling of SAXS data. (A) Solution scattering profiles. Solution X-ray scattering profiles for Pacsin $1^{\Delta SH3}$ (red) and full-length Pacsin 1 (green) are shown. The inset presents the Guinier plot for the low-angle region. Radius of gyration (R_g) and the maximum diameter of the dimeric protein (D_{max}) was calculated based on the scattering data, or from the crystal structure in case of Pacsin $1^{F\text{-}BAR}$ ('xtal').(B) SAXS-based shape reconstruction of Pacsin $1^{\Delta SH3}$. *Ab initio* free atom modeling of the SAXS data yielded the shape reconstructions for Pacsin $1^{\Delta SH3}$ shown as surfaces. The grey envelope represents the averaged solution using 40 independent models. The red envelope represents the averaged and filtered solution (most reproducibly modeled volume). The crystal structure of Pacsin $1^{F\text{-}BAR}$ was docked into the envelope manually and is shown as ribbons. Two orthogonal views are shown. (C) SAXS-based shape reconstruction of Pacsin $1^{full\text{-}length}$. Modeling and visualization proceeded as described in (B) with the filtered envelope for Pacsin $1^{full\text{-}length}$ shown in green. (D) Superposition of solution structures and the crystal structure. The filtered envelopes determined from the SAXS data were superimposed, and the crystal structure of Pacsin $1^{F\text{-}BAR}$ was docked into the envelopes manually.



	R_g (Å)	D_{max} (Å)
Pacsin $1^{\Delta SH3}$	57.2 ± 0.2	210.5 ± 2.5
Pacsin $1^{full\text{-}length}$	58.3 ± 0.2	214.5 ± 2.5
Pacsin $1^{F\text{-}BAR/xtal}$	50.3 ± 0.2	205.7

Table 2.2. SAXS data statistics

	R_g (Å)[a]	D_{max} (Å)[b]	MM_{seq} (kDa)[c]	MM_{exp} (kDa)[d]	V_p (nm^3)[e]	$I_p^{norm}/I_{std}^{norm,f}$	NSD
PACSIN full-length	58.3 ±0.5	214.5 ±5.0	51.7	112.2±11	181±14	2.17±0.18	2.02
PACSIN ΔSH3	57.2 ±0.5	210.5 ±5.0	43.7	92.3±8	156±11	2.11±0.18	1.46

(a) R_g determined from Grunier equation and data that satisfied $Q_{max}xR_g<1.3$; similar values were obtained using the program GNOM (15).
(b) D_{max} was determined using the program GNOM (15).
(c) MM_{seq} is the molecular mass calculated from the primary sequence.
(d) MM_{exp} is the molecular mass calculated from the scattering data calculated by

$$MM_{protein} = MM_{lysozyme} \times \frac{I^o{}_{protein} \times C_{lysozyme}}{I^o{}_{lysozyme} \times C_{protein}}$$

where $MM_{lysozyme}$ is the molecular mass of lysozyme**,** I^0 is the scattering intensity and C (mg/ml) is the protein concentration (23).
(e) V_p is the excluded volume (Porod volume) calculated using PRIMUS (24). Data with S>0.25 were excluded from this calculation (23, 25).
(f) Normalized intensities for proteins (p) $I^p{}_{norm}$ were calculated by

$$I_p^{norm} = (\frac{I^o}{C}) / MM_{seq}$$

where I^0 is the intensity calculated using the Gunier equation and C (mg/ml) is the protein concentration determined from its absorbance at 280 nm. Lysozyme served as a mono-dispersed protein standard (std). Ideally, the ratio $I^p{}_{norm}/I^{std}{}_{norm}$ equals 1.00 (26, 27).
(g) c_{best} is the lowest c value obtained for a free atom model from *ab initio* shape determinations using DAMMIN.

indicating that the domain is indeed rigid in solution. We speculate that flexibility between the F-BAR and SH3 domains may interfere with efficient lattice or oligomer formation. Although intramolecular occlusion of certain motifs cannot be ruled out, this model would suggest that membrane deformation may not be determined by a single F-BAR domain but may require F-BAR domain multimers. Activation may be achieved by the binding of protein ligands to the SH3 or linker domains of Pacsin, or upon post-translational modifications.

CONCLUSION

The generation and stabilization of high membrane curvature is of fundamental biological importance for cellular membrane trafficking and maintenance of organellar structures, for example the endoplasmic reticulum (ER) (Zimmerberg and Kozlov, 2006). In the tubular ER, Reticulons, membrane-associated proteins, maintain such high curvature by the stable insertion of helical elements into one leaflet of the bilayer, providing a more permanent anchorage (Hu et al., 2008; Voeltz et al., 2006). In contrast, endocytosis and membrane trafficking relies on the reversible association of proteins at the lipid bilayer resulting in its deformation, fission, or fusion.

From our structural and functional data, we propose a model in which tubulation and Pacsin-specific membrane bulging activities rely on the distinct curvatures and particular distribution of positively charged residues within a Pacsin dimer, in addition to the insertion of the wedge loops into the membrane. Similar to other F-BAR domain-containing proteins, Pacsin's full deformation potential is likely to rely on higher-order packing on the membrane, tunable and alternative lattices stabilizing different degrees of curvature. At the same time, the association of Pacsin with itself and the membrane is rather labile, based on the weak interactions observed crystallographically and its salt and dilution sensitivity in tubulation assays (data not shown). The wedge loop is crucial for the vesiculation activity and we speculate that Pacsin may contribute more directly to fission by the insertion of this motif that resembles fusion and fission loops in other proteins.

Recently, it has been shown that Dynamin-mediated fission proceeds through a hemi-fission event (Bashkirov et al., 2008; Pucadyil et al., 2008). Scission succeeded spontaneously upon GTPase-triggered disassembly of Dynamin coats that initially introduced an unstable intermediate. Pacsin may facilitate initial steps that will aid Dynamin-mediated vesicle scission, or may play a regulatory role by forming

metastable lattices on the membrane while stressing the bilayer via wedge loop insertion. It has been shown that perturbing Pacsin function under intense action potential stimulation decreases the pool of endocytic vesicles and increases the number of cisternae at the synapse (Andersson et al., 2008). While Pacsin's role in recruiting Dynamin may explain such an effect on vesicle recycling, it may also suggest a more central role for Pacsin under these conditions. In *Drosophila*, the F-BAR domain of Pacsin was recently shown to be required for the expansion of the postsynaptic membrane system while the SH3 domain serves both an inhibitory as well as a targeting role (Kumar et al., 2009). These results are consistent with our findings and support our model. It remains to be investigated if there is an apparent redundancy in the activities of Pacsin and other factors involved in endocytosis such as BAR domain-containing proteins, Dynamin, and EHD-type ATPases, and how their activities are coordinated and utilized in the cell to drive vesicle fission.

REFERENCES

Adams P., et al. (2002) PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. D 58:1948-1954.

Andersson F., Jakobsson J., Löw P., Shupliakov O., Brodin L. (2008) Perturbation of syndapin/PACSIN impairs synaptic vesicle recycling evoked by intense stimulation. J. Neurosci. 28:3925-3933.

Anggono V., et al. (2006) Syndapin I is the phosphorylation-regulated dynamin I partner in synaptic vesicle endocytosis. Nat. Neurosci. 9:752-760.

Bashkirov P.V., et al. (2008) GTPase cycle of dynamin is coupled to membrane squeeze and release, leading to spontaneous fission. Cell 135:1276-1286.

Braun A., et al. (2005) EHD proteins associate with syndapin I and II and such interactions play a crucial role in endosomal recycling. Mol. Biol. Cell. 16:3642-3658.

Campelo F., McMahon H.T., Kozlov M.M. (2008) The hydrophobic insertion mechanism of membrane curvature generation by proteins. Biophys. J. 95:2325-2339.

DiProspero N.A., et al. (2004) Early changes in Huntington's disease patient brains involve alterations in cytoskeletal and synaptic elements. J. Neurocytol. 33:517-533.

Dommersnes P.G., Fournier J.B. (2002) The many-body problem for anisotropic membrane inclusions and the self-assembly of "saddle" defects into an "egg carton". Biophys. J. 83:2898-2905.

Emsley P., Cowtan K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D 60:2126-2132.

Farsad K., et al. (2001) Generation of high curvature membranes mediated by direct endophilin bilayer interactions. J Cell Biol 155:193-200.

Frost A., et al. (2008) Structural basis of membrane invagination by F-BAR domains. Cell 132:807-817.

Fotin A., et al. (2004) Molecular model for a complete clathrin lattice from electron cryomicroscopy. Nature 432:573-579.

Gallop J., et al. (2006) Mechanism of endophilin N-BAR domain-mediated membrane curvature. EMBO J. 25:2898-2910.

Kabsch W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Cryst. 26:795-800.

Kessels M.M., Qualmann B. (2004) The syndapin protein family: linking membrane trafficking with the cytoskeleton. J. Cell Sci. 117:3077-3086.

Kumar V., et al. (2009) Syndapin Promotes Formation of a Postsynaptic Membrane System in Drosophila. Mol Biol Cell 12(8):221-236.

Henne W.M., et al. (2007) Structure and analysis of FCHo2 F-BAR domain: a dimerizing and membrane recruitment module that effects membrane curvature. Structure 15:839-852.

Hu J., et al. (2008) Membrane proteins of the endoplasmic reticulum induce high-curvature tubules. Science 319:1247-1250.

Itoh T., et al. (2005) Dynamin and the actin cytoskeleton cooperatively regulate plasma membrane invagination by BAR and F-BAR proteins. Dev. Cell 9:791-804.

Jeffries C.M., Whitten A.E., Harris S.P., Trewhella J. (2008) Small-angle X-ray scattering reveals the N-terminal domain organization of cardiac myosin binding protein C. J. Mol. Biol. 377:1186-1199.

Konarev P.V., Volkov V.V., Sokolova A.V., Koch M.H.J., Svergun D.I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. J. Appl. Cryst. 36:1277-1282.

Laskowski R.A., Moss D.S., Thornton J.M. (1993) Main-chain bond lengths and bond angles in protein structures. J. Mol. Biol. 231:1049-1067.

Masuda M., et al. (2006) Endophilin BAR domain drives membrane curvature by two newly identified structure-based mechanisms. EMBO J 25:2889-2897.

Modregger J., Ritter B., Witter B., Paulsson M., Plomann M. (2000) All three PACSIN isoforms bind to endocytic proteins and inhibit endocytosis. J. Cell Sci. 113:4511-4521.

Modregger J., DiProspero N.A., Charles V., Tagle D.A., Plomann M. (2002) PACSIN 1 interacts with huntingtin and is absent from synaptic varicosities in presymptomatic Huntington's disease brains. Hum. Mol. Genet. 11:2547-2558.

Murray D., Arbuzova A., Honig B., McLaughlin S. (2002) The role of electrostatic and nonpolar interactions in the association of peripheral proteins with membranes. Current Topics in Membranes, eds Simon S.A., McIntosh T.J. (Academic Press), Vol 52, pp 271-379.

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

Petoukhov M.V., Konarev P.V., Kikhney A.G., Svergun D.I. (2007) ATSAS 2.1 - towards automated and web-supported small-angle scattering data analysis. J. Appl. Cryst. 40:S223-S228.

Petoukhov M.V., Svergun D.I. (2006) Joint use of small-angle X-ray and neutron scattering to study biological macromolecules in solution. Eur. Biophys. J. 35:567-576.
Porod G. (1982) General Theory. Small-angle X-ray scattering, ed Kratky O (Academic Press, London), pp 17-51.

Pucadyil T.J., Schmid S.L. (2008) Real-time visualization of dynamin-catalyzed membrane fission and vesicle release. Cell 135:1263-1275.

Qualmann B., Kelly R.B. (2000) Syndapin isoforms participate in receptor-mediated endocytosis and actin organization. J. Cell Biol. 148(5):1047-1062.

Qualmann B., Roos J., DiGregorio P.J., Kelly R.B. (1999) Syndapin I, a synaptic dynamin-binding protein that associates with the neural Wiskott-Aldrich syndrome protein. Mol. Biol. Cell 10:501-513.

Reynwar B.J., et al. (2007) Aggregation and vesiculation of membrane proteins by curvature-mediated interactions. Nature 447:461-464.

Sheldrick G.M. (2008) A short history of SHELX. Acta Crystallogr. A 64:112-122.

Shimada A., et al. (2007) Curved EFC/F-BAR-domain dimers are joined end to end into a filament for membrane invagination in endocytosis. Cell 129:761-772.

Svergun D.I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J. Appl. Cryst. 25:495-503.

Svergun D., Barberato C., Koch M.H.J. (1995) CRYSOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. J. Appl. Cryst. 28:768-773.

Svergun D.I., Petoukhov M.V., Koch M.H. (2001) Determination of domain structure of proteins from X-ray solution scattering. Biophys. J. 80:2946-2953.

Takei K., McPherson P.S., Schmid S.L., De Camilli P. (1995) Tubular membrane invaginations coated by dynamin rings are induced by GTP-gamma S in nerve terminals. Nature. 374:186-190.

Takei K., Slepnev V.I., Haucke V., De Camilli P. (1999) Functional partnership between amphiphysin and dynamin in clathrin-mediated endocytosis. Nat. Cell Biol. 1:33-39.

Taraban M., et al. (2008) Ligand-induced conformational changes and conformational dynamics in the solution structure of the lactose repressor protein. J. Mol. Biol. 376:466-481.

Terwilliger T.C. (2000) Maximum-likelihood density modification. Acta Crystallogr. D 56:965-972.

Thompson J.D., Gibson T.J., Plewniak F., Jeanmougin F., Higgins D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25:4876-4882.

Tsujita K., et al. (2006) Coordination between the actin cytoskeleton and membrane deformation by a novel membrane tubulation domain of PCH proteins is involved in endocytosis. J. Cell Biol. 172:269-279.

Voeltz G.K., Prinz W.A., Shibata Y., Rist J.M., Rapoport T.A. (2006) A class of membrane proteins shaping the tubular endoplasmic reticulum. Cell 124:573-586.

Volkov V.V., Svergun D.I. (2003) Uniqueness of ab initio shape determination in small-angle scattering. J. Appl. Cryst. 36:860-864.

Wang Q., Kaan H.Y., Hooda R.N., Goh S.L., Sondermann H. (2008) Structure and plasticity of Endophilin and Sorting Nexin 9. Structure 16:1574-87.

Weissenhorn W. (2005) Crystal structure of the endophilin-A1 BAR domain. J. Mol. Biol. 351:653-661.

Yarar D., Surka M.C., Leonard M.C., Schmid S.L. (2008) SNX9 activities are regulated by multiple phosphoinositides through both PX and BAR domains. Traffic 9:133-146.

Zimmerberg J., Kozlov M.M. (2006) How proteins produce cellular membrane curvature. Nat. Rev. Mol. Cell Biol. 7:9-19.

Zimmerberg J., McLaughlin S. (2004) Membrane curvature: how BAR domains bend bilayers. Curr. Biol. 14:R250-252.

CHAPTER 3

MOLECULAR MECHANISM AND ACTIVATION OF PACSIN-MEDIATED MEMBRANE DEFORMATION

ABSTRACT

Endocytosis is a fundamental process in neuronal transmission and membrane trafficking. The formation of vesicles at the plasma membrane is mediated by enzymes such as Dynamin that catalyze the final fission step, the actin cytoskeleton, and proteins that sense or induce membrane curvature. The F-BAR domain-containing protein Pacsin-1 contributes to this process, and has been shown to induce a spectrum of membrane deformation including tubules and tube constrictions *in vitro*. The full-length protein has reduced activity compared to the isolated F-BAR domain, suggesting an inhibitory role for the C-terminal SH3 domains. Here we show that the three different human Pacsin isoforms have unique membrane sculpting potentials and inhibition levels *in vitro*. Activation of Pacsin-1's membrane deformation capacity is achieved by engaging the SH3 domains with the non-catalytic proline-rich domain of Dynamin. Depending on the experimental conditions, we observed distinct membrane morphologies ranging from tubules with variable diameter to small vesicles, which may suggests a more versatile role for Pacsin in sculpting cellular membranes. A theoretical assessment of Pacsin F-BAR domain binding to membranes suggests a mixed mechanism involving scaffolding and the insertion of amphipathic motifs into the lipid bilayer.

INTRODCUTION

Local differences and dynamic changes in curvature are hallmarks of cellular membranes, contributing to the identity of organelles such as the endoplasmic reticulum and to mechanisms in membrane trafficking and endocytosis (Shibata et al., 2009). Peripheral and integral membrane proteins have been identified that either promote or stabilize membrane curvature at different sites in the cell and for various cellular functions. For example, endocytosis, the uptake of material into cells, relies on the coordinated interplay of protein coats that membrane vesicles budding, and other proteins that stabilize the bud-neck structure and recruit the large GTPase Dynamin, which subsequently will facilitate the final fission (Doherty and McManhon, 2009; Henne et al., 2010; Pucadyil et al., 2009; Ferguson et al., 2009). In addition, reorganization of the actin cytoskeleton via the recruitment and activation of Wiskott-Alrich Syndrome proteins (WASP) provides a driving force in this process.

Proteins containing a BAR (Bin/amphyphysin/rvs) domain have emerged as facilitators of membrane trafficking and the fission process by directly stabilizing tubular membrane structures *in vitro* and in cells. They can be divided into three distinct structural classes based on their deformation activity and structures: BAR and N-BAR domain-containing proteins (e.g. Endophilin, Amphyphysin, Sorting nexin 9, and APPL1) that prefer highly curved membranes; F-BAR domain-containing proteins (e.g. CIP4, FChO2, Pacsin/Syndapin) that often are associated with wider tubes; and inverse or I-BAR domain-containing proteins (e.g. IRSp53; MIM) that induce membrane invaginations (Farsad et al., 2001; Peter et al., 2004; Takei et al., 1999; Itoh et al., 2006; Tsujita et al., 2006; Mattila et al., 2007).

The general fold of BAR domains consists of three helices that form a six-helix bundle in a dimeric assembly, the predominant quaternary structure in solution (Peter,

2004). The overall preference for distinct membrane curvatures is partially encoded in the particular folds of the different subfamilies. The dimeric BAR domains present an overall crescent shape. In the case of BAR and N-BAR domains, the concave surface is lined with positively charged residues and other motifs involved in membrane interactions (Gallop et al., 2006; Jao et al., 2010; Masuda et al., 2006; Peter et al., 2004; Weissenhorn, 2005). The intrinsic curvature of N-BAR domains is higher than in F-BAR domains characterized to date, and the lower degree of protein curvature of the latter often matches their preference for wider membrane tubes (Shimada et al., 2007; Henne et al., 2007; Frost et al., 2008). In contrast, it is a convex surface in I-BAR proteins that mediates membrane interactions and promotes filopodia formation (Saarikangas et al., 2009).

Recently, some exceptions to these correlations have been reported for F-BAR domain-containing proteins. In addition to its canonical function of stabilizing wide tubes, the F-BAR domain of FCHo2 also facilitates the formation of highly curved tubes (Henne et al., 2010; Henne et al., 2007). srGAP2, a protein involved in neuronal migration and morphogenesis, contains a F-BAR domain based on its primary sequence but induces I-BAR-like membrane protrusions (Guerrier et al., 2009). Another example is Pacsin (also known as Syndapin) that has been shown to induce a wide range of membrane deformation, in particular membrane tubes of various diameters, pearling structures and invaginations (Wang et al., 2009; Rao et al., 2010; Shimada et al., 2010). Structural and functional analyses revealed multiple Pacsin-unique features that are likely contributing to its unique morphogenetic potential (Rao et al., 2010; Shimada et al., 2010; Edeling et al., 2009; Wang et al., 2009; Plomann et al., 2010). Its F-BAR domain shows a distinct lateral curvature in addition to its concave surface, and these geometric constraints may contribute to the resulting Pacsin-dependent membrane morphologies and its potential to form higher-order

lattices on lipid bilayers (Wang et al., 2009) (Figure 3.1.A). Another striking feature is a short insertion in helix 2 that forms a amphipathic wedge proposed to dip into the acyl chain layer of one lipid bilayer leaflet (Rao et al., 2010; Shimada et al., 2010; Wang et al., 2009; Plomann et al., 2010) (Figure 3.1.A). Hydrophobic insertion of amphipathic helices or loops has been identified as one of the main driving forces in the generation of membrane curvature (Gallop et al., 2006; Zimmerberg and Kozlov, 2006; Campelo et al., 2008). Other factors that may contribute are protein oligomerization and electrostatic interactions of the curved scaffold with the membrane(Frost et al., 2008; Yarar et al., 2007a; Yarar et al., 2007b; Wang et al., 2008; Wang et al., 2009; Shimada et al., 2007).

Pacsin is required for neuromorphogenesis by influencing cellular motility (Edeling et al., 2009; Dharmalingam et al., 2009). On a cellular level, Pacsin is contributes to clathrin-dependent endocytosis and the formation of synaptic vesicles, functions that require membrane localization and complex formation with the endocytotic machinery and the cytoskeleton via its C-terminal SH3 domain (Kessels et al., 2002; Modregger et al., 2000; Qualmann et al., 2000 ; Qualmann et al., 2000; Kumar et al., 2009; Dharmalingam et al., 2009). In addition to binding N-WASP/WASP among others, the SH3 domain of Pacsin engages the proline-rich domain (PRD) of Dynamin. This interaction is sensitive to the phosphorylation state of the PRD, providing a mechanism for the spatial and temporal regulation of Pacsin-mediated membrane deformation (Andersson et al., 2008; Anggono et al., 2006).

Recently, we reported the autoinhibition of Pacsin1 *in vitro*, which was consistent with the suppressed activity of its fruit fly-homolog *in vivo* (Kumar et al., 2009; Wang et al., 2009). Both studies identified the SH3 domain as an autoinhibitory feature. Based on small-angle X-ray scattering (SAXS) experiments, we proposed a compact structure for full-length Pacsin, in which the SH3 domains fold back onto the

F-BAR domain dimer and that would be disrupted by ligand binding to the SH3 domain (Wang et al., 2009). Confirmation of such a model came from the crystal structure of full-length Pacsin and functional studies revealing that the PRD of Dynamin activates Pacsin's tubulation activity (Rao et al., 2010).

Here, we demonstrate that the three human isoforms of Pacsin, the brain-specific Pacsin 1 and the more widely expressed Pacsin 2 and 3, show distinct membrane deformation and autoinhibition potentials *in vitro*. While Pacsin 2 and 3 appear more active with regard to tubulation, all three isoforms are subject to further activation by addition of the PRD of Dynamin in a phosphorylation-dependent manner. Depending on the way liposomes were prepared, Pacsin 1 stabilizes more vesicular or more tubular structures in the reconstituted system. Full activity requires an intact amphipathic wedge loop and the SH3 domain ligand-binding site. While N-BAR domain-mediated deformation relies predominantly on hydrophobic insertion of protein motifs into one of the bilayer leaflets, calculations based on the structural analysis of Pacsin suggest a hybrid mechanism involving electrostatic interactions, wedge-loop insertion and protein oligomerization.

MATERIALS AND METHODS

Protein expression and purification

The coding regions corresponding to full-length human Pacsin 1 2, and 3, Pacsin $1^{F\text{-}BAR}$ (residues 1-325), Pacsin $2^{F\text{-}BAR}$ (residues 1-324) and Pacsin $3^{F\text{-}BAR}$ (residues 1-324) were amplified by standard PCR and cloned into a modified pET28a expression plasmid (Novagen) yielding N-terminally hexahistidine-tagged SUMO fusion proteins. The hexahistidine-tagged SUMO-moiety was cleavable using the protease Ulp-1 from *S. cerevisiae*. The coding regions corresponding to Dynamin1 PRD domain(residue

724-822) was cloned into pGEX6p1 expression plasmid yielding N-terminally GST fusion protein.

Native and selenomethionine-derivatized proteins were overexpressed in *Escherichia coli* cells BL21 (DE3) (Novagen) or in T7 Crystal Express (NEB), respectively. For the expression of native proteins, cells were grown in Terrific Broth (TB) media supplemented with 50 mg/l kanamycin at 37°C. Selenomethionine-derivatized proteins were produced in cells grown in M9 minimal media supplemented with 50 mg/l kanamycin and 40 mg/l selenomethionine. At a cell density corresponding to an absorbance of 0.8 at 600 nm, the temperature was reduced to 18°C, and protein production was induced with 1 mM IPTG. After 16 hours, cells were harvested by centrifugation, resuspended in NiNTA buffer A (25 mM Tris-Cl, pH 8.2, 500 mM NaCl and 20 mM imidazole), and flash-frozen in liquid nitrogen.

After cell lysis by sonication and removal of cell debris by centrifugation, clear lysates were loaded onto a NiNTA columns (HisTrap; GE Healthcare) equilibrated in NiNTA buffer A. The resin was washed with 20 column volumes of NiNTA buffer A, and proteins were eluted on a single step of NiNTA buffer A supplemented with 500 mM imidazole. Proteins were buffer exchanged into desalting buffer (25 mM Tris-HCl, pH 7.4, 400 mM NaCl), and affinity tags were removed by incubation with the protease Ulp-1 at 4°C overnight. Cleaved proteins were collected in the flow-through during NiNTA affinity chromatography, and were subjected to size exclusion chromatography on a Superdex 200 column (GE Healthcare) equilibrated in gel filtration buffer (25 mM Tris-HCl, pH 7.4, 400 mM NaCl). Proteins were concentrated on a Centricon ultrafiltration device (10 kDa cutoff; Millipore) to a final concentration of ~0.5-1 mM (~25-50 mg/ml). Protein aliquots were frozen in liquid nitrogen and stored at -80°C.

Small angle X-ray scattering (SAXS) and SAXS-based shape reconstruction

SAXS experiments were carried out at the Advanced Photon Source (APS, BESSRC-CAT, beamline 12-IDC, Argonne, IL) at the electron energy of 12 KeV and 4°C. Prior to data collection, samples were centrifuged for 10 min at 13,200 xg. Scattering data were collected in triplicates at protein concentrations between 0.5-4 mg/ml. Exposure times (usually 2 seconds) yielding good signal-to-noise but minimal radiation damage was determined experimentally. Background scattering was collected from low salt gel filtration buffer (25 mM Tris-HCl, pH 7.4, 150 mM NaCl), and scattering data was background-corrected, averaged, scaled, and analyzed using the programs IGOR PRO (version 5.04B) , GNOM, and CRYSOL (Svergun 1992, Svergun et al.,1995). Only data with $S_{max}*R_g<1.3$, computed from Guinier plots at low angle regions, were considered for further analysis. Kratky plots, experimental molecular weight determinations, and the normalized scattering intensity were used to assess the folded-state of the proteins and overall data quality. Distance distribution functions were determined using the program GNOM (Svergun et al., 1995).

Ab initio free atom modeling was performed by using the program GASBOR22 with the distance distribution function as the fitting target (Svergun et al., 2001). Proteins were found to be dimeric under the experimental conditions as confirmed by sedimentation equilibrium analytical ultracentrifugation and multi-angle light scattering (data not shown), and P2 symmetry was applied during modeling. Forty independent simulations were carried out for each protein. Solutions were superimposed, averaged, and filtered using the program DAMAVER (Volkov and Svergun, 2003). Data statistics are summarized in Table 3.2.

Liposome preparation

Folch fraction 1 lipids (Sigma Aldrich) and synthesis lipid mixture were dissolved in chloroform and stored at -20°C. For liposome preparations, chloroform was evaporated under a nitrogen atmosphere for 15 min, and the samples were subjected to vacuum desiccation for 2 hrs at 60 mtorr. Dry films were hydrated in low salt buffer (25 mM Tris-HCl, pH 7.4 and 50 mM NaCl) to a concentration of 10 mg/ml.

To prepare large multi-lamellar vesicles (LMVs) via sonication methods, hydrated liposomes were sonicated for 15 min followed by 8 freeze-thaw cycles using liquid nitrogen. For most applications, liposomes were diluted to a concentration of 2 mg/ml in low salt buffer and were incubated in a heat block at 30°C for 1 hr before exposure to protein solutions. This method yields LMVs with an average diameter of 300-600 nm, and the occasional occurrence of larger vesicles (≥1 μm). For the preparation of large unilamellar vesicles (LUVs) via freeze and thaw method, hydrated liposomes were subjected to 6 freeze-thaw cycles followed by extrusion through filters with pore sizes 800 nm. LUV preparations were diluted to a concentration of 2 mg/ml, and used immediately after extrusion. The liposomes that are prepared from rapid solvent exchange method follow the procedures as previously described (Buboltz and Feigenson, 1999)

Liposome co-pelleting assay

Liposome samples (1 mg/ml) were incubated in the presence or absence of protein (10 μM) in 40 μl low salt buffer (corresponding to a 1:1 lipid-protein ratio) for 20 min at room temperature. Samples were centrifuged in a Optima MAX-E ultracentrifuge (Beckman) equipped with a TLA-100 rotor at 87,000 rpm at 4°C for 45 min. Pellets

and supernatants were separated, pellets were resuspended in 40 μl low salt buffer, and both fractions were analyzed by SDS-PAGE and Coomassie Blue staining.

Negative Staining Electron Microscopy (EM)

Liposome samples (2 mg/ml) were incubated in the presence or absence of protein (0.2-1 mg/ml or 4-20 μM) in low salt buffer for 1 min at 25°C, if not indicated otherwise. The sample (8 μl) was applied to a Carbon-formvar-coated copper grid (EMS) and incubated for 2 min. Excess liquid was carefully removed by using a wet absorbent tissue (Kimwipes; Kimberly Clark). The grids were incubated three times with 2% filtered uranyl acetate solution for 5 sec and air-dried. Negative staining was performed at 25°C (except for samples prepared at 4°C, which were stained in the cold room). Membrane morphologies were examined on a FEI Morgagni Transmission Electron Microscope and a FEI Tecnai G2 F20 TEM/STEM with the electron energy set to 80 kV and 200 kV, respectively. Representative images were taken on an AMT camera with a direct magnification of 11kx-36kx. Distance and size measurements were taken by using the program ImageJ.

RESULTS AND DICUSSIONS

Isoform-specific activity of Pacsin

Previously, we reported an SH3-dependent autoinhibiton mechanism for the brain-specific Pacsin isoform 1 *in vitro*. We extended the analysis to the other two human isoforms, Pacsin 1 and 2. The purified, isolated F-BAR domains or full-length proteins were incubated with large multilamellar vesicles (LMVs; freeze-thaw/sonication method; Folch fraction I lipids) and analyzed by negative-stain electron microscopy (EM). As observed before, the F-BAR domain of Pacsin 1 produces three distinct membrane morphologies under these conditions (Figure 3.1.B): wide tubes (open triangle), narrow tubes (dashed arrow) and pearling or beads-on-a-

string structures (solid arrows). Twisted, narrow tubes occur upon incubation with F-BAR domain-containing protein FCHo, which has been attributed to the observation that residues predicted to mediate membrane interactions do not line up perfectly with the concave surface of the F-BAR domain (Henne et al., 2010; Henne et al., 2007). This is also the case for Pacsin's F-BAR domain, and the pearling structures may represent an extreme case driven by the additional insertion of the wedge loop (Figure 3.1).

While the F-BAR domain of Pacsin 2 and 3 induced similar morphologies under these conditions, the autoinhibition of the full-length proteins appeared less pronounced or absent, respectively (Figure 3.1.B). A minor but noticeable difference in these micrographs is the absence of narrow tubes that have been observed with the F-BAR domain of Pacsin 1 and FCHo2. A potential explanation may involve variations in the degree of lateral curvature and/or flexibility of the distal tips within the F-BAR domain of Pacsin 1 and 2 (Figure 3.1.C), as also suggested by a recent crystallographic analysis of different Pacsin isoforms (Plomann et al., 2010). Considering the divergence in sequence conservation in the linker regions and the high degree of conservation of the F-BAR and SH3 domains, the nature of the linker segment may contribute to the different morphogenic potential and degree of autoregulation.

Autoinhibition and activation mechanism for Pacsin

Reduction of the membrane deformation activity of Pacsin 1 was also achieved when the isolated SH3 domain was added to the F-BAR domain in trans (Figure 2), supporting the proposed model in which the SH3 binds to the F-BAR domain and dampens its activity (Rao et al., 2010;Wang et al., 2009). These data are also in line with our structural analysis.

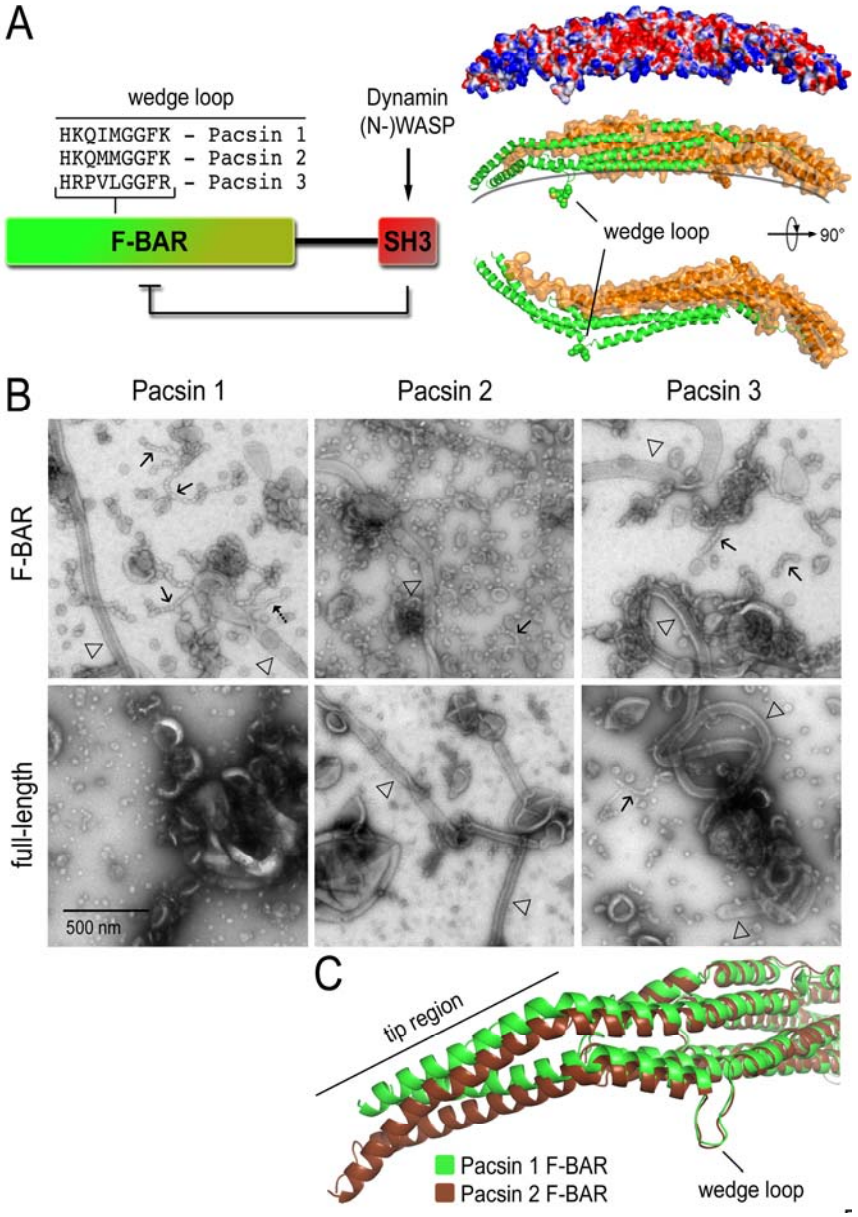


Figure 3.1. Pacsin isoform-specific membrane deformations. (A) Domain organization of human Pacsin isoforms. A schematic diagram of the three Pacsin isoforms is shown. The structure of the Pacsin F-BAR domain dimer. Human Pacsin 1 (Wang et al., 2009) is shown on the right panel.(B) Negative-stain EM of Pacsin-mediated membrane tubulation. 10uM F-BAR domains and full-length proteins of human Pacsin1,2,3 were incubated with 2mg/ml Folch Fraction I lipids. The liposomes are prepared through sonication+freeze/thaw method. The open arrows indicate featureless or striated wide tubes; Dashed arrow indicates thin tubes; Small arrows show beads-on-a-string structures. The micrographs show representative areas from the grids from three independent experiments. (C) Tip flexibility of PACSIN F-BAR domain (Wang et al., 2009)

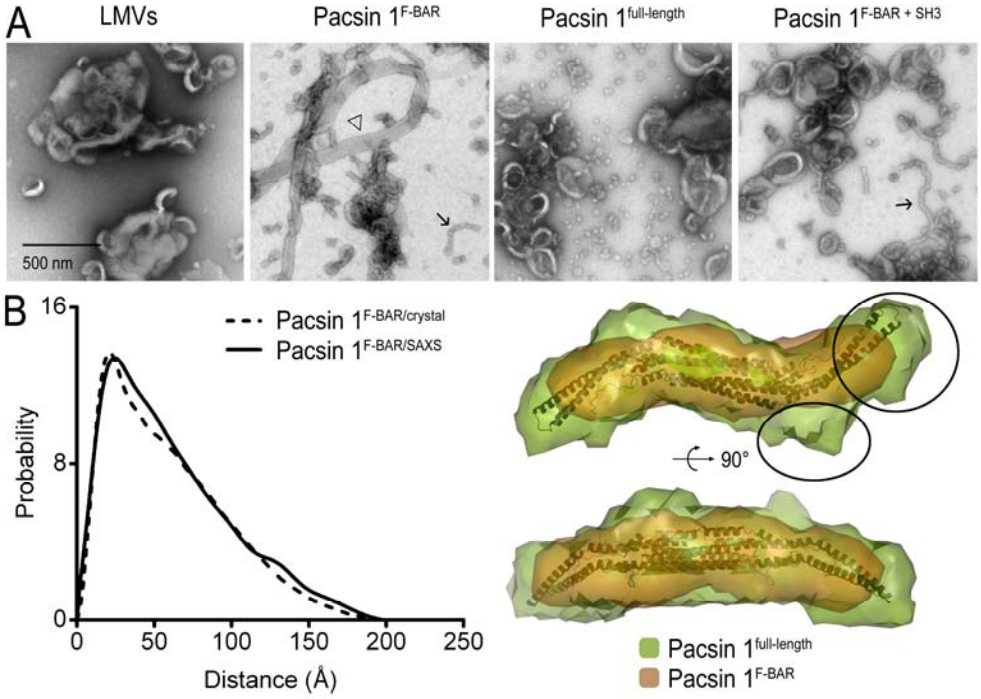


Figure. 3.2. Autoinhibitory role of the SH3 domain in Pacsin 1. (A) Negative-stain EM of SH3 domain inhibitory effect *in trans.* Dark arrow: thin tubes/pearling structures; Open arrow: wide tubes. (B) Small angle X-ray scattering study of PACSIN1$^{F\text{-}BAR}$ and PACSIN1$^{full\text{-}length}$. Left panel: distance distribution function of PACSIN1$^{F\text{-}BAR}$ in solution and in crystal. Right panel: Low resolution solution envelops of PACSIN$^{F\text{-}BAR}$ and PACSIN$^{full\text{-}length}$. the crystal structure is manually docked into the density via PyMol. Circles indicate the extra density in the full length protein that may account for the SH3 domain and linker region. The envelops are generated by ATSAS and visualized by PyMol.

We collected SAXS data on a monodispers sample of Pacsin's F-BAR domain, calculated its distance distribution or P(r) function, and compared it to the P(r) function calculated from the crystal structure of the F-BAR (Figure 3.2.B; Table 3.1). Furthermore, we modeled the SAXS data of the F-BAR domain and compared it to the reconstructions obtained from the full-length protein (Wang et al., 2009) (Figure3.2.B). These analyses of the scattering data revealed a similar radius of gyration, maximum diameter and shape of the isolated F-BAR domain in solution and in crystals. In solution, the shape and dimensions of the isolated F-BAR domain are similar to the

one of the full-length protein and the construct lacking the SH3 domain, which have been studied previously (Wang et al., 2009; Figure 3.2.B). Although the location of the SH3 domain could not be determined explicitly using this approach, it suggests a compact conformation of the full-length protein. A similar mechanism was recently revealed by the crystal structure of the mouse homolog of Pacsin 1 (Rao et al., 2010). One of the proposed positions for the SH3 domain based on the SAXS-based models (Figure 3.2.B; right circle), located close to the F-BAR domain tips, is in good agreement with the crystallographic analysis.

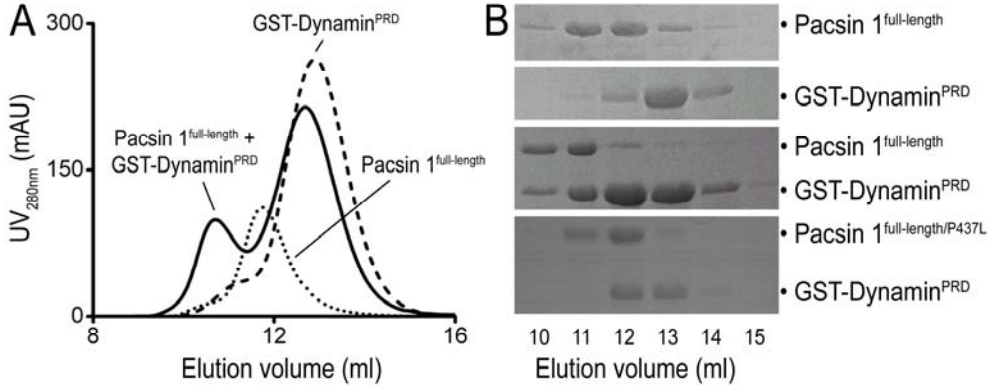


Figure 3.3. Pacsin1 and GST-dynamin^{PRD} forms stable complex *in vitro* (A) Pacsin1 and GST-dynamin^{PRD} co-migrate on the analytical size exclusion column. the binding buffer contains 50mM NaCl and 25mM Tris-HCl at pH 7.5. (B) SDS-PAGE of size exclusion fractions of individual Pacsin1, individual GST-dynamin^{PRD}, and the mixture of both. The comigration band shows up in the fraction 11, with roughly 1:1 pacsin1 to PRD ratio. The band is much weaker in the individual proteins or the mixture with SH3 domain and Pacsin interaction impaired.

The SH3 domain of Pacsin binds to the proline-rich domain (PRD) of proteins involved in endocytosis such as Dynamin, Synaptojanin and WASP/N-WASP (Qualmann et al., 1999; Simpson et al., 1999). A possible activation mechanism involves ligand binding and sequestration of the SH3 domains away from the F-BAR domain by the PRD. We tested this hypothesis in the liposome deformation assay.

Large multilamellar vesicles (LMVs) were incubated with human, full-length Pacsin 1 alone or in presence of the PRD of Dynamin (Dynamin^{PRD}) fused to GST. Pacsin1 and Dynamin^{PRD} domain forms stable complex on the size exclusion column (Figure 3.3). As a control, the isolated GST moiety was co-incubated with Pacsin 1 and LMVs. Full-length Pacsin 1 remained low in tubulation activity in the presence of GST, similar to results obtained with Pacsin 1 alone. Also, the isolated F-BAR domain was insensitive to the presence of GST or GST-Dynamin^{PRD}. In contrast, addition of GST-Dynamin^{PRD} to full-length Pacsin 1 resulted in the appearance of vesicular structures in the micrographs (Figure 3.4).

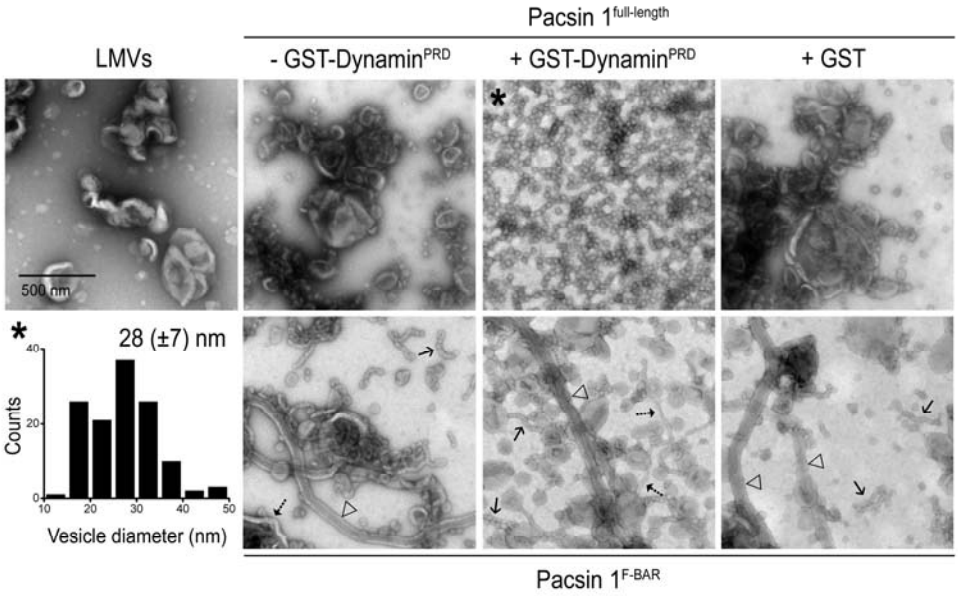


Figure 3.4. Membrane vesiculation mediated by Pacsin-Dynamin^{PRD} complex. Negative-stained EM of Pacsin-dymain^{PRD} treated liposome sample contains massive small homogenous vesicles. Vesicles have more defined shape and boundaries than these observed in the control samples, and much smaller than those generated by highly concentrated N-BARs (Peter et al., 2004). Mixing Pacsin$^{F\text{-}BAR}$ with Dynamin^{PRD} construct results in no apparent differences. Open arrows indicate the wide membrane tubes and black arrows indicate the highly curved membrane structures.

The morphology of the vesicles was homogeneous with an average diameter of 28 ± 7 nm (Figure 3.4), and distinct from the tubulation in cells and of liposomes

comprising only phosphatidylserine observed for mouse Pacsin 1-PRD complexes (Rao et al., 2010). Vesicle formation was protein concentration-dependent and unlikely resulting from over-titration, as no tubular structures was observed at any protein concentration under these conditions (Figure 3.5). Similar results were obtained when a Sumo moiety was used as the fusion partner for the Dynamin^{PRD} (Figure 3.6). Since GST has the propensity to form dimers, the Sumo moiety is believed to present the PRD as a monomeric ligand, indicating that a simple binding mode between the SH3 domain and the PRD is required for modulating Pacsin 1's membrane deformation activity.

Mutation of the central proline residue in the ligand-binding site of the SH3 domain, preventing PRD binding (Figure 3.6B), abolished the formation of small vesicles, and the liposomes appear significantly larger in diameter (~100nm). In addition, efficient vesiculation requires wedge-loop insertion, revealed by the reduced activity of M^{126}K mutant in full-length Pacsin 1 that retains autoregulation and near-wild-type membrane binding (Figure 3.6.B). By using the wedge-loop mutant we see tubular structures with large diameter at low frequency, but less and less homogeneous small vesicles have been observed (Figure 3.6.B).

Although Folch fraction I represents a natural and popular lipid extract, the uncertainty of its exact composition poses a reasonable concern. We repeated the core experiments described above using synthetic lipid mixtures with a defined composition similar in charge to the inner leaflet of the plasma membrane where Pacsin functions *in vivo* (45% POPS/27.5% POPC/27.5% POPE; freeze-thaw/sonication method). Incubation of LMVs generated with this mixture with the F-BAR domain of Pacsin 1 yielded similar deformation types as observed with Folch fraction I lipids, with the exception that tubular structures were observed less frequently and pearling structures comprised the majority in the micrographs.

Autoinhibition of Pacsin 1 and its activation by Dynamin^{PRD} (presented as a GST or Sumo fusion protein) could be reproduced in the defined lipid mixture.

Motivated by the large spectrum of membrane morphologies observed upon

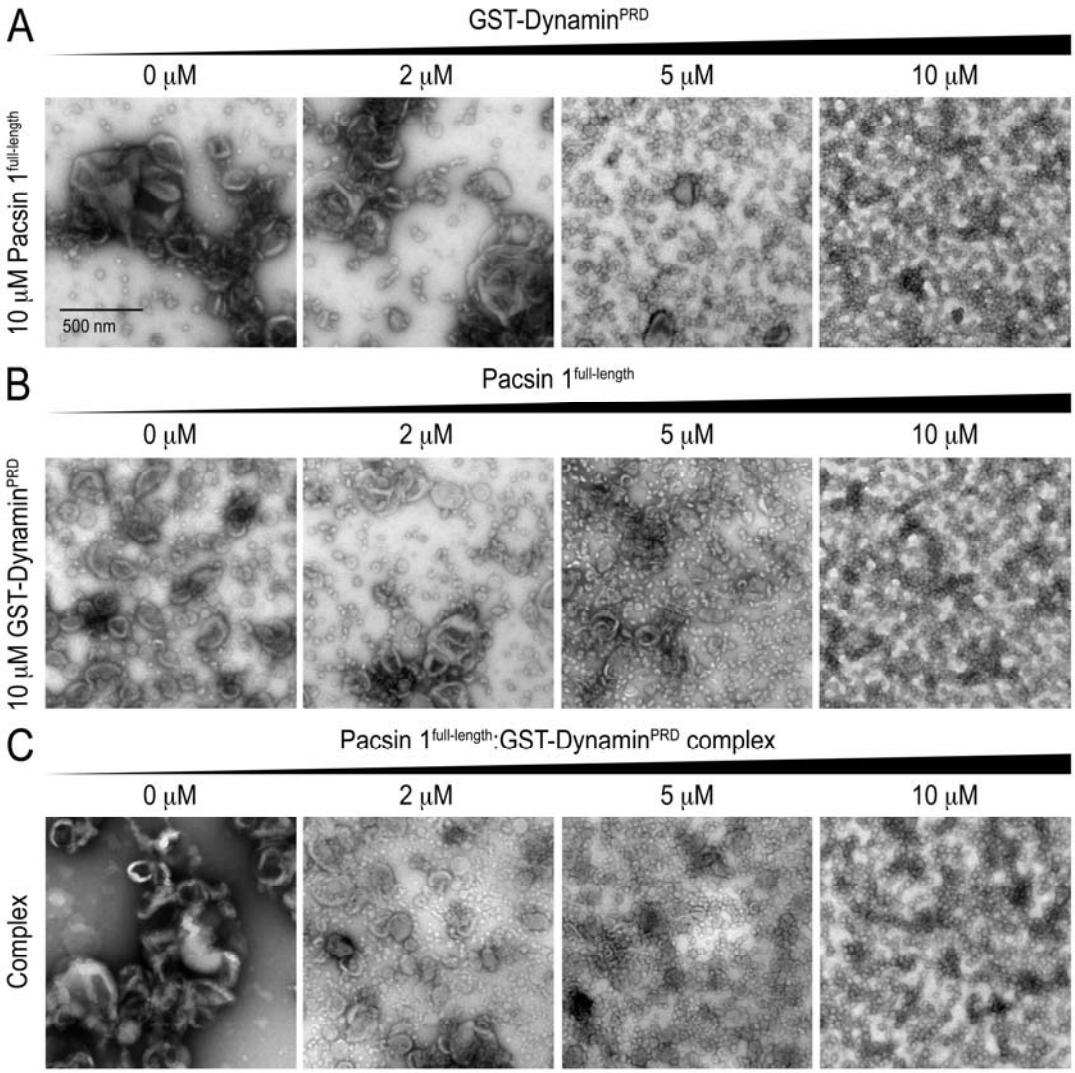


Figure 3.5. Concentration dependence of membrane vesiculation. Negative-stained EM of membrane vesiculation triggered by Pacsin-dynamin^{PRD} complex. (A) 10uM Pacsin1 mixed with 0-10uM Dynamin^{PRD}. Massive vesiculation happens at 5uM dynamin^{PRD} but not 2uM. (B) 10uM Dynamin^{PRD} with 0-2uM Pacsin1. the concentration dependent manner is similar to the one in (A). (C) Pacsin-Dynamin^{PRD} 1:1 mixture causes membrane vesiculation when concentration is higher than 2uM.

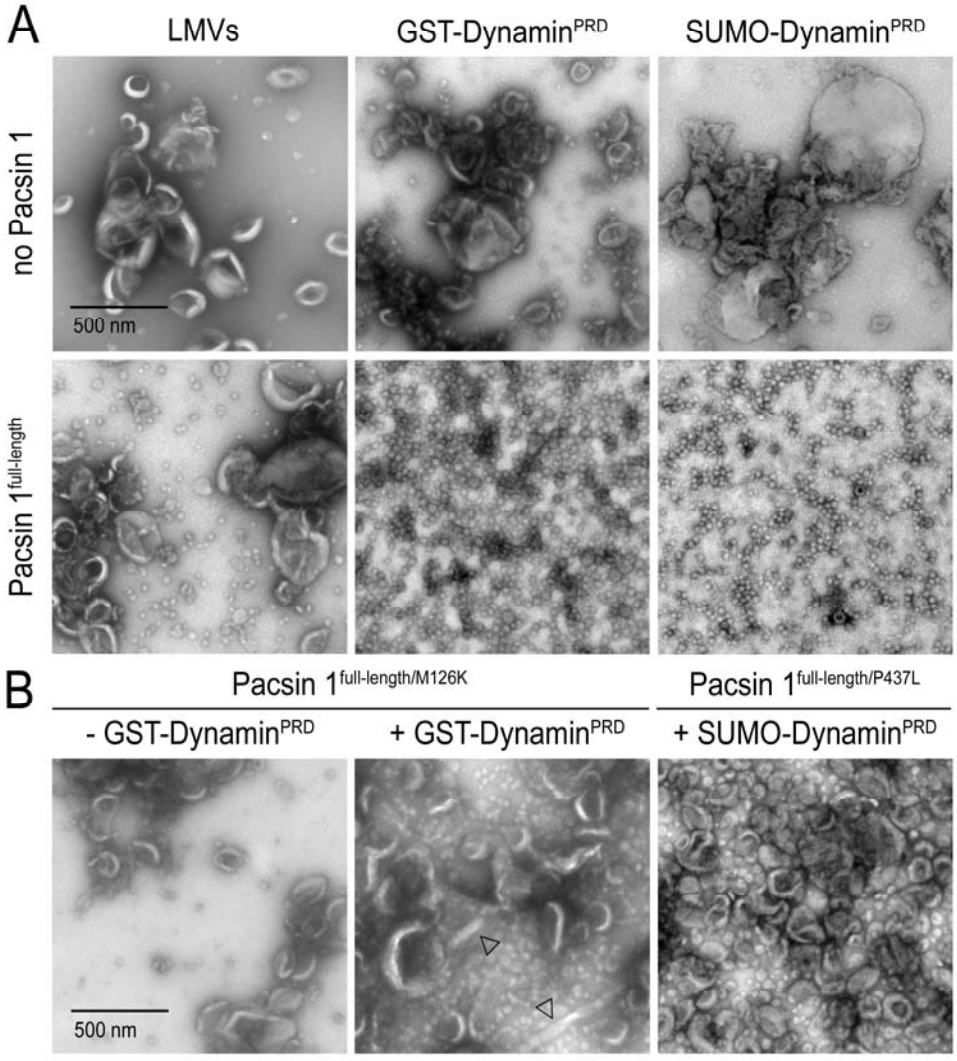


Figure 3.6. Pacsin1 wedge loop insertion and PRD-SH3 interaction are critical for membrane vesiculation. (A) Negative-stained EM of different Dynamin^{PRD} constructs treated liposomes with and without Pacsin1$^{full\text{-}length}$.Both GST tagged and SUMO tagged PRD cause membrane vesiculation when Pacsin1$^{full\text{-}length}$ is present. (B)Neither wedge loop insertion impared mutant M126K (Wang et al., 2009) nor PRD binding deficient mutant P^{435}L (Qualmann et al., 1999) causes membrane vesiculation at presence of dynamin^{PRD}

Pacsin incubation, ranging from vesicular structures to tubules of varying diameter (Rao et al., 2010; Shimada et al., 2010; Wang et al., 2009), we asked whether the state of the liposomes could account for the generation of different morphologies. employed three distinct liposome preparation methods. The protocol used thus far has been reported to generate LMVs. It involves the rehydration of a dried, thin lipid film in buffer followed by sonication and repeated freeze-thaw, which reduces the lamellarity and breaks up bigger aggregates, respectively. Prior to the addition of proteins, liposomes were incubated at 30°C to ensure that the membranes reach equilibrium. The second method is similar but instead of sonication, the liposomes are extruded through a filter with defined pore size. For liposome smaller than 200 nm in diameter, this method produces more unilamellar vesicle. The third method, rapid solvent exchange (RSE), provides a gentle way to exchange chloroform for an aqueous buffer, producing large liposomes with a lamellarity of approximately 1.5(Buboltz et al., 1999). By withdrawing chloroform in the presence of an aqueous buffer under vortexing, this method is also believed to avoid the formation of non-ideally mixed lipid bilayers that may arise from the generation of dry lipid film in the more common methods. Experiments were performed with both, Folch fraction I and a synthetic lipid mixture comprising 45% POPS, 27.5% POPC, and 27.5% POPE.

While many other factors may alter the morphogenic poteintial (e.g. lipid composition, protein/lipid ration, temperature, buffer conditions), we kept those variables constants with one exception. We noticed a sensitivity of the Pacsin 1 F-BAR domain to EDTA when RSE-derived liposomes were used, but not in the case of the freeze-thaw/sonication method. In the presence of the chelating agent, RSE-derived liposomes were deformed showing a spectrum of morphologies as described above, whereas predominantly wide-tubes were observed in the absence of EDTA

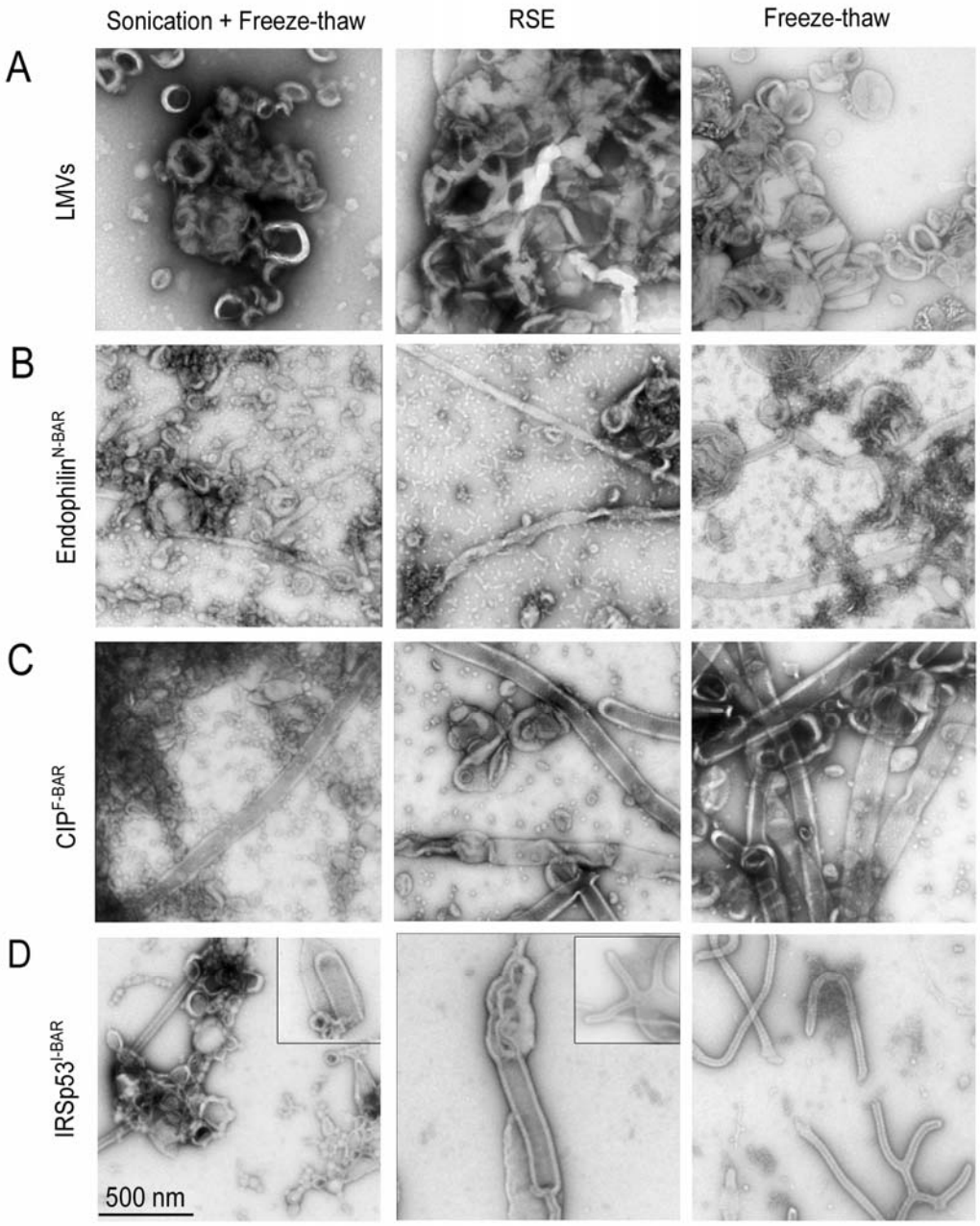


Figure 3.7. BAR domain mediated membrane tubulation depends on the liposome preparation methods. Negative stained EM of (A) Large Multilamanar Vesicles prepared by three different methods (B)deformed membrane states by Endophilin$^{N\text{-}BAR}$. (C) deformed membrane states by CIP4$^{F\text{-}BAR}$. (D) deformed membrane states by IRSp53$^{I\text{-}BAR}$.

The canonical F-BAR protein CIP4 generates massive amount of wide membrane tubes with the RSE and Freeze/Thaw liposome samples (Figure 3.7.C). The number of tubes in the sonicated samples is much less, although the tube morphology maintains the same. The different activities we observed for CIPF-BAR can be explained by the fact that it preferably interacts with larger lipid vesicles (Shibata et al., 2007), which is less populated in the sonicated vesicles. Similarly, Endophilin N-BAR domain generates much longer membrane tubes with the RSE and FT liposomes, and short tubes with the sonicated liposomes (Figure 3.7.B). Furthermore, the sample background is filled with small vesicles with heterogeneous sizes, which is consistent with the report that N-BAR domain causes liposome vesiculation at higher protein concentration (Peter et al., 2004). Unexpectedly, we observed both outward tubular structures and inward tubular structures in the IRSp53I-BAR domain treated samples, regardless the lipid preparation method(Figure 3.7.D). The fact that I-BAR domain can generate outward membrane tubes is much intriguing, since crystal structure shows that the it has convex surface positively charged which can only stabilize negative membrane curvature(Mallita et al., 2007).

The full-length PACSIN1 remains auto-inhibited in the all three liposome preparation methods, despite that both PACSINF-BAR and the full-length protein have promoted tubulation activity when treated with RSE and Freeze-Thaw liposomes (Figure 3.8.B and 3.8.C). The consistent observations of suppressed activity in full length protein suggest that the auto-inhibition is intrinsically encoded in the protein structures but not membrane. The promoted full-length protein activity is most likely due to the fact that it has stronger membrane interacting ability with larger vesicles in the RSE and Freeze-Thaw liposomes, which stabilize itself in the activated conformation and slightly releases the tubulation activity. Consistently, Pacsin1-Dynamin^{PRD} complex generate large amount of small homogenous lipid vesicles in

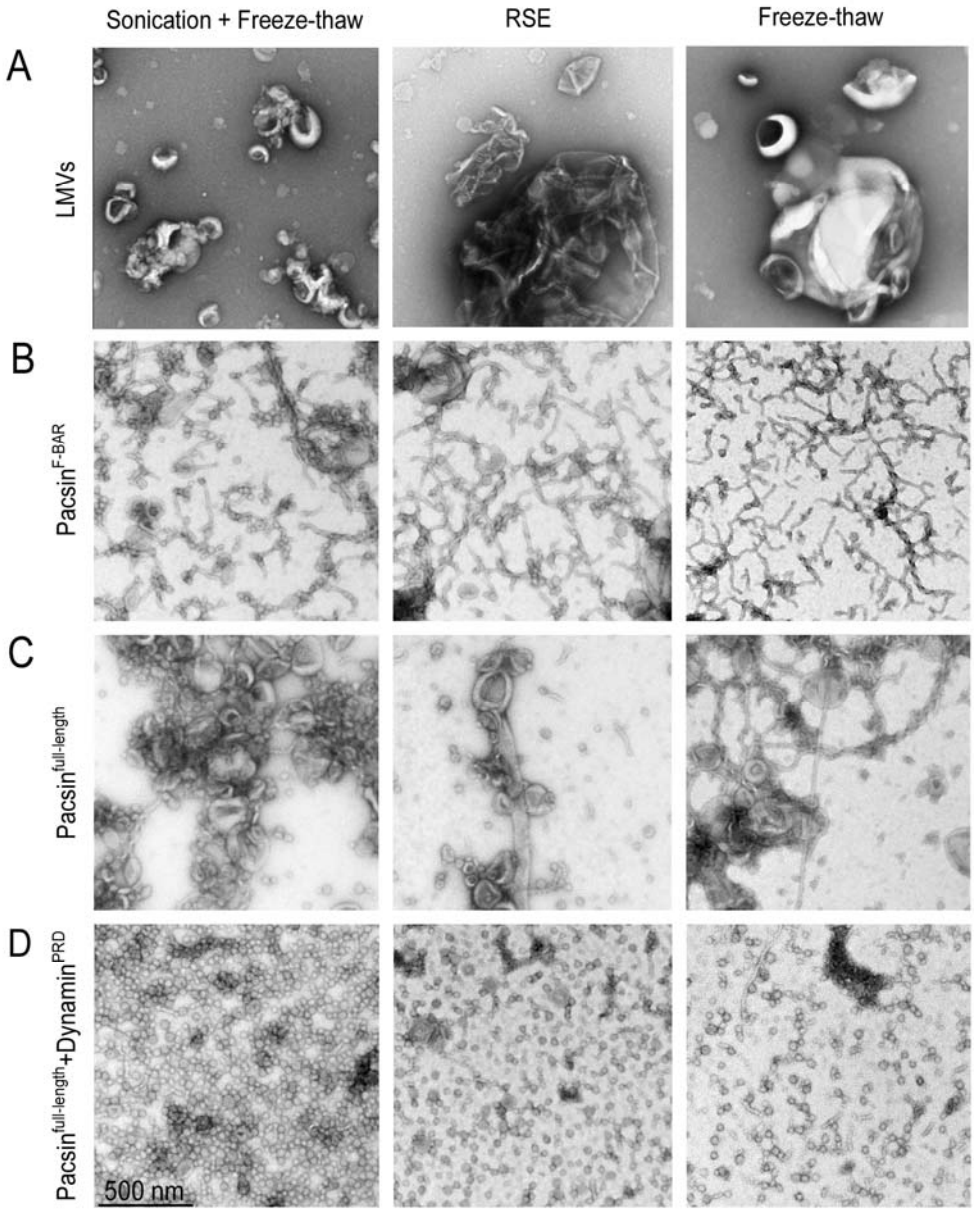


Figure 3.8. Robust auto-inhibition and activation of PACSIN1. Negative stained EM of (A) Large Multilamanar Vesicles prepared by three different methods (B)deformed membrane states by Pacsin$^{F\text{-}BAR}$. (C) deformed membrane states by Pacsin$^{full\text{-}length}$. (D) deformed membrane states by Pacsin$^{full\text{-}length}$-Dynamin^{PRD} complex.

all three liposome samples. This suggests that the unique vesiculation activity is not due to the specific membrane properties but the special protein function. The

molecular mechanism of Pacsin-Dynamin^{PRD} remains elusive. It is possible that the positively charged PRD domain forms a continuous membrane interacting surface with F-BAR that can provide extra energy required for vesiculation.

The effect of scaffolding in Pacsin-mediated membrane deformation

It has been shown that N-BAR-mediated tubulation is predominantly driven the insertion of hydrophobic motifs into the acyl chain layer of one membrane leaflet (Jao et al., 2010; Gallop et al., 2006; Masuda et al., 2006). Based on theoretical estimations, the shape of the protein and electrostatic interactions between protein and membrane provide a minor component to curvature generation (Campelo et al., 2008; Campelo et al., 2010). We applied similar calculation to Pacsin's F-BAR domain to assess the relative contributions of scaffolding, driven by shape and charge complementarily, and wedge loop insertion (Campelo et al., 2008; Campelo et al., 2010).

In the scaffolding mechanism, the bending energy originates from direct protein-lipid headgroup interactions, which usually involves 20-25 interaction pairs (Frost et al., 2008). The interactions that maintain the protein lattices on the membrane are minor, limited to 1-2 hydrogen bonds per molecule (Frost et al., 2008; Shimata et al., 2007; Wang et al., 2009). According to our estimations, the maximum energy density due to peripheral lipid binding of the F-BAR domain of Pacsin (and similar domains) is approximately $0.13\,k_BT$. In comparison, generation of a tube with a radius R=10 nm, which corresponds to an energy density of approximately $0.125k_BT$, would require full coverage of the membrane surface with F-BAR domains, if scaffolding were the main driving force in membrane deformation. In addition, it appears unlikely that the more curved, vesicular structures can be generated solely by the scaffolding mechanism.

Effect of hydrophobic insertions into the a bilayer leaflet

Experimentally, the membrane deformation potential of the Pacsin 1 F-BAR domain is sensitive to mutations in the amphipathic wedge loop and ionic strength of the solvent suggesting insertion of the loop into one leaflet of the lipid bilayer (Wang et al., 2009). In order to assess the contribution of hydrophobic insertion on Pacsin-mediated membrane sculpting, we estimated the wedge loop bending potential, taking into account the coupling of the two membrane leaflets (Campelo et al., 2008). The quantity ε describes the ratio between the surface of the outer and inner leaflet upon membrane bending. Insertion of hydrophobic motifs such as amphipathic helices or loops will counteract the initial surface is mismatch. By comparing ε and the surface area of these inserted motifs relative to the total membrane interaction surface of the protein scaffold will provide an estimate with regard to the contribution of the wedge loop insertion to the bending process.

The N-BAR domains from Endophilin and Amphiphysin can effectively convert flat lipid bilayers into tubules with a radius R=20-25 nm *in vitro*. Based on the crystal structures of the two, the insertion motifs of the N-BAR domains from Endophilin and Amphyphysin account for approximately 50% and 25% (0.5 and 0.25, expressed as fractions) of the total membrane interaction interface, respectively (Gallop et al., 2006; Peter et al., 2004; Campelo et al., 2008; Weissenhorn, 2005). Both values are higher than the estimated excess surface ratio (ε~0.17-0.22) that would initially be generated by the bending of a flat membrane to a tube with a radius R=20-25 nm. This implies that the hydrophobic units in N-BAR domains are sufficient to create or counteract the surface area mismatch in a bent bilayer, which is consistent with the conclusions that have been drawn from the elastic model of laterally coupled monolayers (Campelo et al., 2008).

Unlike the amphipathic helices in Amphiphysin that span a total area of 12 nm^2 (25% of the total membrane interaction area), the two wedge loops in the dimeric F-BAR domain of Pacsin span only an area of 1.5 nm^2. Assuming the extreme case of 100% membrane coverage with the F-BAR domain, the highest surface occupancy of these insertion units is only 5%-7%. This number is far below the expected ε=0.5 for a tube with radius R=10 nm. Similar conclusion can also be obtained based on the elastic model of membrane monolayers. On an uncoupled monolayer, 10-12% of the membrane surface must be occupied by the insertion motif in order to generate a curvature with R=10 nm. Based on these estimations, it is unlikely that Pacsin's potential of generating highly curved membrane surface is solely driven by the insertion mechanism.

Coupled scaffolding-insertion mechanism

Based on the above analysis, we conclude that the F-BAR domain of Pacsin employs both, the scaffolding and hydrophobic insertion mechanisms to deform membranes. The two mechanisms are tightly coupled in the stabilization of highly curved membranes. While the relative degree of the contributions requires further elucidation, the scaffolding mechanism may play multiple roles. In addition to providing a large portion of the membrane bending and binding free energy, it also facilitates the insertion of the hydrophobic wedge loop contributing the energy, which is ultimately required to create a high degree of curvature. These results are consistent with experimental data describing the salt sensitivity of Pacsin-membrane interactions (Wang et al., 2009). Furthermore, both mechanisms may contribute distinctly to the deformation of membranes by providing a driving force for counteracting the bilayer mismatch between the two leaflets that would arise upon curvature generation. In addition, the overall protein shape and charge distribution will determine the types of

membrane morphologies Pacsin's F-BAR domain prefers to interact with or to stabilize (Wang et al., 2009).

Protein-mediated lipid demixing and bimodality in membrane deformation.

Instigated by the apparent inconsistencies observed for Pacsin-mediated membrane sculpting by us and others (Rao et al., 2010; Shimada et al., 2010; Wang et al.,2009), we provide a possible explanation based on energetic considerations. The system's total free energy is a summation of the protein-membrane interaction energy, and the internal energies of the protein and bilayer:

$$E_{total} = I_{protein} + I_{bilayer} + E_{protein-bilayer}$$

We first assume that within the scope of elastic Gaussian theory, both internal energy terms are constant. As a consequence, the membrane-protein interaction represents the major energy term in the system. From plotting the radii of tubes and vesicles against the energy density required for protein-induced membrane deformations, it is obvious that there are energetically equivalent structures that could potentially coexist. In this simplistic scenario, a cylindrical structure with a radius R and a spherical structure with a radius of ~$2R$ represent systems with theoretically similar free energy. It is worth noting, that both structures also have similar mean curvature. These estimations suggest that the system can be subject to bimodality, producing either narrow membrane tubules or vesicular structures, consistent with our experimental observations. Bimodality, in contrast to bistability, does not assume an identical origin. Indeed, we were able to show modulation of the deformation morphologies by preparing liposomes following different protocols suggesting that the state of the starting material, which is likely to be a heterogeneous mixture with regard to liposome size and lamellarity, may determine the preferred structure stabilized by Pacsin. Differences in the frequency in which tubules and vesicles occur may indicate

differences of the membrane properties of the initial liposomes from which these structures arise.

Much of the current analysis assumes that the lipid composition and properties remain locally unchanged during the membrane deformation process. Yet, the presentation of a positively charged electrostatic surface, for example the concave side of a BAR domain, to an ideally mixed bilayer comprising charged and neutral lipids will promote the rearrangements of lipids (Khelashvili et al., 2009; Khelashvili et al., 2008; Mbamala et al., 2005; May et al., 2000; Heimburg et al., 1999). The reorganization within the lipid bilayer is driven through both attractive as well repulsive forces. It is conceivable that lipid demixing alters membrane behavior, which ultimately may influence the membrane morphologies and their frequency of occurrence upon F-BAR domain binding. Vice versa, non-ideally mixed bilayers due to macroscopic and nanoscopic phase separation may react differently to interactions with the protein, and the different liposome preprations methods used here may differ in this parameter.

Different regimes of mixing may have distinct tendencies towards a particular shape. In addition, inter-lipid interaction energy may no longer be negligible when the membrane radius becomes as small as several folds of the membrane thickness. Furthermore, the nature of the lipid headgroup and acyl chains may confer preferred partition propensities to either spherical or cylindrical bilayers. Although a potential F-BAR domain-mediated effect on lipid phase behavior awaits experimental validation, it provides a theoretical rationalization for the observed bimodality in Pacsin-mediated membrane deformation, namely the formation of tubes and vesicular structure from the same liposome preparations. Based on the *in vitro* experiments described here, the state of the input material, the liposomes, has a significant influence on the membrane deformation potential and membrane morphology. In addition, the ratio between

protein and lipid, ultimately determining the degree of protein coverage on the membranes, also contributes to the different morphogenic potential of Pacsin's F-BAR domain.

CONCLUSION

It is worth noting that canonical N-BAR and F-BAR domain comprising only one principle curvature predominantly produce tubular membrane morphologies that roughly correlate with the diameter of the concave surface of the protein scaffold (Farsad et al., 2001; Frost et al., 2008; Itoh et al., 2006; Peter et al., 2004; Shimada et al., 2007; Takei et al., 1999). Structural variation in form of bent tips of dimeric F-BAR domains and the distribution of charged residues that do not exactly coincide with the concave surface appear to introduce additional membrane morphologies such as vesicular and twisted structures, tubes with varying diameter and even invaginations (Guerrier et al., 2009; Henne et al., 2010; Henne et al., 2007; Shimada et al., 2010; Wang et al., 2009). We could show that different preparation methods for liposomes using identical lipid mixtures can produce different protein-mediated morphologies, which suggests that the properties of the input liposomes influence the tubulation activity of F-BAR proteins.). The reorganization within the lipid bilayer is driven through both attractive as well repulsive forces. It is conceivable that lipid demixing alters membrane behavior, which ultimately may influence the membrane morphologies and their frequency of occurrence upon F-BAR domain binding. Vice versa, non-ideally mixed bilayers due to macroscopic and nanoscopic phase separation may react differently to interactions with the protein, and the different liposome preprations methods used here may differ in this parameter. Taken together, both intrinsic protein structure and membrane bilayer properties contribute to the protein mediated membrane deformation process. It is still unclear what specific aspects of

membranes influence the output of membrane deformation. The nature parameters such as membrane composition, bending rigidity, temperature and phases all contribute to this process. More comprehensive characterization of membranes under differenet preparation methods can provide further insights in understanding this complicated process.

REFERENCES

Andersson F., Jakobsson J., Löw P., Shupliakov O., Brodin L. (2008) Perturbation of syndapin/PACSIN impairs synaptic vesicle recycling evoked by intense stimulation. J. Neurosci. 28:3925-3933.

Anggono V., et al., (2006) Syndapin I is the phosphorylation-regulated dynamin I partner in synaptic vesicle endocytosis. Nat. Neurosci. 9:752-760.

Campelo F., Fabrikant G., McMahon H.T., Kozlov M.M. (2010) Modeling membrane shaping by proteins: focus on EHD2 and N-BAR domains. FEBS Lett. 584(9):1830-9.

Campelo F., McMahon H.T., Kozlov M.M 2008) The hydrophobic insertion mechanism of membrane curvature generation by proteins. Biophys. J. 95(5):2325-39.

Clayton E.L., Anggono V., Smillie K.J., Chau N., Robinson P.J., Cousin M.A. (2009) The phospho-dependent dynamin-syndapin interaction triggers activity-dependent bulk endocytosis of synaptic vesicles. J. Neurosci. 29(24):7706-17.

Dharmalingam E., Haeckel A., Pinyol R., Schwintzer L., Koch D., Kessels M.M., Qualmann B.(2009) F-BAR proteins of the syndapin family shape the plasma membrane and are crucial for neuromorphogenesis. J. Neurosci. 29(42):13315-27.

DiProspero N.A., et al., (2004) Early changes in Huntington's disease patient brains involve alterations in cytoskeletal and synaptic elements. J. Neurocytol. 33:517-533.

Doherty G.J., McMahon H.T. (2009) Mechanisms of endocytosis. Annu. Rev. Biochem. 78:857-902.

Edeling M.A., Sanker S., Shima T., Umasankar P.K., Höning S., Kim H.Y., Davidson L.A., Watkins S.C., Tsang M., Owen D.J., Traub L.M. (2010) Structural requirements for PACSIN/Syndapin operation during zebrafish embryonic notochord development. PLoS One;4(12):e8150.

Farsad K., Ringstad N., Takei K., Floyd S.R., Rose K., De Camilli P. (2001) Generation of high curvature membranes mediated by direct endophilin bilayer interactions. J. Cell Biol. 155: 193–200

Frost A., Perera R., Roux A., Spasov K., Destaing O., Egelman E.H., De Camilli P., Unger V.M.(2008) Structural basis of membrane invagination by F-BAR domains. Cell. 2008 132(5):807-17.

Guerrier S., Coutinho-Budd J., Sassa T., Gresset A., Jordan N.V., Chen K., Jin W.L., Frost A., Polleux F. (2009) The F-BAR domain of srGAP2 induces membrane protrusions required for neuronal migration and morphogenesis. Cell. 138(5):990-1004.

Itoh T., Erdmann K.S., Roux A., Habermann B., Werner H., De Camilli P. (2006) Dynamin and the actin cytoskeleton cooperatively regulate plasma membrane invagination by BAR and F-BAR proteins. Dev.Cell. 9(6):791-804.

Jao C.C., Hegde B.G., Gallop J.L., Hegde P.B., McMahon H.T., Haworth I.S., Langen R. (2010) Roles of amphipathic helices and the BAR domain of endophilin in membrane curvature generation. J. Biol. Chem. 2010 Apr 23. Paper in press

Khelashvili G., Harries D., Weinstein H. (2009) Modeling membrane deformations and lipid demixing upon protein-membrane interaction: the BAR dimer adsorption. Biophys. J. 97(6):1626-35.

Khelashvili G., Weinstein H., Harries D. (2008) Protein diffusion on charged membranes: a dynamic mean-field model describes time evolution and lipid reorganization. Biophys. J. 94(7):2580-97.

Kumar V., Fricke R., Bhar D., Reddy-Alla S., Krishnan K.S., Bogdan S., Ramaswami M. (2009) Syndapin promotes formation of a postsynaptic membrane system in Drosophila Mol. Biol. Cell. 2009 20(8):2254-64.

Marks B., Stowell M.H., Vallis Y., Mills I.G., Gibson A., Hopkins C.R., McMahon H.T. (2001). GTPase activity of dynamin and resulting conformation change are essential for endocytosis. Nature. 410(6825):231-5.

Mattila P.K., Pykäläinen A., Saarikangas J., Paavilainen V.O., Vihinen H., Jokitalo E., Lappalainen P. (2007) Missing-in-metastasis and IRSp53 deform PI(4,5)P2-rich membranes by an inverse BAR domain-like mechanism. J. Cell Biol. 176(7):953-64.

McMahon H.T., Kozlov M.M., Martens S. (2010) Membrane curvature in synaptic vesicle fusion and beyond. Cell. 140(5):601-5.

Mettlen M., Pucadyil T., Ramachandran R., Schmid S.L. (2009) Dissecting dynamin's role in clathrin-mediated endocytosis. Biochem. Soc. Trans. 37(Pt 5):1022-6.

Merzlyak E.M. et al., Bright far-red fluorescent protein for whole-body imaging. 2007 Nat. Methods. 4(9):741-6.

Miyawaki A., Llopis J., Heim R., McCaffery J.M., Adams J.A., Ikurak M., Tsien R.Y. (1997). Fluorescent indicators for Ca2+ based on green fluorescent proteins and Calmodulin. Nature. 388 (6645): 882–7

Miyawaki A., Tsien R.Y. (2000) Monitoring protein conformations and interactions by fluorescence resonance energy transfer between mutants of green fluorescent protein. Methods Enzymol. 327: 472-500

Mizuno N., Jao C.C., Langen R., Steven A.C. (2010) Multiple modes of endophilin-mediated conversion of lipid vesicles into coated tubes: implications for synaptic endocytosis. J. Biol. Chem. 2010 May 18. Paper in press.

Modregger J., Ritter B., Witter B., Paulsson M., Plomann M. (2000) All three PACSIN isoforms bind to endocytic proteins and inhibit endocytosis. J. Cell. Sci. 113:4511-4521.

Modregger J., DiProspero N.A., Charles V., Tagle D.A., Plomann M. (2002) PACSIN 1 interacts with huntingtin and is absent from synaptic varicosities in presymptomatic Huntington's disease brains. Hum. Mol. Genet. 11:2547-2558.

McMahon H.T., Kozlov M.M., Martens S. (2010) Membrane curvature in synaptic vesicle fusion and beyond. Cell. 140(5):601-5.

Plomann M., Wittmann J.G., Rudolph M.G. (2010) A hinge in the distal end of the PACSIN 2 F-BAR domain may contribute to membrane-curvature sensing. J. Mol. Biol. 400(2):129-36.

Qualmann B., Kelly R.B. (2000) Syndapin isoforms participate in receptor-mediated endocytosis and actin organization. J. Cell Biol. 148(5):1047-1062.

Qualmann B., Roos J, DiGregorio PJ, Kelly RB (1999) Syndapin I, a synaptic dynamin-binding protein that associates with the neural Wiskott-Aldrich syndrome protein. Mol. Biol. Cell 10:501-513.

Rao Y., Ma Q., Vahedi-Faridi A., Sundborger A., Pechstein A., Puchkov D., Luo L., Shupliakov O., Saenger W., Haucke V. (2010) Molecular basis for SH3 domain regulation of F-BAR-mediated membrane deformation. Proc. Natl. Acad. Sci. USA 107(18):8213-8

Shimada A., et al. (2007) Curved EFC/F-BAR-domain dimers are joined end to end into a filament for membrane invagination in endocytosis. Cell 129:761-772.

Shimada A., Takano K., Shirouzu M., Hanawa-Suetsugu K., Terada T., Toyooka K., Umehara T., Yamamoto M., Yokoyama S., Suetsugu S. (2010) Mapping of the basic amino-acid residues responsible for tubulation and cellular protrusion by the EFC/F-BAR domain of pacsin2/Syndapin II. FEBS Lett. 19;584(6):1111-8.

Simpson F., Hussain N.K., Qualmann B., Kelly R.B., Kay B.K., McPherson P.S., Schmid S.L. (1999) SH3-domain-containing proteins function at distinct steps in clathrin-coated vesicle formation. Nat. Cell Biol. 1: 119–124

Takei K., Slepnev V.I., Haucke V., De Camilli P. (1999) Functional partnership between amphiphysin and dynamin in clathrin-mediated endocytosis.Nat. Cell Biol. 1(1):33-9.

Yarar D., Surka M.C., Leonard M.C., Schmid S.L. (2007a) SNX9 activities are regulated by multiple phosphoinositides through both PX and BAR domains. Traffic 12(1):133-46..

Yarar D., Waterman-Storer C.M., Schmid S.L. (2007b) SNX9 couples actin assembly to phosphoinositide signals and is required for membrane remodeling during endocytosis. Dev. Cell. 13(1):43-56.

Wang Q., Kaan H.Y., Hooda R.N., Goh S.L., Sondermann H. (2008) Structure and plasticity of Endophilin and Sorting Nexin 9. *Structure* 16:1574-87.

Wang Q., Navarro M.N.S., Peng G., Goh S.L., Judson B.T., Sondermann H. (2009) Molecular mechanism of membrane tabulation and restriction by F-BAR domain protein pacsin. (2009) Proc. Nat. Acad. Sci. 123(4):124-133

Weissenhorn W (2005) Crystal structure of the endophilin-A1 BAR domain. *J. Mol. Biol.* 351:653-661.

CHAPTER 4

STRUCTURE AND PLASTICITY OF SORTING NEXIN 9 AND ENDOPHILIN

ABSTRACT

Endophilin and Sorting Nexin 9 (Snx9) play key roles in endocytosis by membrane curvature sensing and remodeling via their Bin/Amphiphysin/Rvs (BAR) domains. BAR and the related F-BAR domains form dimeric, crescent-shaped units that occur N- or C-terminally to other lipid binding, adaptor or catalytic modules. In crystal structures, the PX-BAR unit of Snx9 ($Snx9^{PX\text{-}BAR}$) adopts an overall compact, moderately curved conformation. SAXS-based solution studies revealed an alternative, more curved state of $Snx9^{PX\text{-}BAR}$ in which the PX domains are flexibly connected to the BAR domains, providing a model for how Snx9 exhibits lipid-dependent curvature preferences. In contrast, Endophilin appears to be rigid in solution with the SH3 domains being located at the distal tips of a BAR domain dimer with fixed curvature. We also observed tip-to-tip interactions between the BAR domains in a trigonal crystal form of $Snx9^{PX\text{-}BAR}$ reminiscent of functionally important interactions described for F-BAR domains.

INTRODUCTION

BAR domain-containing proteins are implicated in endo- and exocytosis regulating membrane trafficking and cytoskeletal rearrangements (Itoh and De Camilli, 2006). BAR domains consist of a three-helix bundle that dimerizes forming a crescent shape with surface properties that allow them to sense and induce membrane curvature (Peter et al., 2004; Figure 4.1). There are structural variations within the BAR domain superfamily, that contribute to distinct membrane curvature preferences for each subfamily: N-BAR proteins such as Endophilin and Amphiphysin contain a N-terminal amphipathic helix that contributes to their membrane bending potential (Gallop et al., 2006; Masuda et al., 2006; Zimmerberg and Kozlov, 2006). F-BAR proteins form a distantly related sub-family, inducing or stabilizing membrane deformations with lower degree of curvature distinct from N-BAR domain proteins (Henne et al., 2007; Shimada et al., 2007; Frost et al., 2008). The recently characterized I-BAR domains form rather flat dimers and are thought to stabilize negative curvatures (Lee et al., 2007; Mattila et al., 2007; Millard et al., 2005).

From the available crystal structures, there appears to be a preferred degree of curvature within particular BAR domains with APPL1 being more curved than Endophilin and Amphiphysin (Figure 4.1) (Gallop et al., 2006; Li et al., 2007; Masuda et al., 2006; Pylypenko et al., 2007; Zhu et al., 2007). F-BAR domains are significantly larger in diameter and have a lower degree of curvature and their preference for larger vesicles can be rationalized on a structural basis (Frost et al., 2008). The crystal structure of the membrane binding module of Snx9, a Phox-homology (PX) and BAR domain-containing Sorting Nexin, depicts a less curved BAR domain dimer compared to other BAR domain structures, consistent with the induction of wider membrane tubes by Snx9 in tubulation assays (Pylypenko et al., 2007; Yarar et al., 2008). Interestingly, Snx9 was also shown to prefer more curved

vesicles in the absence of phosphatidylinositol phosphates (PIPs) but displaying affinity for a broad range of membrane curvatures in the presence of the appropriate phospholipids binding to the PX domain (Pylypenko et al., 2007). Such a phenomenon could suggest plasticity within this domain, yet little is known about the structural flexibility within BAR domain dimers.

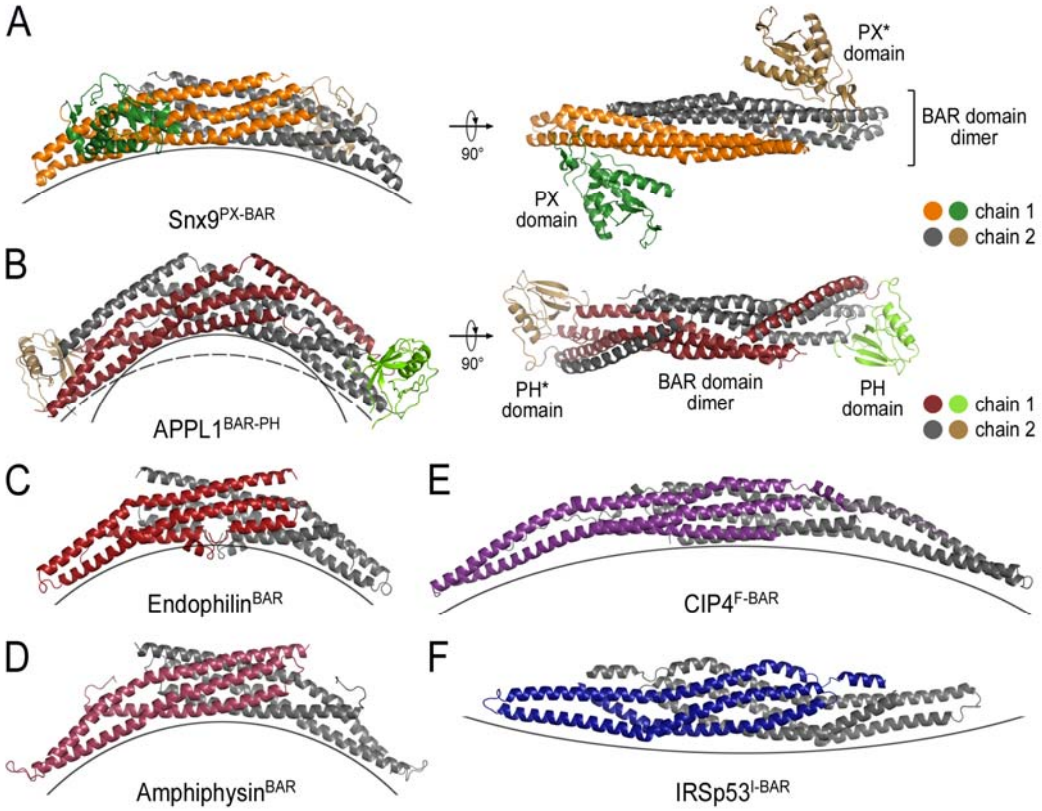


Figure 4.1. Structural comparison of BAR domain and BAR domain-related proteins. Crystal structures of Snx9$^{PX\text{-}BAR}$ (A) (Pylypenko et al., 2007; this study) and of the BAR-PH module of APPL1 (B) (PDB: 2ELB, Li et al., 2007; Zhu et al., 2007) are shown in two orthogonal views in their dimeric state. Dimeric crystal structures of isolated N-BAR domains from Endophilin (C) (PDB: 1X03, Masuda et al., 2006) and Amphiphysin (D) (PDB: 1URU, Peter et al., 2004), the F-BAR domain from CIP4 (E) (PDB: 3EFK, Shimada et al., 2007) and the I-BAR domain of IRSp53 (F) (PDB: 1Y2O, Millard et al., 2005) are shown. The degree of curvature of BAR domain dimers is indicated as grey lines.

Sorting nexins form a family of peripheral membrane proteins characterized by a conserved Phox-homology (PX) domain, and are involved in intracellular membrane

trafficking (Carlton and Cullen, 2005; Worby and Dixon, 2002). Similarly to endophilin and amphiphysin, Snx9 (also known as SH3PX1) contains an N-terminal SH3 domain with affinity for dynamin, synaptojanin, and N-WASP providing a physical link to vesicle scission and endocytosis (Badour et al., 2007; Ringstad et al., 1997; Shin et al., 2007; Yarar et al., 2007). The SH3 domain also binds to the non-receptor activated Cdc42-associated kinase (ACK) associating Snx9 with the degradation of epidermal growth factor receptors (Lin et al., 2002; Yeow-Fong et al., 2005). Receptor down-regulation requires dimerization of the C-terminal BAR domain and ACK-mediated phosphorylation of Snx9 (Childress et al., 2006). The SH3 domain connects to the C-terminal PX-BAR module via a low complexity region that mediates binding to clathrin light chains and the appendage domain of AP-2, linking Snx9 to clathrin-mediated endocytosis (Lundmark and Carlsson, 2002, 2003; Soulet et al., 2005).

In Snx9 (and related proteins Snx18 and Snx30), the PX domain is located adjacent to the BAR domain and displays broad specificity for mono-, doubly-, or triply phosphorylated PIPs, namely PI(3)P, PI(4,5)P2, PI(3,4)P2, PI(3,5)P2, PI(3,4,5)P3 (Pylypenko et al., 2007; Yarar et al., 2008) (Figure 4.1.A). In the crystal structure, the PX domains form lobes that interact with the lateral surface close to the distal tips of the BAR domains (Pylypenko et al., 2007) (Figure 4.1.A). The domains form a continuous surface for membrane interactions spanning the concave side of the BAR domains and an area extending from the canonical PIP binding site. An additional motif, the conserved yoke domain, connects the BAR and PX domains contributing to the inter-domain packing. The relative domain organization in this structure is different than that observed for the two-domain membrane-interacting module of the Rab-effector APPL1 (Li et al., 2007; Zhu et al., 2007) (Figure 4.1.B). The BAR and Pleckstrin homology (PH) domains in APPL1 are connected via a long

helical linker that packs against the three-helix bundle of the adjacent BAR domain placing the PH domains of one protomer at the distal tip of the BAR domain from its dimerization partner, displaying a continuous membrane interface (Li et al., 2007). In addition, a predicted amphipathic helix in Snx9, located N-terminally to the PX and yoke domain, was shown to be functionally homologous to the N-terminal helix in N-BAR domains crucial for generating membrane tubules (Pylypenko et al., 2007).

As a prerequisite for its membrane remodeling activity, Snx9 has to undergo an oligomerization switch mediated by PI(4,5)P2-containing membranes producing a complex that stimulates actin polymerization (Yarar et al., 2007). Oligomerization has also been described as a major driving force for membrane tubulation by F-BAR domains (Frost et al., 2008; Henne et al., 2007; Shimada et al., 2007). As evidenced by electron microcopy, F-BAR domains can form lattices on model membranes, mediated by tip-to-tip and lateral interactions between their helical bundles (Frost et al., 2008; Shimada et al., 2007). Concerted changes within the lattice allow F-BAR domains to interact with membrane tubes of various diameters (Frost et al., 2008). In addition, tip-to-tip interactions between F-BAR domain dimers have been observed in crystals (Shimada et al., 2007). It is less clear to-date whether BAR domains form similar arrangements for sensing or tubulating membranes.

Here we report crystal and solution structures of the $Snx9^{PX\text{-}BAR}$ lacking a short helix at the N-terminus of the yoke domain. The crystal structures by and large resemble the previously published structure despite the truncation of the yoke domain. In a second, new crystal form, tip-to-tip interactions between adjacent BAR domain dimers were observed suggesting a mode of oligomerization of Snx9 similar to F-BAR domain proteins. In SAXS-based solution studies, a more elongated and curved conformation of $Snx9^{PX\text{-}BAR}$ was observed, in addition to the compact structure seen in crystals. In the curved state, the PX domains appear to be flexibly linked to the BAR

domain dimer. In contrast, Endophilin exists in a rigid conformation with comparable degree of curvature in crystal lattices and in solution. Using SAXS, we determined a low-resolution structure of full-length Endophilin and could position the SH3 domains at the distal tips of the BAR domains. Our data suggests that particular BAR domains might be dynamic, and that hinge and inter-domain motions might modulate their preference for tubules and vesicles with different degree of curvature.

MATERIALS AND METHODS

Protein purification

cDNA corresponding to the human sequence of full-length Snx9 (Howard et al., 1999) or the isolated PX-BAR module ($Snx9^{PX\text{-}BAR}$; residues 230-595 of human Snx9) was amplified by standard PCR and was cloned into a modified pProExHT expression plasmid (Invitrogen) yielding a Precision Protease-cleavable N-terminally hexahistidine-tagged protein. *Escherichia coli* cells BL21 (DE3) (Novagen) were transformed by electroporation. Cells were grown in TB medium supplemented with 100 mg/l Ampicillin at 37°C. The temperature was reduced to 18°C at a cell density corresponding to an absorbance of 1.0 at 600 nm, and protein production was induced with 1 mM IPTG. Proteins were expressed for 12–16 hr.

Cells were collected by centrifugation for 20 min at 4000 x g, resuspended in NiNTA buffer A (25 mM Tris-Cl pH 8.2, 500 mM NaCl, 20 mM imidazole, and 5 mM β-mercaptoethanol) and frozen in liquid nitrogen. Cell suspensions were thawed in a water bath at 25°C. After cell lysis using sonication, cell debris was removed by centrifugation at 40,000 x g for 1 hr at 4°C. Clear lysates were loaded onto HisTrap NiNTA columns (GE Healthcare) equilibrated in NiNTA buffer A. The resin was washed with 20 column volumes of NiNTA buffer A, and proteins were eluted on a gradient from 20 to 500 mM imidazole in NiNTA buffer A over 15 column volumes.

Fractions containing His-tagged protein were pooled and buffers were exchanged using a Fast Desalting column (GE Healthcare) equilibrated in protease buffer (25 mM Tris-Cl pH 7.5, 300 mM NaCl, and 5 mM β-mercaptoethanol). Precision Protease was added and reactions were incubated over night at 4°C. Uncleaved proteins and protease were removed by NiNTA affinity chromatography and cleaved, pure protein was recovered in the flow-through. Proteins were further subjected to size exclusion chromatography on a Superdex200 column (GE Healthcare) equilibrated in gel filtration buffer (25 mM Tris-Cl pH 7.5, 100 mM NaCl, and 2 mM DTT). Fractions containing protein were pooled and concentrated on a Centricon ultrafiltration device (Millipore) to a final concentration of approximately 30 mg/ml. Protein concentrations were determined by measuring UV extinction at 280 nm shortly after purification using the exact buffer used for gel filtration chromatography as blank. In the time frame of the purification and concentration determination we observed only a small error from buffer aging (~1%/hour). From each purification, we stored buffer samples at -80°C, under identical conditions as for protein samples, that served as buffer blanks in SAXS experiments. Protein aliquots were frozen in liquid nitrogen and stored at -80°C.

Selenomethionine-substituted $Snx9^{PX\text{-}BAR}$ was expressed as a hexahistidine-tagged SUMO-fusion protein. Purification proceeded as described above with the exception that the affinity and purification tag was removed with the SUMO-specific protease Ulp-1 from *S. cerevisiae*. cDNA corresponding to $Endophilin^{BAR}$ (residues 1-256) and full-length Endophilin (isoform A1 from mouse) were cloned into the modified pProExHT vector (see above) and purification proceeded as outlined for Snx9.

Lysozyme standards (Sigma) were prepared from commercially available protein by weight determination, and concentrations were double-checked by measuring UV extinctions.

Analytical ultracentrifugation

Sedimentation equilibrium experiments were performed using a XL-I analytical ultracentrifuge (Beckman Coulter) equipped with an AN-60 Ti rotor. Proteins at concentrations between 30-140 μM in assay buffer (25 mM Tris-HCl [pH 7.5], 250 mM NaCl, 1 mM TCEP) were analyzed at centrifugations speeds of 8,000 rpm, 14,000 rpm and 25,000 rpm. Data collection was carried out at 301 nm, followed by data analysis and global equilibrium fitting using the program Ultrascan (Demeler, 2005). Satisfactory fitting quality was only obtained for a monomer-dimer equilibrium. In the final steps, a more refined global fitting was applied using the Levenberg-Marquardt NLS algorithm for calculation of disassociation constants.

Crystallization, X-ray Data Collection, and Structure Solution

Native crystals of $Snx9^{PX\text{-}BAR}$ were obtained by hanging drop vapor diffusion by mixing equal volumes of protein (5-30 mg/ml) and reservoir solution (1.8 M ammonium sulfate, 0.1 M Tris-base pH 8.5, 5 % PEG 400, 0.05 M MgSO4, and 15 % xylitol.) followed by incubation at 20°C. Crystals with block-shaped and rod-shaped morphology grew under identical conditions and appeared within 1 to 2 day with typical dimensions of 0.3 mm x 0.3 mm x 0.3 mm (orthorhombic system), and 0.3 mm x 0.02 mm x 0.02 mm (rhombohedral system). Crystals were flash-frozen in liquid nitrogen and kept at 100K during data collection. Selenomethionine-substituted protein did not crystallize spontaneously. Streak seeding using native crystals was used to initiate crystallization.

Crystallographic statistics for data collection are shown in Table 1. Data sets were collected using synchrotron radiation at the Cornell High Energy Synchrotron Source (CHESS, Ithaca, beamline A1 and F2). Data reduction was carried out with the software package HKL2000 (Otwinowski and Minor, 1997). The space groups were determined to be $C222_1$ with a = 67.0 Å, b = 144.0 Å, c = 103.3 Å and $R3_2$ with a = 131.8 Å, b = 131.8 Å, c = 569.1 Å. The asymmetric unit consists of one molecule, bound to a sulfate ion (space group $C222_1$) or 3 molecules (space group $R3_2$). The initial structure was determined using experimental phases obtained from selenomethionine-substituted crystals, and the phases were calculated using Single-wavelength Anomalous Dispersion (SAD). Structure solution was carried out using the software SOLVE/RESOLVE within the software package Phenix (Adams et al., 2002; Terwilliger and Berendzen, 1999). The majority (80%) of the $Snx9^{PX\text{-}BAR}$ model was built automatically in Phenix. Refinement using CNS (Brunger et al., 1998) and O (Jones et al., 1991) yielded the final model. Molecular replacement using Phenix was applied to solve the low-resolution structure in the rhombohedral space group. Illustrations were made in Pymol (DeLano Scientific, Palo Alto, CA).

Small-angle X-ray scattering (SAXS) data collection and processing

SAXS experiments were carried out at the Advanced Photon Source (APS; BESSRC-CAT beamline 12-IDC, Argonne, IL) at an electron energy of 12KeV. Low and high angle data (with S_{min}~0.25 and S=0.15-0.55, respectively) were collected separately. Proteins were filtered through 0.2 μm membrane filters (Millipore) and loaded into a flow cell. Scattering data were collected in triplicates at protein concentrations between 2-10 mg/ml. Taking into consideration the dissociation constant for dimerization, only samples between 7.5-10 mg/ml (resulting in ~10% monomer fraction) were used for further analysis. Exposure times were optimized by

dose-response experiments, and only exposures (2 seconds) showing no sign of radiation damage and aggregation were used. Background scattering was collected from gel filtration buffers used during protein purification. Scattering data was background-corrected, averaged and scaled using the program PRIMUS (Konarev et al., 2003), and analyzed using the programs GNOM and CRYSOL (Svergun et al., 1995; Svergun, 1992). Only scattering data with $S_{max}*R_g<1.3$, computed from Guinier plots at low angle regions, was considered for further analysis. Kratky plots, experimental molecular weight determinations, and the normalized scattering intensity I^0 were used to assess the folded-state of the proteins and overall data quality using the program IGOR PRO (version 5.04B). Distance distribution functios [P(r)] and D_{max} were determined using the program GNOM (Svergun, 1992).

Free-atom and rigid-body modelling

Ab initio free atom modeling was performed using the program GASBOR22 (Svergun et al. 2001) with the distance distribution function as fitting target. All proteins were found to be dimeric under the conditions used and P2 symmetry was used during modeling. 40 independent simulations are carried out for each protein. Solutions were then superimposed, averaged, and filtered using the program DAMAVER (Volkov and Svergun, 2003). Calculated scattering curves and goodness of fit (χ) were calculated using the program CRYSOL, as described earlier (Sondermann et al., 2005; Svergun et al., 1995).

The conformational space occupied by the *ab initio* models and heterogeneity of the solutions were illustrated by an unrooted distance tree diagram constructed using the program NEIGHBOR of the PHYLIP package (http://evolution. genetics.washington.edu/phylip) and on the basis of the NSD matrix produced by

DAMAVER. Tree-files were then visualized in TreeView (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html).

Atomic models with decreased or increased curvature were constructed manually from crystal structures using PyMol, and were scored against the experimental data in CRYSOL, based on the χ values. The model with the lowest χ value was selected for automated rigid body modeling using the program BUNCH (Petoukhov and Svergun, 2005). During this optimization process, the BAR domain dimer was treated as a single domain with fixed orientation, and the two PX domains were allowed to move relative to the BAR domain.

RESULTS AND DISCUSSIONS

Crystal structure of Snx9$^{PX\text{-}BAR}$

Crystallization trials carried out with Snx9$^{PX\text{-}BAR}$, the membrane binding and dimerization module of Snx9 (residues 230-595 of human Snx9), yielded two crystal forms with distinct growth morphologies and diffraction properties. For the first form, crystals grown with selenomethionine-substituted protein diffracted X-rays to a maximal resolution of 2.0 Å (Table 4.1). The structure was solved by single anomalous dispersion (SAD) phasing, using a data set collected at the selenium anomalous scattering peak wavelength. Snx9$^{PX\text{-}BAR}$ crystallized in the space group $C222_1$ with one molecule in the asymmetric unit. The biologically relevant dimer is completed by a symmetry-related molecule burying ~5700Å^2 at the dimer interface (Figure 4.2.B). The packing and overall conformation is very similar to the structure of a longer construct of Snx9 containing the complete yoke domain (rmsd of ~0.80 across 2234 atoms) (Pylypenko et al., 2007). Local differences were observed at the unpaired distal helical tips of the BAR domain dimer (Figure 4.2.C).

The BAR domains line up to form a central 6-helix bundle via both hydrophobic and complementary electrostatic interactions (Figure 4.2.B and 4.3.A). Interestingly, each BAR domain contributes six tyrosine residues to the homotypic dimerization surface (in addition to two buried tyrosine residues in the PX-BAR interface) (Figure 4.3.C). Phosphorylation of Snx9 by Ack2 has been reported to be required for epidermal growth factor receptor degradation (Childress et al., 2006; Lin et al., 2002), and our analysis may suggest a structural model for phosphorylation-dependent changes in conformation or oligomerization. The C-terminal tail of the BAR domain was resolved in the crystal structure and forms important contacts by reaching across the last helix of an adjacent BAR domain fold, consistent with the finding that deletion of the C-terminal 13 residues results in loss of dimerization, ACK2 binding and epidermal growth factor receptor degradation (Childress et al., 2006) (Figure 4.3.B).

The BAR domains line up to form a central 6-helix bundle via both hydrophobic and complementary electrostatic interactions (Figure 4.2.B and 4.3.A). Interestingly, each BAR domain contributes six tyrosine residues to the homotypic dimerization surface (in addition to two buried tyrosine residues in the PX-BAR interface) (Figure 4.3.C). Phosphorylation of Snx9 by Ack2 has been reported to be required for epidermal growth factor receptor degradation (Childress et al., 2006; Lin et al., 2002), and our analysis may suggest a structural model for phosphorylation-dependent changes in conformation or oligomerization. The C-terminal tail of the BAR domain was resolved in the crystal structure and forms important contacts by reaching across the last helix of an adjacent BAR domain fold, consistent with the finding that deletion of the C-terminal 13 residues results in loss of dimerization, ACK2 binding and epidermal growth factor receptor degradation (Childress et al., 2006) (Figure 4.3.B).

Figure 4.2. Structure of Snx9PX-BAR.(A) Domain organization of Sorting Nexin 9 (Snx9) and Endophilin. Schematic presentations and binding partners of the domains in Endophilin and Snx9 are shown.(B)Crystal structure of Snx9PX-BAR. Two orthogonal views of Snx9PX-BAR (residues 230-595 of human Snx9; space group C2221) are shown with the PX domains colored in green and brown and the BAR domains colored in orange and grey. The membrane interaction surface is indicated as dashed line.(C)Local differences in the structures of Snx9. The structure of Snx9PX-BAR (residues 230-595) was superimposed onto Snx9PX-BAR including the N-terminal amphipathic helix and helix 1YN of the yoke domain (residues 214-595; PDB:2RAJ; Pylypenko et al., 2007).

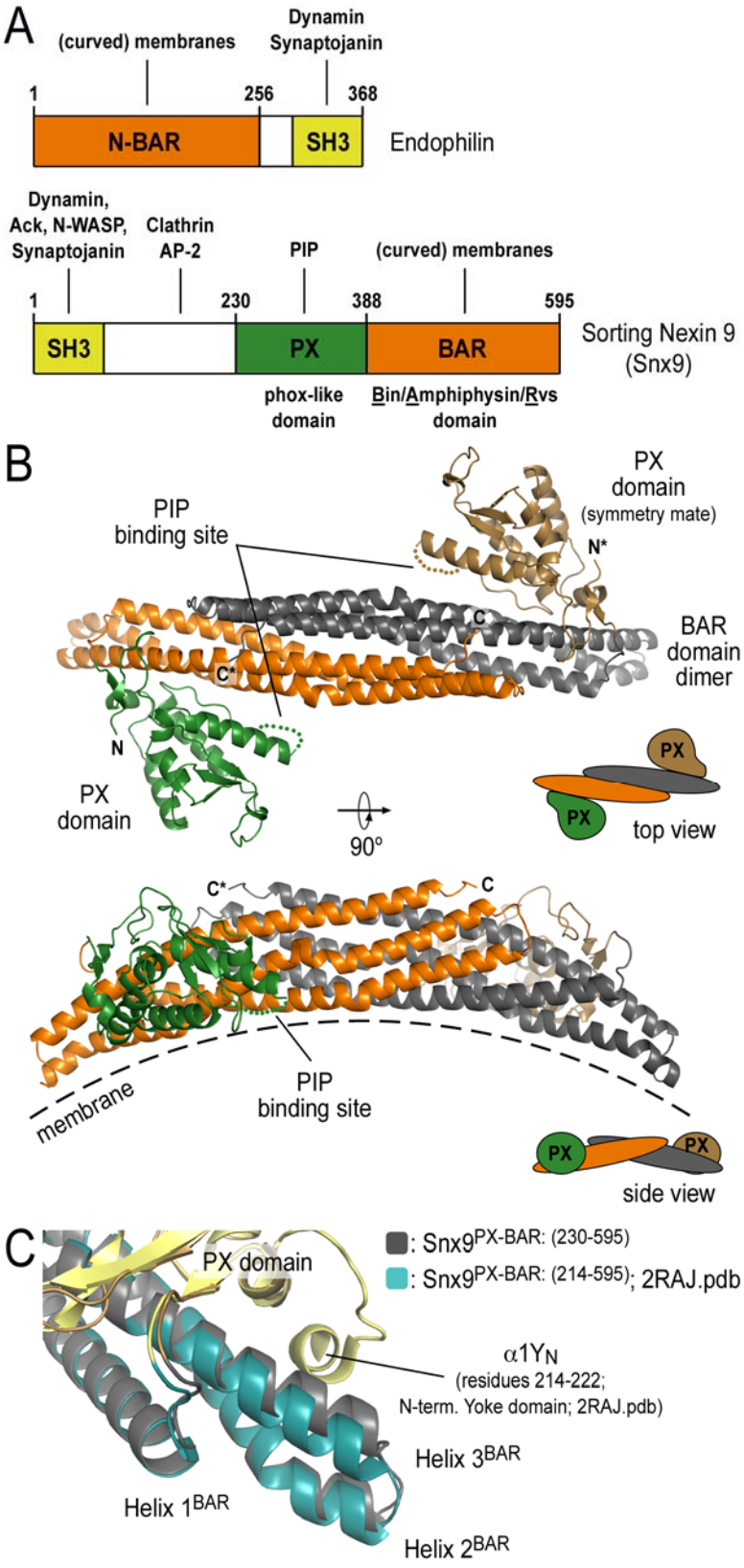

A
(curved) membranes
Dynamin
Synaptojanin
1
256
368
N-BAR
SH3
Endophilin
Dynamin,
Ack, N-WASP,
Synaptojanin
Clathrin
AP-2
PIP
(curved) membranes
1
230
388
595
SH3
PX
BAR
Sorting Nexin 9
(Snx9)
phox-like
domain
Bin/Amphiphysin/Rvs
domain
B
PIP
binding site
PX
domain
(symmetry mate)
N*
C
C*
BAR
domain
dimer
N
PX
domain
90°
PX
PX
top view
C*
C
membrane
PIP
binding site
PX
PX
side view
C
PX domain
: Snx9PX-BAR: (230-595)
: Snx9PX-BAR: (214-595); 2RAJ.pdb
α1YN
(residues 214-222;
N-term. Yoke domain; 2RAJ.pdb)
Helix 3BAR
Helix 1BAR
Helix 2BAR

Table 4.1. Data Collection and Refinement Statistics		
Structure	**Snx9 PX-BAR** (selenomethionine)	**Snx9 PX-BAR** (native)
Space Group	$C222_1$	$R3_2$
Unit Cell	a=67.0 Å, b=144.0 Å, c=103.3 Å α=β=γ=90°	a=131.8 Å, b=131.8 Å, c=569.1 Å α = β =90°, γ =120°
X-ray source	CHESS (F2)	CHESS (A1)
Wavelength (Å)	0.9792	0.9771
Resolution (Å)	50-2.08 (2.15-2.08)	50-4.1 (4.25-4.10)
Measured reflections (#)	437,532	204,578
Unique reflect. (#)	30,311	15,467
Data redundancy	14.4 (14.0)	13.2 (12.8)
Completeness (%)	100.0 (100.0)	100.0 (100.0)
R_{sym} (%)	11.2 (35.0)	10.9 (88.8)
I / σ_I	20.6 (9.6)	25.9 (3.0)
Current Model Refinement Statistics		
Phasing	SAD	Molecular Replacement
Molecules/AU	1	3 (with 67% solvent)
R_{work} / R_{free} (%)	22.3 / 24.7	36.3/36.8
Free R test set size (#/%)	2072/6.8	682/4.4
Number of protein atoms	2883	Phaser statistics (likelihood gain/Z-score):
Number of ligand atoms/solvent	5/143	1st $Snx9^{PX-BAR}$: 315.3/16.42
Rmsd bond length (Å)	0.005	2st $Snx9^{PX-BAR}$: 875.8/25.44
Rmsd bond angles (°)	1.10	3rd $Snx9^{BAR}$: 906.1/9.74

In a previously published structure, the N-terminal 15 residues of the yoke domain that were omitted in our construct form an additional short helix buttressing against the tip of the distal BAR domain (Pylypenko et al., 2007). This helix does not contribute directly to the PX-BAR domain interface ($\alpha 1Y_N$; Figure 4.3.C) and is not required for the formation of the conformation observed in the crystal lattices.

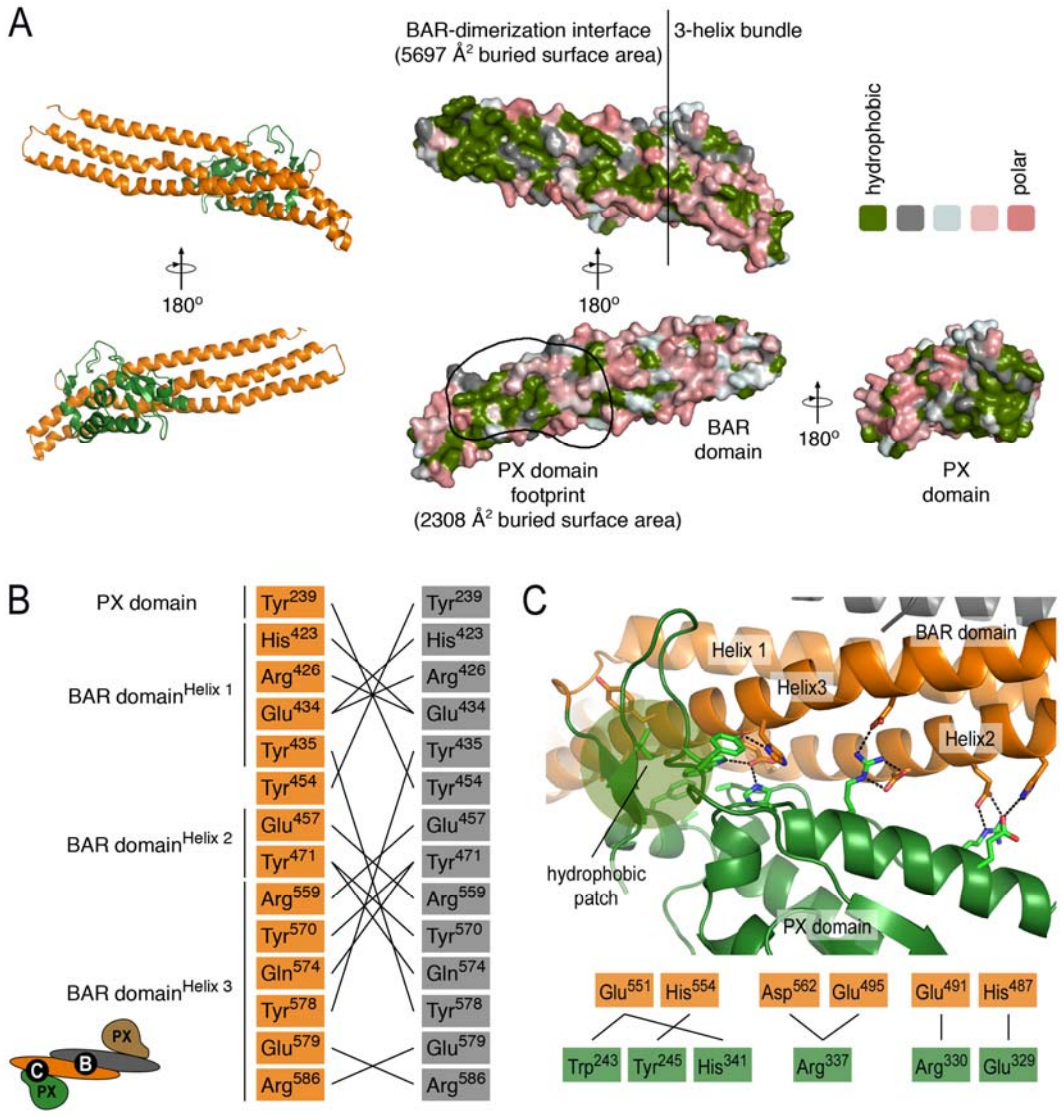


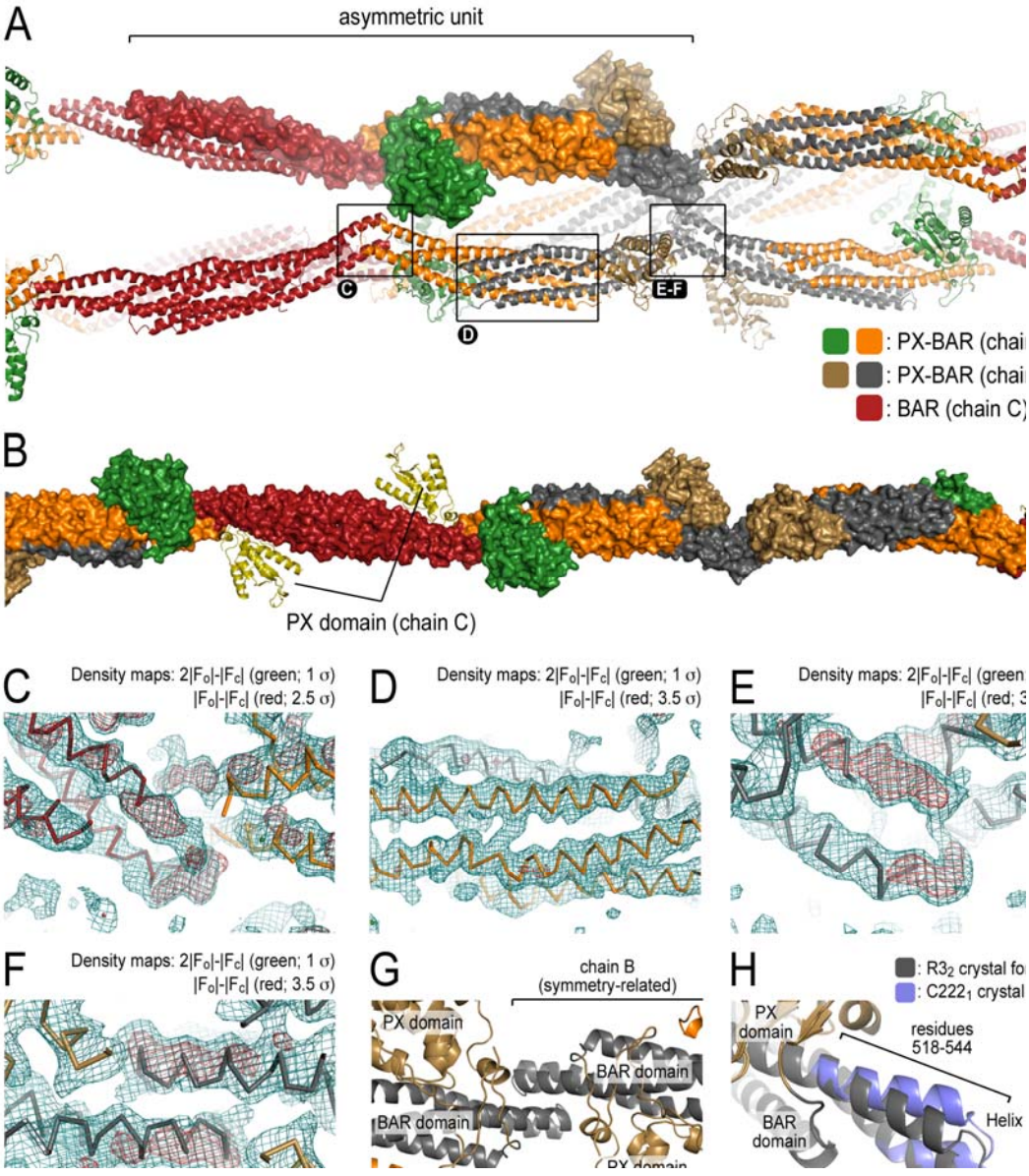
Figure 4.3. Interfacial surfaces between the BAR domains and the BAR and PX domains in the crystal structure of Snx9$^{PX\text{-}BAR}$. (A) Characterization of the BAR domain dimer and the PX-BAR domain interfaces of Snx9. Domains are shown as solid surface and colored according to their hydrophobicity. Hydrophobic residues are shown in green. Buried surface areas were calculated using CNS (Brunger et al., 1998). (B) Polar and charge interactions in the BAR domain dimerization interface. A schematic representation of polar residues contributing to the homo-dimerization surface is shown. (C) Polar and charge interactions in the PX-BAR domain interface. Key residues in the interface between the PX domain (green) and the BAR domain (orange) are shown. In addition to a hydrophobic patch at one corner of the interface, a large number of polar and charge interactions contribute to the interface.

Tip-to-tip interaction observed in a second Snx9$^{PX\text{-}BAR}$ crystal lattice

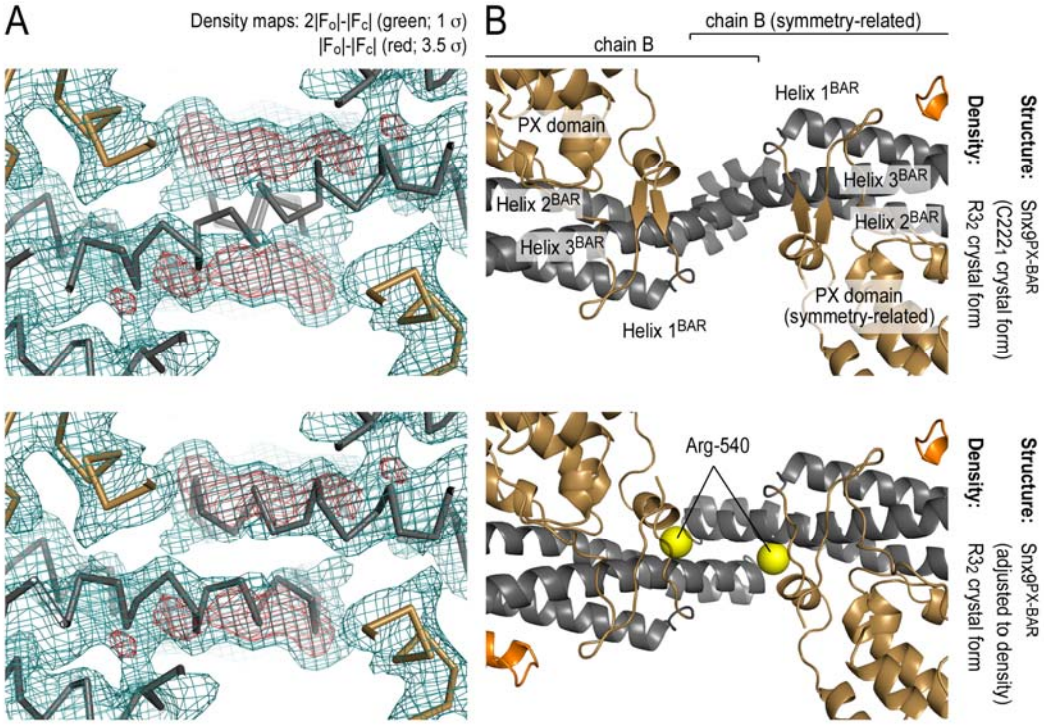
We also obtained a second, rod-shaped crystal form that grew under identical conditions using the native protein. Despite the low resolution (4.1 Å), data quality was sufficient to determine the space group ($R3_2$) and packing of molecules (3 molecules/asymmetric unit). Molecular replacement using the high-resolution structure as search model was only successful when protein backbone clashes were permitted or when the distal tip of the BAR domain was truncated, suggesting that lattice formation involved subtle conformational changes in this region. We could place two PX-BAR domain units into the electron density with high confidence (Table 4.1). A third BAR domain could be located with lower scores with the corresponding PX domain not contributing to the solution (Figure 4.4.A and 4.4.B). Given the relatively low resolution and the lack of experimental phases, model optimization was limited to rigid body refinement. Inspection of electron density maps revealed additional un-interpreted density in the region where the BAR domain helices had been truncated to avoid clashes (Figure 4.4.C and 4.4.E), whereas maps indicated good agreement of the model with the crystallographic data in the 6-helix bundle region (Figure 4.4.D). In the final model, the helical tips of the BAR domains were adjusted in accordance to the maps (Figure 4.4.F and 4.4.G). The differences between the structures from the two crystal forms were illustrated in Figure 4.4.H.

Although the exact interactions between the predominantly polar or charged residues in the tip-interaction interface could not be determined, one residue, Arginine-540, caught our attention. This residue is conserved in Snx9 and the other two members of this Sorting Nexin subfamily, Snx18 and Snx30 (Haberg et al., 2008). It is located on the helix that is most distal to the membrane interacting surface (helix 3 of the BAR domain).

Figure 4.4. Tip-to-tip interactions between BAR domains in rhombohedral crystals of Snx9PX-BAR. (A)Crystal packing in Snx9PX-BAR crystals solved in a rhombohedral space group (R32). Two molecules of Snx9PX-BAR (colored as in Figure 2B) and an isolated BAR domain (colored in red) were placed into the electron density using molecular replacement and are shown as solid surfaces. Molecules from adjacent asymmetric units are shown in ribbon presentation. The boxed regions are shown in C-F as close-up views. (B)Isolated filament of Snx9PX-BAR. Molecules lined up along the long crystal axis are shown as surface presentation. The position of the PX domain of chain C was not included in the molecular replacement solution. Its position relative to the BAR domain is shown on the basis of the high-resolution crystal structure shown in Figure 2B. The filament formed by Snx9PX-BAR dimers is twisted with approximately four dimers completing a 180° rotation. (C)Electron density for tip-to-tip interactions between BAR domains within one asymmetric unit. Density is shown for the region where the BAR domain tips of protomer C (red C - trace) and protomer A (orange C α -trace) convene. BAR domain tips (residues 518-544 of Snx9) had been removed from the search model prior to molecular replacement to avoid clashes. Electron density maps shown have amplitudes of (|Fo| - |Fc|) or (2|Fo| - |Fc|), with Fo and Fc being the observed and calculated structure factors, and were contoured at 2.5 and 1 , respectively. (D)Electron density for the 6-helix bundle region of a Snx9PX-BAR dimer. Electron density maps were contoured at 3.5 (|Fo| - |Fc|) and 1 . (2|Fo| - |Fc|). (E)Electron density for tip-to-tip interactions between crystal symmetry-related BAR domains. Maps were contoured as in (D). (F)Modeling of the tip region of the BAR domains based on the observed electron density. Molecules without truncation of the helical tips were superimposed onto molecules from the molecular replacement solution. Residues 518-544 were treated as rigid body and were placed into the electron density maps. (G)Ribbon presentation of tip-to-tip interactions between BAR domains. The orientation is identical to the view shown in (F). (H)Comparison of crystal structures of Snx9PX-BAR. Conformational changes in the tip regions (residues 518-544) of Snx9PX-BAR are shown after superimposition of the molecules obtained from orthorhombic crystals and from rrhombohedral crystals.


A
asymmetric unit
C
D
E-F
: PX-BAR (chai
: PX-BAR (chai
: BAR (chain C)
B
PX domain (chain C)
C
Density maps: 2|Fo|-|Fc| (green; 1 σ)
|Fo|-|Fc| (red; 2.5 σ)
D
Density maps: 2|Fo|-|Fc| (green; 1 σ)
|Fo|-|Fc| (red; 3.5 σ)
E
Density maps: 2|Fo|-|Fc| (green
|Fo|-|Fc| (red; 3
F
Density maps: 2|Fo|-|Fc| (green; 1 σ)
|Fo|-|Fc| (red; 3.5 σ)
G
chain B
(symmetry-related)
PX domain
BAR domain
BAR domain
H
: R32 crystal fo
: C2221 crystal
PX
domain
residues
518-544
BAR
domain
Helix

In the longer Snx9$^{PX\text{-}BAR}$ structure (residues 214-595) (Pylypenko et al., 2007), Arg-540 is engaged in contacts with the N-terminal helix of the yoke domain. This helix packs against the tip region of the BAR domain and might contribute to the stabilization of the particular conformation observed in this crystal form (Figure 4.5.A). In the crystal form with $R3_2$ symmetry, helix 3 of the adjacent BAR domain partially occupies the space of the N-terminal yoke domain helix, and would displace it in the context of longer constructs (Figure 4.5.B). This competitive binding mode might be utilized by Snx9 to positively or negatively regulate oligomerization or membrane association.



S

Figure 4.5. Comparison between structures of Snx9$^{PX\text{-}BAR}$. (A) Comparison of the conformation in the tip region of the BAR domains in crystal structures of Snx9$^{PX\text{-}BAR}$. Structures obtained from orthorhombic (blue) and rhombohedral (grey) crystals and the previously published structure of Snx9$^{PX\text{-}BAR\text{-}long}$ (residues 214-595; PDB: 2RAJ; cyan) (Pylypenko et al., 2007) were superimposed based on their Cα-positions. (B) Position of the yoke domain. form. A second, symmetry-related molecule is shown. Two different views of the tip-to-tip interaction region are shown.

Tip-to-tip interactions appeared to be rather weak with a buried surface area of ~1220 $Å^2$, but might be strengthened and relevant when the BAR domains interact with the membrane. Although a monomer-dimer equilibrium dominates in solution, this model is consistent with the observation that Snx9 oligomerizes on the surface of vesicles (Yarar et al., 2007). A similar mode of interaction mediated by their tips has been described for F-BAR domains (Frost et al., 2008; Shimada et al., 2007). The packing described here is the first structural evidence for similar interactions within the BAR domain subfamily.

Solution scattering profiles of Endophilin and Snx9

For the PX-BAR module of Snx9, two modes of vesicle binding have been described (Pylypenko et al., 2007). In the presence of phosphoinositides that bind to the PX domains, Snx9 showed no preference for a particular curvature of membranes. In contrast, in the absence of PIPs $Snx9^{PX\text{-}BAR}$ had a clear binding preference for vesicles with high curvature (50-100 nm), suggesting distinct curvature adaptations of the protein. One mechanism for achieving such curvature bias would be the formation of distinct helical lattices of rigid domains, driven by tip-to-tip and lateral interactions, on membrane tubes accommodating different degrees of membrane curvature as described for F-BAR domains (Frost et al., 2008). An alternative mechanism could involve a conformational change increasing the intrinsic curvature of the BAR domain dimer.

To explore the conformation and intrinsic flexibility of BAR domain-containing proteins in solution, we carried out small-angle X-ray scattering (SAXS) experiments. We collected X-ray scattering data from monodispersed solutions of $Endophilin^{BAR}$, $Snx9^{PX\text{-}BAR}$, and the respective full-length proteins. Typical scattering

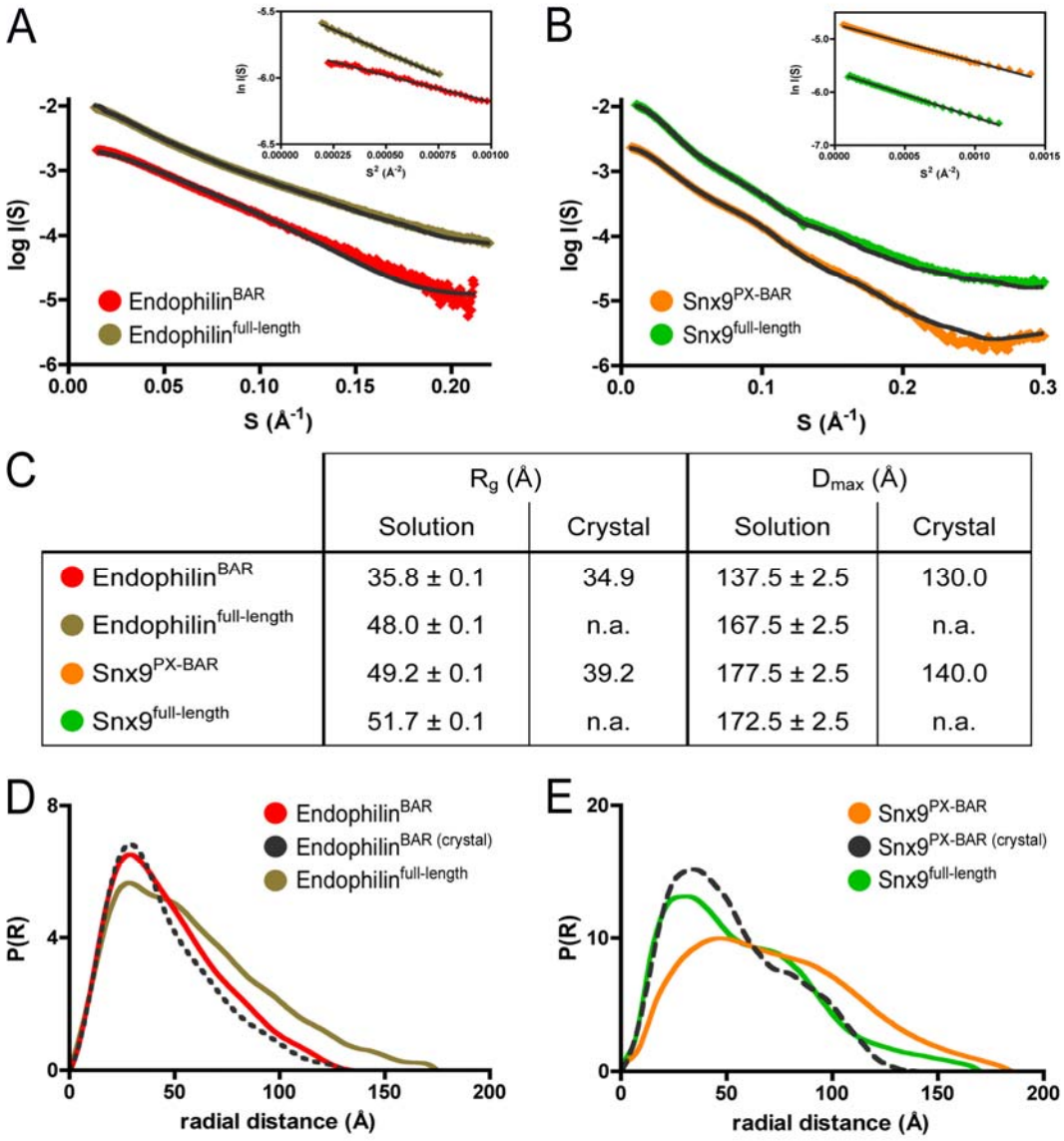
curves, measured to S values corresponding to 28 Å and 21 Å resolution, respectively, are shown in Figures 4.5.A and 4.5.B. In addition to choosing exposure times that showed no radiation damage, careful inspection of the Guinier plots at low angles indicated good data quality and no significant protein aggregation (inset in Figures 4.6.A and 4.6.B) (Konarev et al., 2003). In addition, molecular weight determinations based on the scattering data corroborate good data quality (Table 4.2).

From the scattering profiles molecular geometry parameters such as the radius of gyration (R_g) and maximum diameter of a molecule or assembly (D_{max}) can be obtained in a model-independent manner (Wall et al., 2000) (Figure 4.6.C). For Endophilin^{BAR}, we observed a good agreement of these values calculated from the scattering data with the dimensions calculated from the crystal structure (R_g=35.8 Å [solution] *vs* 34.9 Å [crystal]; D_{max}=137.5 Å [solution] *vs* 130.0 Å [crystal]) (Svergun et al., 1995). The dimensions of full-length Endophilin were larger with an R_g=48.0 Å and D_{max}=167.5 Å suggesting an extended conformation. In contrast, Snx9$^{PX\text{-}BAR}$ appeared to be elongated in solution with an increase in R_g of 10 Å and in D_{max} of 38 Å compared to the values calculated from the crystal structure (R_g=49.2 Å [solution] *vs* 39.2 Å [crystal]; D_{max}=177.5 Å [solution] *vs* 140.0 Å [crystal]). In contrast, the full-length protein did not appear to be extended further having a similar R_g compared to Snx9$^{PX\text{-}BAR}$ (R_g=51.7 Å) and a shorter maximum diameter (D_{max}=172.5 Å) despite the presence of two additional domains (residues 1-229).

Based on the scattering profile, distance distribution functions [P(r)] were calculated by inverse Fourier transformation, providing low-resolution shape information (Figures 4.6.D and 4.6.E). The P(r) function for Endophilin^{BAR} was in good agreement with the theoretical P(r) function calculated from the crystal structure showing a distribution with a sharp peak at shorter distances gradually dropping off to the maximum distance (Figure 4.6.D). Such a profile is characteristic for a rod-shaped

protein. For full-length Endophilin, a profile was observed with a peak at the same position as observed for Endophilin^{BAR} indicative of some structural resemblance between the isolated BAR domain and full-length protein. The P(r) function for Endophilin$^{full\text{-}length}$ had a bimodal profile with a second peak at larger radii and larger maximum distance compared to the P(r) function for Endophilin^{BAR}, suggesting that full-length Endophilin is an elongated molecule consisting of two folded domains (Figure 4.6.D).

As already indicated by the difference in R_g and D_{max}, the experimentally determined P(r) function for Snx9$^{PX\text{-}BAR}$ in solution deviated significantly from the P(r) function calculated from the crystal structure. The peaks for the solution state shifted towards larger radii with a significantly larger D_{max} (see above). R_g and D_{max} of full-length Snx9 suggested that the SH3 domain and low complexity region fold back onto Snx9$^{PX\text{-}BAR}$ in solution, an interpretation that was corroborated by the P(r) function. The profile partially resembled features of the one calculated from the Snx9$^{PX\text{-}BAR}$ crystal structure, which may indicate that intra-molecular interactions involving the additional domains constrain the conformation of full-length Snx9 (Figure 4.7.E). Analytical ultracentrifugation experiments yielded comparable dissociation constants describing a monomer-dimer equilibrium for full-length Snx9 and Snx9$^{PX\text{-}BAR}$ ($K_d^{full\text{-}length}$=7.9 μM and $K_d^{PX\text{-}BAR}$=4.2 μM) (Figure 4.8). Compaction and interaction of the SH3 domain and low-complexity region with the PX-BAR unit are likely to occur within one protomer of a dimer, and did not contribute to the dimerization of Snx9 that appears to be solely driven by the BAR domains. Intramolecular interactions in full-length Snx9 are likely to contribute to the autoinhibition of Snx9 (Worby et al., 2002), similar to mechanisms that have been proposed for Amphiphysin and Pick1, two BAR domain-containing proteins involved in membrane trafficking (Farsad et al., 2003; Madsen et al., 2008).



C

	R_g (Å)		D_{max} (Å)	
	Solution	Crystal	Solution	Crystal
Endophilin^{BAR}	35.8 ± 0.1	34.9	137.5 ± 2.5	130.0
Endophilin$^{full-length}$	48.0 ± 0.1	n.a.	167.5 ± 2.5	n.a.
Snx9^{PX-BAR}	49.2 ± 0.1	39.2	177.5 ± 2.5	140.0
Snx9$^{full-length}$	51.7 ± 0.1	n.a.	172.5 ± 2.5	n.a.

Figure 4.6. SAXS data for Endophilin and Snx9. (A)Scattering profiles from BAR domain of Endophilin and full-length Endophilin. X-ray scattering curves of are shown after averaging and solvent-subtraction. Theoretical scattering profiles calculated from the *ab initio* models with the lowest χ values are shown (black line). The inset shows Guinier at the low angle region. (B)Scattering profiles from solution of the PX-BAR unit of Snx9 and full-length Snx9.. (C) Radius of gyration (R_g) and maximum diameter (D_{max}) determined from solution scattering and crystal structures. Values based on available crystal structures were calculated using the program CRYSOL.(D). Distance distribution functions [P(r)] for Endophilin. P(r) curves of Endophilin^{BAR} (red) and Endophilin$^{full-length}$ (olive) were calculated from SAXS data or from the crystal structure of Endophilin^{BAR} (dashed line). (E) Distance distribution functions [P(r)] for Snx9. P(r) curves of Snx9^{PX-BAR} (orange) and Snx9$^{full-length}$ (green) were calculated from SAXS data or from the crystal structure of Snx9^{PX-BAR} (dashed line).

SAXS-based shape reconstructions for Endophilin

To investigate the origin for the larger D_{max} and discrepancy between the experimental scattering curve and P(r) function for $Snx9^{PX\text{-}BAR}$ and the corresponding profiles obtained from the crystal structure, we calculated low-resolution *ab initio* models using the program GASBOR (Svergun et al., 2001). In this approach, the positions of spherical scattering units representing residues in a protein are adjusted by a simulated annealing procedure to optimally match the data in real space using the P(r) function as a fitting target. Similar results were obtained when fitting was performed against the scattering intensity (data not shown). The proteins in this study were shown to be dimeric in solution by analytical ultracentrifugation (Gallop et al., 2006), allowing us to impose 2-fold symmetry during the modeling (trials in the absence of symmetry yielded comparable results; data not shown). For each protein, 40 independent simulated annealing runs were carried out. The accuracy of each model was assessed using a normalized residual coefficient χ that provides a value for the goodness of fit (Svergun, 1999). Low values (close to 1.0) indicate a good agreement of the models with the experimental data. While deviations from the ideal value might indicate systematic errors from differences between samples and buffer blanks, they may also indicate conformational heterogeneity. Superposition, averaging of all models, and filtering of the most robustly modeled regions generated a consensus envelope, which we refer to as "filtered volume" (Volkov and Svergun, 2003). The envelopes colored in grey represent the sum of all models.

The model for $Endophilin^{BAR}$ derived from SAXS data matched the shape of the crystal structure remarkably well (Figure 4.7A). χ values for individual solutions deviated slightly from the ideal value (1.0) (Supplemental Figure 5B), which may be due to flexible regions in the protein. The N-terminal amphipathic helix, accounting for ~10% of the protein, has been shown to be disordered in solution (Gallop et al.,

2006) and may not be modeled accurately. In addition, normal mode calculations suggested flexibility of the unpaired tips within the BAR domain dimers (data not shown). Given the striking similarity between crystal structures and the shape reconstruction of Endophilin^{BAR}, modeling results appeared to be robust despite somewhat suboptimal χ values. The shape reconstruction for full-length Endophilin

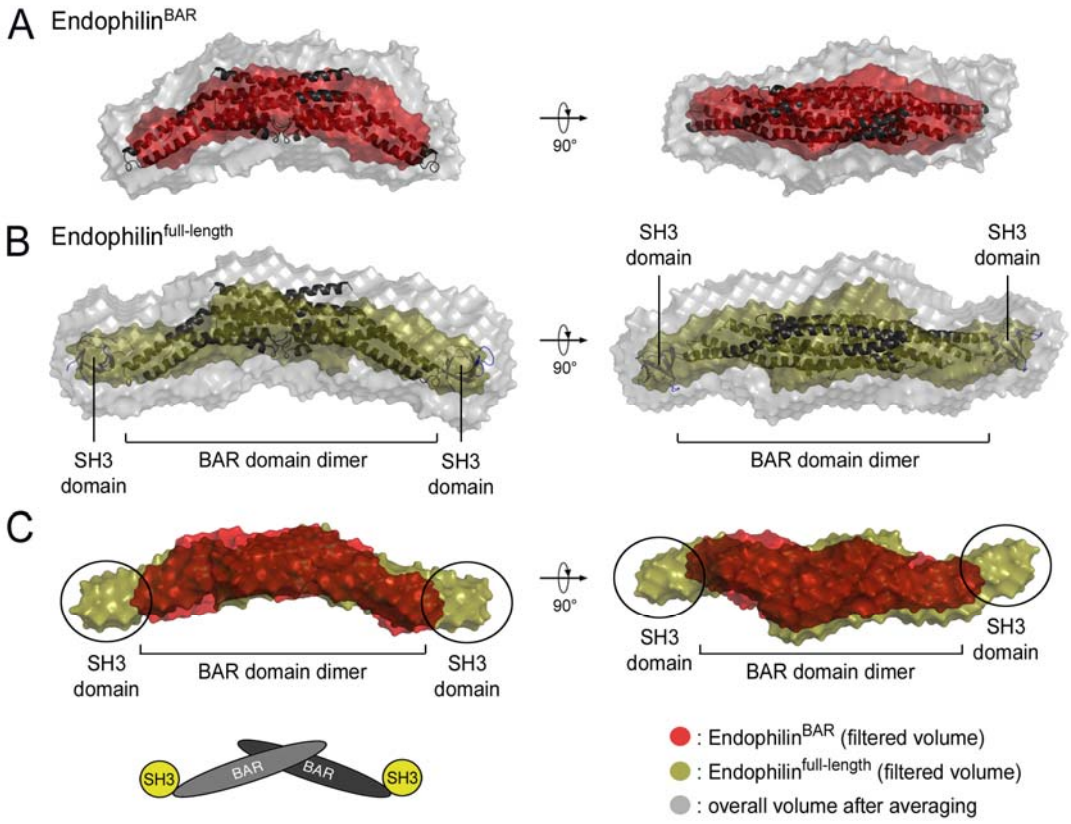


Figure 4.7. SAXS-based shape reconstructions of Endophilin^{BAR} and full-length Endophilin. (A)Endophilin^{BAR}. Models were calculated using the program GASBOR (Svergun et al., 2001) and the P(r) function as fitting target. The overall volume after averaging (grey) and the filtered volume (red) were calculated from 40 independent models using the program DAMAVER. The crystal structure of an Endophilin^{BAR} dimer (PDB: 1X03; black ribbon; Masuda et al., 2006) was manually docked into the envelope. Two orthogonal views are shown.(B)Full-length Endophilin. Averaging (producing the grey envelope), filtering (producing the olive envelope) and manual docking were carried out as described above. The crystal structure of an isolated SH3 domain (PDB: 3C0C) was manually docked into the extra density not occupied by the crystal structure of the BAR domain dimer.(C)Comparison of shape reconstructions for Endophilin^{BAR} and full-length Endophilin. Averaged and filtered envelopes shown in (A) and (B) were superimposed manually.

Table 4.2 Small angle X-ray scattering data statistics

	R_g (Å)[a]	D_{max} (Å)[b]	MM_{seq} (kDa)[c]	MM_{exp} (kDa)[d]	V_p (nm^3)[e]	$I_p^{norm}/I_{std}^{norm}$[f]	χ^2_{best}[g]
Endophilin$^{full-length}$	48.0 ± 0.5	167.5 ± 2.5	79.8	65.2 ± 10	114 ± 12	0.79 ± 0.16	1.91
Endophilin^{BAR}	35.8 ± 0.5	137.5 ± 2.5	56.4	50.5 ± 8	87 ± 9	0.93 ± 0.18	2.02
SNX9$^{full-length}$	51.7 ± 0.5	172.5 ± 2.5	133.2	115.0 ± 18	235 ± 24	0.87 ± 0.17	1.77
SNX9^{PX-BAR}	49.2 ± 0.5	177.5 ± 2.5	84.6	79.5 ± 12	150 ± 15	0.94 ± 0.19	1.58

(a) R_g determined from Guinier equation and data that satisfies $Q_{max} \times R_g < 1.3$; similar values were obtained using the program GNOM (Svergun, 1992).
(b) D_{max} was determined using the program GNOM (Svergun, 1992).
(c) MM_{seq} is the molecular mass of a dimeric protein calculated from the primary sequence.
(d) MM_{exp} is the molecular mass calculated from the scattering data calculated by

$$MM_{protein} = MM_{lysozyme} \times \frac{I^o_{protein} \times C_{lysozyme}}{I^o_{lysozyme} \times C_{protein}}$$

where $MM_{lysozyme}$ is the molecular mass of lysozyme**,** I^0 is the scattering intensity at zero angle (calculated from the linear region of the Guinier plot) and C (mg/ml) is the protein concentration (Mylonas and Svergun, 2007).
(e) V_p is the excluded volume (Porod volume) calculated using PRIMUS (Konarev et al., 2003). Porod volumes (in nm^3) for globular proteins are approximately twice the molecular weight (in kDa) (Petoukhov et al., 2006; Porod, 1982). (f) Normalized intensities for proteins (p) I^p_{norm} were calculated by

$$I_p^{norm} = (\frac{I^o}{C}) / MM_{seq}$$

where I^0 is the intensity calculated using the Guinier equation and C (mg/ml) is the protein concentration determined from its absorbance at 280 nm. Lysozyme served as a mono-dispersed protein standard (std). Ideally, the ratio $I^p_{norm}/I^{std}_{norm}$ equals 1.00 (Jeffries et al., 2008; Taraban et al., 2008). Cumulative errors reported here were based on lysozyme purity, measured deviations from buffer oxidation rates, and the estimated error for protein concentration determination by UV spectroscopy.
(g) χ^2_{best} is the lowest χ^2 value obtained for a free atom model from *ab initio* shape determinations using GASBOR.

(Figure 4.7.B) yielded a curved structure in the central region that reproduced the curvature of the isolated BAR domain very well (Figure 4.7). Extra density in the filtered volume not accounted for by the BAR domain dimer revealed that the SH3 domains of Endophilin are located at the tips of the BAR domain dimer, similar in position to the PH domains in APPL1 (Li et al., 2007; Zhu et al., 2007) (Figure 4.6.B and 4.7.C; Figure 4.1.B). For proteins with more than one lipid-binding domain (e.g. APPL1 and Snx9), the observed conformations maximize the surface that interacts with membranes. Alternatively, membrane binding could be a regulatory step, triggering conformational changes that might displace adjacent domains from the tips of the BAR unit. Interestingly, analysis of the electrostatic potential of the EndophilinA2 SH3 domain (PDB code: 3C0C; unpublished) revealed a predominantly negative potential of the domain (data not shown), incompatible with localization of this domain to the negatively charged membrane.

The robustness of the modeling and rigidity of Endophilin in solution was corroborated by investigating the structural relationship of the individual models using the output of the averaging process. After superimposing the independently modeled envelopes, pair-wise normalized spatial discrepancy (NSD) values (Volkov and Svergun, 2003) were calculated as part of the DAMAVER routine. We illustrated the conformational relationship of the models using an unrooted tree presentation on basis of the NSD values. (Figure 4.8.A-C). The dendrogram for Endophilin^{BAR} and Endophilin$^{full\text{-}length}$ showed a tight clustering of the models within a single node. The dendrogram for Endophilin$^{full\text{-}length}$ was slightly wider compared to the one computed for the isolated BAR domain, possibly due to the bigger size of the system leading to less convergence imposed by a constant step number in the annealing procedure and/or more conformational degrees of freedom (Figure 4.9.A-B).

Structure and plasticity of Snx9 in solution

In order to elucidate the molecular basis for the differences between the experimentally determined and crystal structure-derived P(r) function of $Snx9^{PX\text{-}BAR}$ we subjected the scattering data of $Snx9^{PX\text{-}BAR}$ to a similar analysis as described above. In contrast to Endophilin, $Snx9^{PX\text{-}BAR}$ models clustered into two nodes (Figure 4.8.C). Inspection of representative scattering profiles from models of the two nodes and comparison with the experimental scattering curve revealed small but significant deviations of the models from the experimental data (Figure 4.8.D-F). Although most parts of the experimental and model-based scattering curves overlapped almost perfectly (Figure 4.8.D), there were regions in the mid-resolution interval (corresponding to S~0.04-0.13 A^{-1}) in which models within each node systematically deviated from the scattering data. This was most visible in difference plots (Figure 4.8.E-F). All scattering curves derived from models in node 1 deviated from the experimental scattering curve, and did so in an inverse, non-overlapping way when compared to curves derived from models in node 2. The models in both nodes have comparable accuracy (χ values) with regard to the experimental data that was similar to the values obtained for the Endophilin models. Therefore, the clustering and deviations of the models from the scattering curve were exclusively based on differences in conformation (based on NSD values), and not on overall fitting accuracy (based on χ). The clustering of the models into two nodes was specific for Snx9 since similar calculations using Endophilin yielded a single, predominant conformation in solution, which is in agreement with crystal structures (Figure 4.6-4.8).

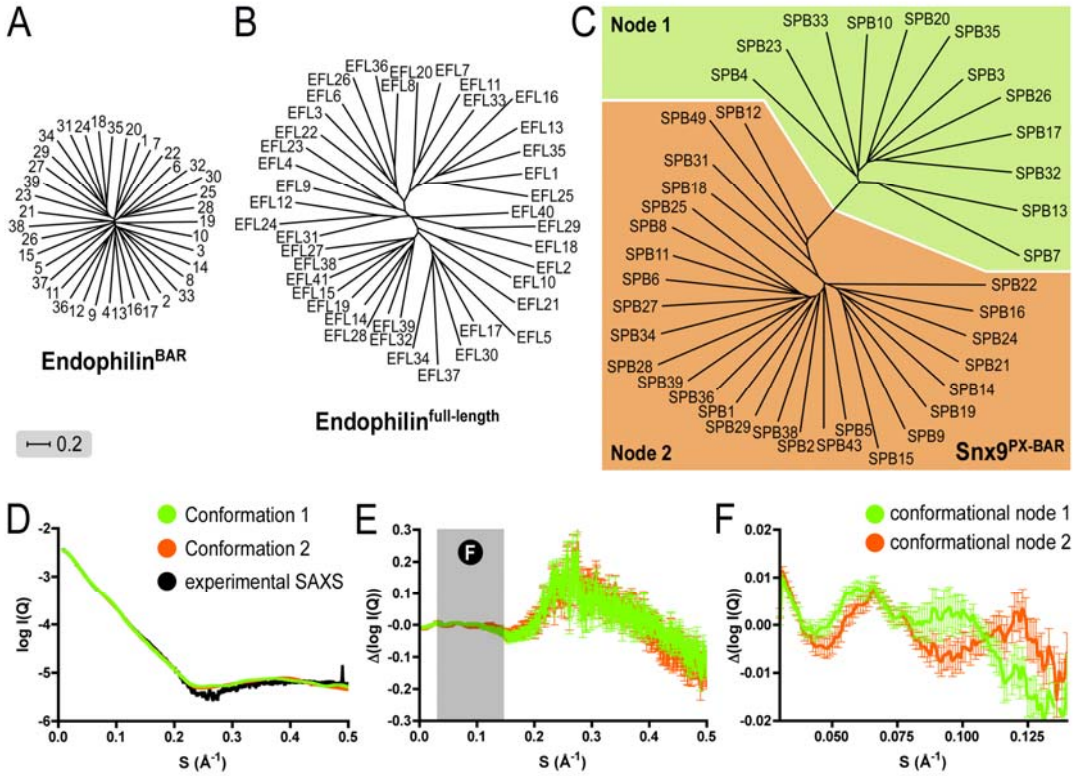


Figure 4.8. Modeling and conformational heterogeneity of Snx9^{PX-BA}(A)Structural comparisons of individual models from Endophilin^{BAR} shape reconstructions. Unrooted distance tree diagrams were constructed on the basis of the NSD matrix produced during superposition and averaging of the individual shape reconstructions.(B) Structural comparisons of individual models for Endophilin$^{full-length}$. Analysis was carried out as described in (A) (C)Structural comparisons of individual models for Snx9^{PX-BAR}. Analysis was carried out as described in (A). Models fell into one of two nodes of the dendrogram. Nodes are colored in light green (conformational ensemble 1) or orange (conformational ensemble 2). (D)Theoretical scattering profiles of SAXS-based models for Snx9^{PX-BAR} from both nodes. The calculated scattering curve for models with the lowest χ value (fitted against the experimental data) from node 1 (green) and 2 (orange) are shown. The experimental scattering curve is shown in black. (E) Difference scattering profiles for models in node 1 and 2. Scattering curves calculated from shape reconstructions in node 1 (light green) or 2 (orange) were subtracted from the experimental data. The average and standard deviations of difference scattering profiles are shown. The region in grey is shown as close-up view in (F). (F)Close-up view of difference scattering profiles in a region showing significant deviation from the experimental scattering profile. Curves are colored as before (E). A region of the scattering profile corresponding to S values from 0.025-0.15 Å^{-1} is shown.

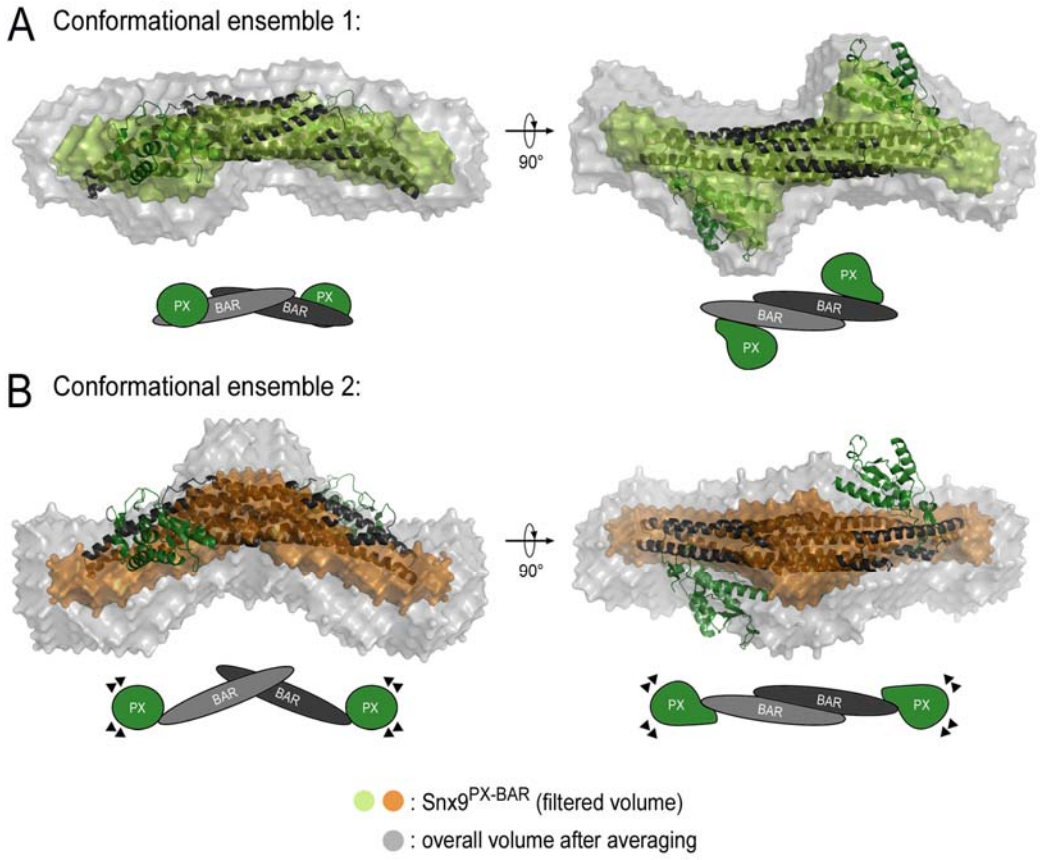


Figure 4.9. SAXS-based shape reconstructions for Snx9^{PX-BAR}. (A)Shape reconstruction based on models from conformational ensemble 1. The overall volume after averaging (grey) and the filtered volume (green) were calculated from models in node 1 of the NSD-based dendrogram (see Figure 4.7.C). The crystal structure of Snx9^{PX-BAR} was manually docked into the shape reconstruction. (B)Shape reconstruction based on models from conformational ensemble 2. The overall volume after averaging (grey) and the filtered volume (orange) calculated from models in node 2 of the NSD-based dendrogram (see Figure 7C) are shown.

Averaging using all models was deemed to be inappropriate and in fact resulted in misleading results due to misalignment of models from different clusters. Based on these observations, we decided to restrict our analysis to the envelopes created from separately averaging the models in node 1 and 2, respectively. Averaging of models in node 1 yielded a conformation that resembled features of the state observed in the crystal structure of Snx9^{PX-BAR} with regard to overall curvature and relative position of the BAR and PX domains (Figure 4.9.A). In contrast, averaging of models in the second node produced an envelope that was more curved than the

conformation in the crystal lattices suggesting that the BAR domain of Snx9 might adopt different degrees of curvature (Figure 4.9.B). In addition, there was no defined density in the filtered volume that could accommodate the PX domains indicating higher mobility of these relative to the BAR domain dimer in this conformation.

Rigid body modeling of the solution state of Snx9$^{PX\text{-}BAR}$

Based on the *ab initio* modeling of conformational states of Snx9$^{PX\text{-}BAR}$ in solution, we formulated a model that incorporates changes within the BAR domain dimer and in the position of PX domains relative to the BAR domains (Figure 4.9). BAR domain curvature can be modulated in two non-mutually exclusive ways, hinge motions involving pivoting of rigid BAR domains or kinking within the three-helix bundles. In the structure of APPL1, the most curved conformation of a BAR domain dimer described to-date, the angle formed between two rigid BAR domains is smaller than that observed in Endophilin or Amphiphysin, which argues that different curvatures can be obtained by variations in the dimer interface (Figure 4.2 and 4.4). For simplicity, we modeled changes in curvature as scissor-type hinge motions between two rigid BAR domain monomers (Figure 4.10.A and 4.10.B). Keeping the PX domain positions fixed relative to the BAR domain, we manually increased the curvature of the BAR domain dimer, scoring the models against the experimental scattering data. Starting at a relatively high χ value for the crystal structure (suggesting differences between the solution and crystalline state of Snx9), models with increased curvature fit significantly better to the scattering curve, with a minimum in χ at approximately 95°. The hinge angles estimated on basis of the *ab initio* models were close to 100° for the more curved envelope (conformational ensemble 2) and approximately 140° for the conformation resembling the crystal structure (conformational ensemble 1).

Next, we combined the manual optimization of the BAR module curvature with an automated search of PX domain orientations using the program BUNCH (Petoukhov and Svergun, 2005). Allowing the PX domains to adopt new positions relative to the BAR domains resulted in a significant drop in χ indicating a better agreement of the new conformation with the solution scattering data (Figure 4.10.C). Further optimization might not be possible given the restrictions in degrees of freedom of this approach and the possibility that the PX domains are flexibly linked to a more curved BAR domain unit as suggested by the *ab initio* modeling (Bernado et al, 2007). In addition, whereas the *ab initio* modeling approach enabled us to deconvolute two conformational ensembles, the rigid body modeling was complicated by the presence of a heterogeneous mixture of states. In this case, the final model would only represent a snapshot of a dynamic and heterogeneous ensemble (Figure 4.10.C). Yet, the relatively low χ values for more curved states with dislodged PX domains suggested that this conformation is a major form in solution contributing to the scattering profile.

Taken together, the modeling against the SAXS data suggested that $Snx9^{PX\text{-}BAR}$ could adopt a more curved state in solution. This result might explain the preference for smaller vesicles when the PX domain is not engaged in the binding (in the absence of PIPs) (Pylypenko et al., 2007). Under these conditions, it would mainly be the BAR domains contributing to membrane binding, and its intrinsic curvature would determine its preference for smaller vesicles. In the presence of appropriate PIPs, the PX domains will be engaged in membrane binding of Snx9 rendering it insensitive to membrane curvature. The functional role for these different modes of binding awaits further experimental elucidation.

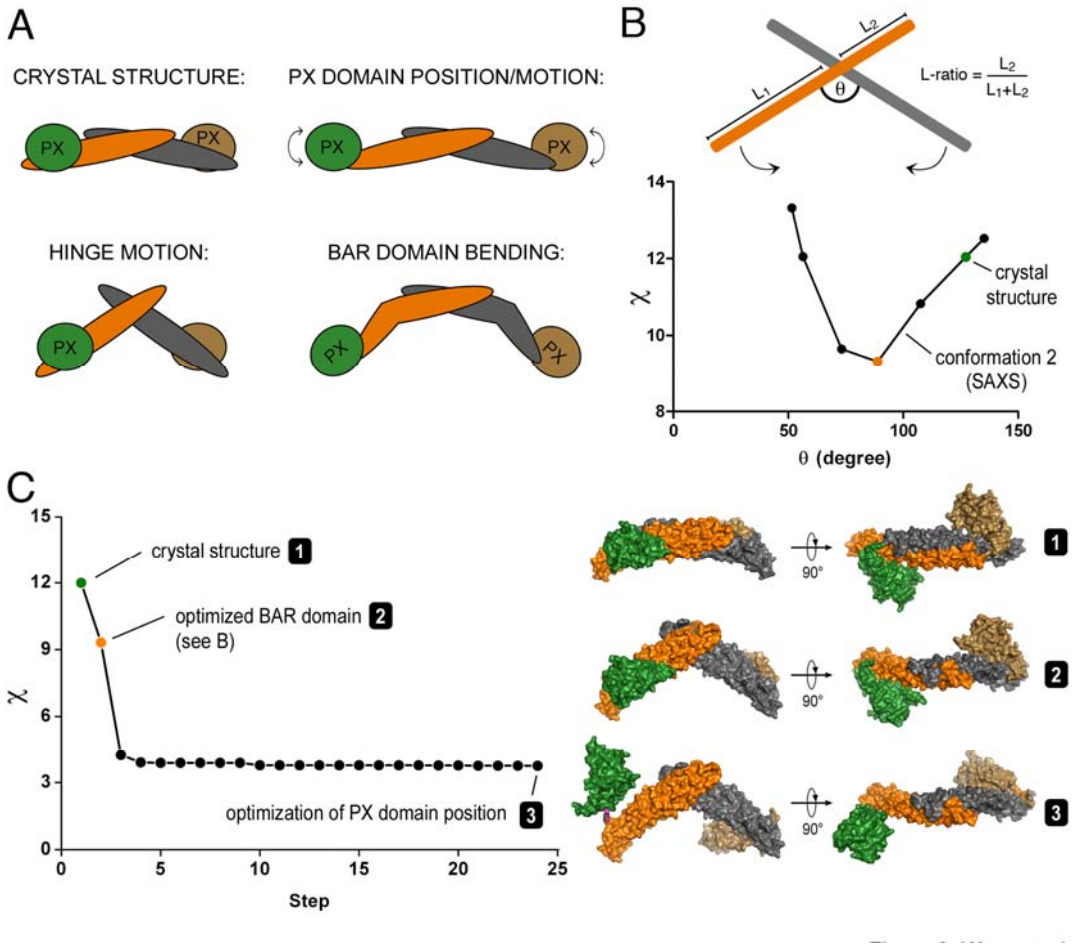


Figure 4.10. Modeling conformational changes of Snx9$^{PX\text{-}BAR}$ in solution.(A)Possible modes for structural flexibility in Snx9$^{PX\text{-}BAR}$. Cartoon presentations illustrate possible conformational changes in Snx9$^{PX\text{-}BAR}$, including changes in the position and dynamics (indicated by the arrows) of the PX domain relative to the BAR domains and bending or hinging of the BAR domain dimer.(B)Snx9$^{PX\text{-}BAR}$ curvature fitting based on experimental SAXS data. For simplicity, we created Snx9$^{PX\text{-}BAR}$ models with increased curvature manually by pivoting of the BAR domains (see diagram). Models with different θ values were scored against the experimental scattering data with χ indicating the goodness of fit. Curvatures estimated from the crystal structure of Snx9$^{PX\text{-}BAR}$ and from the SAXS-based reconstructions using models in node 2 (Figure 6C) are indicated.(C)Rigid body modeling the SAXS data for Snx9$^{PX\text{-}BAR}$ in solution. Modeling started with the crystal structure of Snx9$^{PX\text{-}BAR}$ (green; structure 1). In the first step, the curvature of the BAR domain dimer was optimized as described in (B) (orange, structure 2) with the position of the PX domain relative to the BAR domains being fixed. Modeling the SAXS data by optimizing the position of the X domain was performed using the program BUNCH (Petoukhov and Svergun, 2005). Relatively low χ values were obtained after 20 steps of simulation and allowing flexibility between the PX and BAR domains (structure 3).

CONCLUSION

Recently, SAXS studies have been successfully used to elucidate the quaternary structure and domain organization, alternative states, and conformational heterogeneity of proteins in solution (Bernado et al., 2008; Krukenberg et al., 2008; Putnam et al., 2007; Sondermann et al., 2005; Tsutakawa et al., 2007). Here, reconstructions of the solution state of $Snx9^{PX\text{-}BAR}$ in combination with a cluster analysis of the models revealed an alternative conformation, different from crystal structures with regard to curvature and inter-domain flexibility. The conformational flexibility of $Snx9^{PX\text{-}BAR}$ and structural differences between the isolated PX-BAR module and the full-length protein might also explain the differences observed in tubulation experiments (Pylypenko et al., 2007; Yarar et al., 2008). $Snx9^{PX\text{-}BAR}$ produced tubules with a diameter of 20 nm, smaller than the intrinsic BAR domain curvature observed in crystal structures, and the authors proposed a pitched association of Snx9 molecules (Pylypenko et al., 2007). The more curved conformation described here suggests an alternative model in which the BAR domain may accommodate smaller vesicles by adopting different degrees of curvature. In contrast, the full-length protein produced 40 nm tubules requiring higher protein concentration, suggesting that the SH3 domain and low complexity region of Snx9 are not only autoinhibitory but may also contribute to determine the conformation and mode of oligomerization of Snx9 (Yarar et al., 2008).

To the best of our knowledge, this is the first structural evidence that particular BAR domains exist in different states with varying degree of intrinsic curvature. An open question remains whether some BAR domain proteins adjust their structure to the curvature of the membrane and might use this mechanism to change their overall conformation, and in that sense function as curvature sensors. Further studies will be

necessary to elucidate the structures and dynamics of BAR domain-containing proteins in solution and on membranes.

REFERENCES

Adams P., Grosse-Kunstleve R., Hung L., Ioerger T., McCoy A., Moriarty N., Read R., Sacchettini J., Sauter N., Terwilliger T. (2002) PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. D Biol. Crystallogr. 58, 1948-1954.

Badour K., McGavin M.K., Zhang J., Freeman S., Vieira C., Filipp D., Julius M., Mills G.B., Siminovitch, K.A. (2007) Interaction of the Wiskott-Aldrich syndrome protein with sorting nexin 9 is required for CD28 endocytosis and cosignaling in T cells. Proc. Natl. Acad. Sci. USA 104, 1593-1598.

Bernado P., Perez Y., Svergun D.I., Pons, M. (2008) Structural characterization of the active and inactive states of Src kinase in solution by small-angle X-ray scattering. J. Mol. Biol. 376, 492-505.

Brunger A.T., Adams P.D., Clore G.M., DeLano W.L., Gros P., Grosse-Kunstleve R.W., Jiang J.S., Kuszewski J., Nilges M., Pannu, N.S., et al. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. 54 (Pt 5), 905-921.

Carlton J.G., Cullen P.J. (2005) Sorting nexins. Curr. Biol. 15, R819-820.

Childress C., Lin Q., Yang, W. (2006) Dimerization is required for SH3PX1 tyrosine phosphorylation in response to epidermal growth factor signalling and interaction with ACK2. Biochem J 394, 693-698.

Demeler B. (2005) UltraScan A Comprehensive Data Analysis Software Package for Analytical Ultracentrifugation Experiments. Modern Analytical Ultracentrifugation: Techniques and Methods., 210-229.

Farsad K., Slepnev V., Ochoa G., Daniell L., Haucke V., De Camilli, P. (2003) A putative role for intramolecular regulatory mechanisms in the adaptor function of amphiphysin in endocytosis. Neuropharmacology 45, 787-796.

Frost A., Perera R., Roux A., Spasov K., Destaing O., Egelman E.H., De Camilli P., Unger V.M. (2008) Structural basis of membrane invagination by F-BAR domains. Cell 132, 807-817.

Gallop J.L., Jao C.C., Kent H.M., Butler P.J., Evans P.R., Langen R., McMahon, H.T. (2006) Mechanism of endophilin N-BAR domain-mediated membrane curvature. EMBO J. 25, 2898-2910.

Haberg K., Lundmark R., Carlsson S.R. (2008) SNX18 is an SNX9 paralog that acts as a membrane tubulator in AP-1-positive endosomal trafficking. J. Cell Sci. 121, 1495-1505.

Henne W.M., Kent H.M., Ford M.G., Hegde B.G., Daumke O., Butler P.J., Mittal R., Langen R., Evans P.R., McMahon H.T. (2007) Structure and analysis of FCHo2 F-BAR domain: a dimerizing and membrane recruitment module that effects membrane curvature. Structure 15, 839-852.

Howard L., Nelson K.K., Maciewicz R.A.,Blobel C.P. (1999) Interaction of the metalloprotease disintegrins MDC9 and MDC15 with two SH3 domain-containing proteins, endophilin I and SH3PX1. J. Biol. Chem. 274, 31693-31699.

Itoh T., De Camilli P. (2006) BAR, F-BAR (EFC) and ENTH/ANTH domains in the regulation of membrane-cytosol interfaces and membrane curvature. Biochim. Biophys. Acta 1761, 897-912.

Jeffries C.M., Whitten A.E., Harris S.P., Trewhella J. (2008). Small-angle X-ray scattering reveals the N-terminal domain organization of cardiac myosin binding protein C. J. Mol. Biol. 377, 1186-1199.

Jones T.A., Zou J.Y., Cowan S.W., Kjeldgaard (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A 47 (Pt 2), 110-119.

Konarev P.V., Volkov V.V., Sokolova A.V., Koch M.H.J., Svergun D.I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. Journal of Applied Crystallography 36, 1277-1282.

Krukenberg K.A., Forster F., Rice L.M., Sali A., Agard D.A. (2008) Multiple Conformations of E. coli Hsp90 in Solution: Insights into the Conformational Dynamics of Hsp90. Structure 16, 755-765.

Lee S.H., Kerff F., Chereau D., Ferron F., Klug A., Dominguez, R. (2007) Structural basis for the actin-binding function of missing-in-metastasis. Structure 15, 145-155.

Li J., Mao X., Dong L.Q., Liu F., Tong, L. (2007) Crystal structures of the BAR-PH and PTB domains of human APPL1. Structure 15, 525-533.

Lin Q., Lo C.G., Cerione R.A., Yang, W. (2002) The Cdc42 target ACK2 interacts with sorting nexin 9 (SH3PX1) to regulate epidermal growth factor receptor degradation. J. Biol. Chem. 277, 10134-10138.

Lundmark R., Carlsson S.R. (2002). The beta-appendages of the four adaptor-protein (AP) complexes: structure and binding properties, and identification of sorting nexin 9 as an accessory protein to AP-2. Biochem. J. 362, 597-607.

Lundmark R., Carlsson, S.R. (2003) Sorting nexin 9 participates in clathrin-mediated endocytosis through interactions with the core components. J. Biol. Chem. 278, 46772-46781.

Madsen K.L., Eriksen J., Milan-Lobo L., Han D.S., Niv M.Y., Ammendrup-Johnsen I., Henriksen U., Bhatia V.K., Stamou D., Sitte H.H., et al. (2008) Membrane localization is critical for activation of the PICK1 BAR (Bin/amphiphysin/Rvs) domain. Traffic 34(3): 432-441

Masuda M., Takeda S., Sone M., Ohki T., Mori H., Kamioka Y., Mochizuki, N. (2006) Endophilin BAR domain drives membrane curvature by two newly identified structure-based mechanisms. EMBO J 25, 2889-2897.

Mattila P.K., Pykalainen A., Saarikangas J., Paavilainen V.O., Vihinen H., Jokitalo E., Lappalainen P. (2007) Missing-in-metastasis and IRSp53 deform PI(4,5)P2-rich membranes by an inverse BAR domain-like mechanism. J Cell Biol 176, 953-964.

Millard T.H., Bompard G., Heung M.Y., Dafforn T.R., Scott D.J., Machesky L.M., Futterer, K. (2005) Structural basis of filopodia formation induced by the IRSp53/MIM homology domain of human IRSp53. EMBO J. 24, 240-250.

Mylonas E., Svergun, D.I. (2007) Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. Journal of Applied Crystallography 40, s245-249.

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

Peter B.J., Kent H.M., Mills I.G., Vallis Y., Butler P.J., Evans P.R., McMahon H.T. (2004) BAR domains as sensors of membrane curvature: the amphiphysin BAR structure. Science 303, 495-499.

Petoukhov M.V., Monie T.P., Allain F.H., Matthews S., Curry S., Svergun D.I. (2006). Conformation of polypyrimidine tract binding protein in solution. Structure 14, 1021-1027.

Petoukhov M.V., Svergun, D.I. (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data. Biophysical Journal 89, 1237-1250.

Porod, G. (1982). General Theory. In Small-Angle X-Ray Scattering, O. Kratky, ed. (Academic Press, London), pp. 17-51.

Putnam C.D., Hammel M., Hura G.L., Tainer, J.A. (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. Q. Rev. Biophys. 40, 191-285.

Pylypenko O., Lundmark R., Rasmuson E., Carlsson S.R., Rak, A. (2007) The PX-BAR membrane-remodeling unit of sorting nexin 9. EMBO J. 26, 4788-4800.

Ringstad N., Nemoto Y., De Camilli P. (1997). The SH3p4/Sh3p8/SH3p13 protein family: binding partners for synaptojanin and dynamin via a Grb2-like Src homology 3 domain. Proc. Natl. Acad. Sci. USA 94, 8569-8574.

Shimada A., Niwa H., Tsujita K., Suetsugu S., Nitta K., Hanawa-Suetsugu K., Akasaka R., Nishino Y., Toyama M., Chen L., et al. (2007) Curved EFC/F-BAR-domain dimers are joined end to end into a filament for membrane invagination in endocytosis. Cell 129, 761-772.

Shin N., Lee S., Ahn N., Kim S.A., Ahn S.G., Yong Park Z., Chang S. (2007) Sorting nexin 9 interacts with dynamin 1 and N-WASP and coordinates synaptic vesicle endocytosis. J. Biol. Chem. 282, 28939-28950.

Sondermann H., Nagar B., Bar-Sagi D., Kuriyan J. (2005) Computational docking and solution x-ray scattering predict a membrane-interacting role for the histone domain of the Ras activator son of sevenless. Proc. Natl. Acad. Sci. USA 102, 16632-16637.

Soulet F., Yarar D., Leonard M., Schmid, S.L. (2005) SNX9 regulates dynamin assembly and is required for efficient clathrin-mediated endocytosis. Mol Biol Cell 16, 2058-2067.

Svergun D., Barberato C., Koch, M.H.J. (1995) CRYSOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. Journal of Applied Crystallography 28, 768-773.

Svergun, D.I. (1992). Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria. Journal of Applied Crystallography 25, 495-503.

Svergun, D.I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys. J. 76, 2879-2886.

Svergun D.I., Petoukhov M.V., Koch M.H. (2001) Determination of domain structure of proteins from X-ray solution scattering. Biophys. J. 80, 2946-2953.

Taraban M., Zhan H., Whitten A.E., Langley D.B., Matthews K.S., Swint-Kruse L., Trewhella J. (2008). Ligand-induced conformational changes and conformational dynamics in the solution structure of the lactose repressor protein. J. Mol. Biol. 376, 466-481.

Terwilliger T.C., Berendzen, J. (1999) Automated MAD and MIR structure solution. Acta Crystallogr. D Biol. Crystallogr. 55 (Pt 4), 849-861.

Tsutakawa S.E., Hura G.L., Frankel K.A., Cooper P.K., Tainer, J.A. (2007) Structural analysis of flexible proteins in solution by small angle X-ray scattering combined with crystallography. J. Struct. Biol. 158, 214-223.

Volkov V.V., Svergun, D.I. (2003) Uniqueness of ab initio shape determination in small-angle scattering. Journal of Applied Crystallography 36, 860-864.

Wall M.E., Gallagher S.C., Trewhella, J. (2000) Large-scale shape changes in proteins and macromolecular complexes. Annu. Rev. Phys. Chem. 51, 355-380.

Worby C.A., Dixon, J.E. (2002) Sorting out the cellular functions of sorting nexins. Nat. Rev. Mol. Cell Biol. 3, 919-931.

Worby C.A., Simonson-Leff N., Clemens J.C., Huddler D. Jr., Muda M., Dixon, J.E. (2002) Drosophila Ack targets its substrate, the sorting nexin DSH3PX1, to a protein complex involved in axonal guidance. J. Biol. Chem. 277, 9422-9428.

Yarar D., Surka M.C., Leonard M.C., Schmid, S.L. (2008). SNX9 activities are regulated by multiple phosphoinositides through both PX and BAR domains. Traffic 9, 133-146.

Yarar D., Waterman-Storer C.M., Schmid, S.L. (2007). SNX9 couples actin assembly to phosphoinositide signals and is required for membrane remodeling during endocytosis. Dev. Cell 13, 43-56.

Yeow-Fong L., Lim L., Manser E. (2005) SNX9 as an adaptor for linking synaptojanin-1 to the Cdc42 effector ACK1. FEBS Lett 579, 5040-5048.

Zhu G., Chen J., Liu J., Brunzelle J.S., Huang B., Wakeham N., Terzyan S., Li X., Rao Z., Li G., Zhang, X.C. (2007) Structure of the APPL1 BAR-PH domain and characterization of its interaction with Rab5. EMBO J. 26, 3484-3493.

Zimmerberg J., Kozlov, M.M. (2006) How proteins produce cellular membrane curvature. Nat. Rev. Mol. Cell Biol. *7*, 9-19.

CHAPTER 5

STRUACTURE BASIS OF CALCIUM SENSING BY GCAMP2

ABSTRACT

Genetically encoded Ca^{2+} indicators are important tools that enable the measurement of Ca^{2+} dynamics in a physiologically relevant context. GCaMP2, one of the most robust indicators, is a circularly permutated EGFP (cpEGFP)/M13/Calmodulin (CaM) fusion protein, that has been successfully used for studying Ca^{2+} fluxes *in vivo* in the heart and vasculature of transgenic mice. Here we describe crystal structures of bright and dim states of GCaMP2 that reveal a sophisticated molecular mechanism for Ca^{2+} sensing. In the bright state, CaM stabilizes the fluorophore in an ionized state similar to that observed in EGFP. Mutational analysis confirmed critical interactions between the fluorophore and elements of the fused peptides. Solution scattering studies indicate that the Ca^{2+}-free form of GCaMP2 is a compact, pre-docked state, suggesting a molecular basis for the relatively rapid signaling kinetics reported for this indicator. These studies provide a structural basis for the rational design of improved Ca^{2+}-sensitive probes.

INTRODUCTION

Genetically encoded Ca^{2+} indicators (GECIs) enable the monitoring of intracellular signaling events in defined cell lineages within complex multicellular tissues, without disrupting cell-cell contacts or permeabilizing cell membranes (Ledoux et al., 2008; Mao et al., 2008; Miyawaki et al., 1999; Miyawaki et al., 1997; Nausch et al., 2008; Roell et al., 2007; Tallini et al., 2007; Tallini et al., 2006). Several distinct strategies have been employed to achieve the goals of a bright, fast, high signal-to-noise indicator that functions under *in vivo* conditions, including the modification of the green fluorescent protein (GFP) family from the *Aequorea victoria* (Kotlikoff, 2007; Mank and Griesbeck, 2008; Mao et al., 2008; Palmer and Tsien, 2006; Pologruto et al., 2004; Reiff et al., 2005).

Intact GFP has a cage-like structure with 11 β-sheets forming a barrel around a helix containing the fluorescent moiety (Brejc et al., 1997; Ormo et al., 1996; Yang et al., 1996). Protected from solvent access and coordinated by a number of water molecules and polar side chains contributed by the cage, three consecutive residues in the central helix (Ser-Tyr-Gly) undergo post-translational cyclization to form the fluorophore (Tsien, 1998). Spectroscopic and structural studies have illuminated the molecular mechanism underlying GFP fluorescence (Tsien, 1998). Wild-type GFP (class 1 GFP) has a major excitation peak at a wavelength of 395 nm and a minor peak at 475 nm, resulting from a protonated and a deprotonated fluorophore, respectively, with distinct emission properties (503 nm *vs* 508 nm). Neutralizing effects of close-by residues maintain the fluorophore in its protonated state. The fluorophore is subject to reversible ionization facilitated by a surrounding hydrogen bond network involving interactions with the cage wall. Mutation of Ser-65 of the chromophore to threonine yielded an enhanced GFP (EGFP) with a deprotonated fluorophore and a single excitation peak at 489 nm (GFP-S65T; class 2 GFP) (Ormo et al., 1996).

GFP is amenable to large structural rearrangements without disrupting its basic fluorescent properties, including circular permutation (Baird et al., 1999). Progressive improvements in sensor performance have been made using a central circularly permutated EGFP (cpEGFP) moiety flanked by the M13 helix of myosin light chain and Calmodulin (CaM) at the N- and C-terminus, respectively (GCaMP2 or pericam) (Nagai et al., 2001; Nagai et al., 2004; Nakai et al., 2001; Souslova et al., 2007; Tallini et al., 2007). These molecules exploit Ca^{2+}-dependent intramolecular conformational changes to control the efficiency of fluorophore function. However, the structural basis for Ca^{2+}-dependent changes in fluorescence is not understood constraining the rational design and optimization of these proteins.

The construction and optimization of novel Ca^{2+} indicators or other molecular sensors based on this strategy currently requires extensive random mutagensis and screening due to limited understanding of the working mechanism of the current molecules. Here, we determined structures of GCaMP2 in multiple states revealing a surprisingly sophisticated network of interactions between the cpEGFP and CaM moieties that are responsible for the high functionality of this Ca^{2+} probe. These studies will guide the structurally motivated improvement of current molecules and the design of novel indicators.

MATERIALS AND METHODS

Protein expression and purification

The coding regions corresponding to GCaMP2, EGFP, and cpEGFP were amplified by standard PCR and were cloned into the pRSET expression plasmid (Nakai et al., 2001). GCaMP2ΔRSET was cloned into a modified pET28a expression plasmid (Novagen) yielding N-terminally hexahistidine-tagged SUMO fusion proteins. The hexahistidine-tagged SUMO-moiety was cleavable using the protease from Ulp-1 from *S. cerevisiae*.

Transformed *E. coli* cells BL21(DE3) (Novagen) were grown in TB medium supplemented with 50 mg/l antibiotics at 37°C. At a cell density corresponding to an

absorbance of 1.0 at 600 nm the temperature was reduced to 18°C, and protein production was induced with 1 mM IPTG. Proteins were expressed for 12–16 hr. Cells were collected by centrifugation, resuspended in NiNTA buffer A (25 mM Tris-Cl pH 8.2, 500 mM NaCl and 20 mM imidazole). After cell lysis by sonication, cell debris was removed by centrifugation at 40,000 x g for 1 hr at 4°C. Clear lysates were loaded onto HisTrap NiNTA columns (GE Healthcare) equilibrated in NiNTA buffer A. The resin was washed with 20 column volumes NiNTA buffer A, and proteins were eluted in a single step with NiNTA buffer A supplemented with 500 mM imidazole. GCaMP2ΔRSET was incubated with SUMO protease ULP-1 at 4°C overnight for removal of the hexahistidine-SUMO tag, and the cleaved protein was collected in the flow-through during NiNTA affinity chromatography. Proteins were further subjected to size exclusion chromatography on a Superdex200 column (16/60; GE Healthcare) equilibrated in gel filtration buffer (100 mM NaCl, 25 mM HEPES pH 7.4).

Ca^{2+}-free proteins were produced by incubation in gel filtration buffer supplemented with 10 mM EGTA prior the size exclusion column. Fractions containing protein were pooled and concentrated on a Centricon ultrafiltration device (10 kDa cutoff; Millipore). Protein aliquots were frozen in liquid nitrogen and stored at -80°C.

Point mutations were generated using the QuikChange XL Mutagenesis Kit (Stratagene) following the manufacturer's instructions. Expression and purification of mutant proteins was identical to the procedure described for parent proteins. Mutant proteins showed size exclusion profiles in the presence or absence of Ca^{2+} indistinguishable from original proteins.

Analytical size exclusion chromatography was carried out on a Superdex200 column (10/300; GE Healthcare) equilibrated in gel filtration buffer at a protein

concentration of 0.2 mM (initial concentration). EGTA ("-Ca^{2+}"; 10 mM) or $CaCl_2$ (1 mM) was present where indicated.

Crystallization, X-ray data collection, and structure solution

Crystals were obtained by hanging drop vapor diffusion by mixing equal volumes of protein (~10-30 mg/ml) and reservoir solution at 20°C (if not indicated otherwise). For cpEGFP, the reservoir solution comprised 0.1 M Tris pH 8.5, 2.4 M di-ammonium hydrogen phosphate. Crystals for monomeric GCaMP2ΔRSET•Ca^{2+} grew in 20% PEG8000, 0.1 M HEPES pH 7.5, 1.6 M ammonium sulfate, and 1 mM $CaCl_2$ at 4°C. The reservoir solution for crystallization of dimeric GCaMP2•Ca^{2+} consisted of 0.2 M sodium formate, 20% PEG 3350, 10 mM $MgCl_2$, and 1 mM $CaCl_2$ (dimer 1), or 0.2 M $MgCl_2$, 0.1 M Bis-Tris pH 6.5, 25% PEG3350, and 1 mM $CaCl_2$ (dimer 2). During preparation of this manuscript a third crystallization condition has been reported, producing crystals within the same space group and similar unit cell constants (Rodriguez Guilbe et al., 2008). Crystals producing the structure of dimer 1 were grown from partially dimeric protein (as judged by size exclusion chromatography and analytical ultracentrifugation). The structure of dimer 2 was obtained from EGTA-treated, monomeric protein. All crystals were cryo-protected using crystallization solutions supplemented with 20% xylitol, frozen in liquid nitrogen, and kept at 100 K during data collection.

Crystallographic statistics for data collection are shown in Table 1. Data sets were collected using synchrotron radiation at the Cornell High Energy Synchrotron Source (CHESS, Ithaca, beamline A1, wavelength 0.977Å). Data reduction was carried out with the software package HKL2000 (Otwinowski and Minor, 1997). Phases were obtained from molecular replacement using the software package PHENIX (Adams et al., 2002) with structures of GFP-S65T (PDB code: 1EMA)

(Ormo et al., 1996) and Ca^{2+}-bound calmodulin (PDB code: 1MXE) (Clapperton et al., 2002) as search models. Manual refinement in COOT (Emsley and Cowtan, 2004) and minimization using PHENIX (Adams et al., 2002) yielded the final models with good geometry (with all residues being in allowed regions of the respective Ramachandran plots). Illustrations were made in Pymol (DeLano Scientific). Buried surface areas were calculated using the program CNS (Brunger et al., 1998). Electrostatic potentials were calculated using the program APBS (Baker et al., 2001).

Small angle X-ray scattering (SAXS) and SAXS-based shape reconstruction

SAXS data were collected at the Cornell High Energy Synchrotron Source (CHESS, Ithaca, beamline G1) at an electron energy of 8KeV covering the range of momentum transfer $0.019 < S < 0.22$ $Å^{-1}$. Buffer exchange into Ca^{2+}-free (gel filtration buffer supplemented with 10 mM EGTA) and Ca^{2+}-saturated (gel filtration buffer supplemented with 1 mM $CaCl_2$) buffer was carried out by size exclusion chromatography. Protein samples and buffer controls were centrifuged at 13,200 x g for 20 min before data acquisition. Scattering data were collected in triplicates at protein concentrations between 5-20 mg/ml with exposure times of 5 seconds. Buffer background scattering was collected from gel filtration buffers used during buffer exchange. Background-subtraction, averaging, and scaling were carried out using the program PRIMUS (Konarev et al., 2003), and data was analyzed using the programs GNOM and CRYSOL (Svergun et al., 1995; Svergun, 1992). Only scattering data with $S_{max}*R_g<1.3$, computed from Guinier plots at low angle regions, was considered for further analysis. Kratky plots, experimental molecular weight determinations, and normalized scattering intensities were used to assess the folded-state of the proteins and overall data quality using the program IGOR PRO (version 5.04B). Distance

distribution functions (P[r]) and D_{max} were determined using the program GNOM (Svergun, 1992).

Ab initio free atom modeling was performed using the program DAMMIN (Svergun, 1999) with the scattering curve as the fitting target. Forty independent simulations are carried out for each protein. Superposition, averaging, and filtering using the program DAMAVER (Volkov and Svergun, 2003) yielded shape reconstructions. Calculated scattering curves and goodness of fit (χ) were calculated using the program CRYSOL (Svergun et al., 1995).

Analytical ultracentrifugation

Sedimentation velocity experiments were carried out using an XL-I analytical ultracentrifuge (Beckman Coulter) equipped with an AN-60 Ti rotor. Proteins (5-15 μM) were diluted in AUC buffer (100 mM NaCl, 25 mM HEPES pH 7.4, and 1 mM TCEP) and were analyzed at a centrifugation speed of 40,000 rpm. Data collection was carried out at 280 nm, followed by data analysis using the program SedFit (version 11.0) using default parameters.

Absorbance and fluorescence spectroscopy

Absorbance spectra (320-500 nm; 10 mm path length) of purified proteins (20 μM) were recorded in duplicate on a DU730 UV/Vis spectrophotomer (Beckman Coulter) at 25ºC. Emission spectra (495-600 nm) were recorded in triplicate on a fluorescence spectrophotomer (Photon Technology International) at a protein concentration of 100 nM and an excitation wavelength of 488 nm. The Ca^{2+} calibration buffer kit (Biotium, Inc) was used to create accurate assay conditions with regard to Ca^{2+} concentration. Ca^{2+} sensitivity of purified proteins was determined by measuring the fluorescence intensity at 508 nm under Ca^{2+}-free (10 mM EGTA) and

Ca^{2+}-saturated conditions (40 μM Ca^{2+}). Sensitivity to changes in pH was measured at a protein concentration of 10 μM using buffers containing sodium acetate (pH 4.5-5.0), MES (pH 5-6.5), HEPES (pH 6.5-8), or CHES (pH 8.5-10). pK_a values were calculated by fitting the experimental data to the Hill equation.

Pair-wise distance matrix analysis

A distance matrix D is constructed by calculating the pair-wise distances between the main chain α-carbon atoms within GFP-S65T (PDB code: 1EMA) (Ormo et al., 1996), cpEGFP and GCAMP2ΔRSET•Ca^{2+}, respectively. A difference matrix Δ was calculated by subtracting the distance matrix calculated for the structures of cpEGFP or GCAMP2ΔRSET•Ca^{2+} from the matrix calculated using the structure of GFP-S65T. Two matrices used in such analysis have to be of the same dimension (same residue number), and structures were manipulated accordingly by omitting non-homologous residues.

The difference matrices were regularized by applying a Z-score analysis (ΔZ, see algorithm below). The elements with high Z-scores (higher than 2) were considered of significant difference between two structures. In this approach, each atom is described by a vector composed of the relative distances toward all the other atoms in the structure. Therefore it is very sensitive to subtle conformational change between similar structures. Moreover, the calculation of difference matrices is reference-frame independent and bias-free, not relying on a structural alignment of two structures.

Algorithm:

Calculation of a distance matrix from PDB coordinates of a particular structure:

$D_{ij} = d(C_\alpha{}^i, C_\alpha{}^j)$.

Calculation of difference matrices for two structures: $\Delta_{ij} = Abs(D_{ij} - D_{ij}^{EGFP})$.

Calculation of Z-score-regularized matrices: $\Delta_{ij}^{z} = \frac{Abs(\Delta_{ij} - \overline{\Delta}_{ij})}{\sigma(\Delta_{ij})}$ where $\sigma(\Delta)$ and $\overline{\Delta}$ are the standard deviations and average values of all the elements in the matrix Δ.

RESULTS AND DICUSSIONS

Crystal structures of cpEGFP and GCaMP2•Ca^{2+}

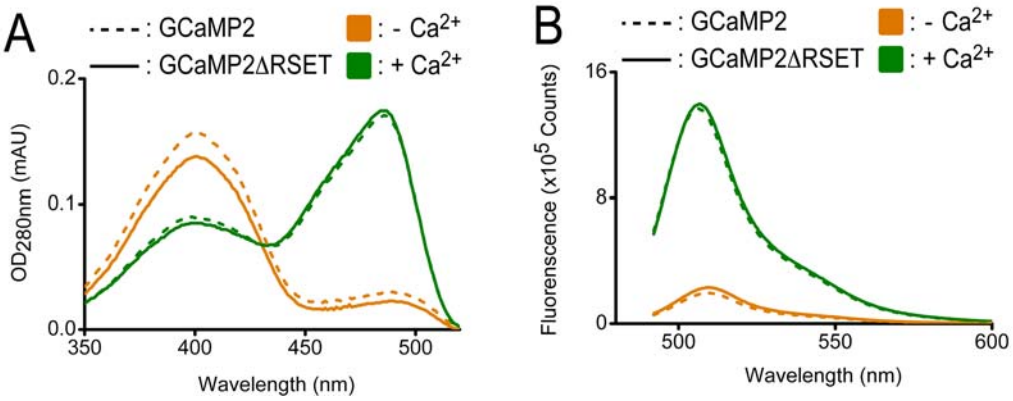


Figure 5.1. Spectroscopic comparison of GCaMP2 and GCaMP2ΔRSET. (A) Absorbance spectra of GCaMP2 (dashed lines) and GCaMP2ΔRSET (solid lines) in the presence (green) and absence (orange) of Ca^{2+}. Absorbance spectra (320-500 nm; 10 mm path length) of purified proteins (20 μM) were recorded at 25ºC. (B) Emission spectra of GCaMP2 and GCaMP2ΔRSET in the presence and absence of Ca^{2+}. Emission spectra (495-600 nm) were recorded at a protein concentration of 100 nM and an excitation wavelength of 473 nm at 25ºC. The Ca^{2+} calibration buffer kit (Biotium, Inc) was used to create accurate assay condition with regard to Ca^{2+} concentration.

Initial studies were carried out with the original GCaMP2 protein containing the N-terminal hexahistidine tag and RSET module (Nakai et al., 2001) (Figure 5.1). Removal of the purification tag and RSET motif yielded crystals of the bright, Ca^{2+}-bound state of GCaMP2. The N-terminal truncations had no effect on basic properties and efficiency of GCaMP2 at room temperature (expression and purification, spectroscopic properties, and Ca^{2+} response; Figure 5.1.A-B) suggesting that these domains mainly affect GCaMP2's folded state under *in vivo* conditions (Tallini et al.,

2006). If not indicated otherwise, the truncated form, GCaMP2ΔRSET, was used in subsequent experiments.

Crystallization trials carried out with native GCaMP2ΔSET and cpEGFP yielded crystal structures of monomeric GCaMP2ΔRSET•Ca^{2+} and the isolated cpEGFP module, respectively (Figure 5.2.B and C;Table 5.1 for details). GCaMP2ΔRSET•Ca^{2+} has a compact structure with the M13 helix forming an intramolecular interaction with the CaM moiety (Figure 5.2.C). The cpEGFP module in GCaMP2 and the isolated domain are very similar in structure to GFP-S65T (rms deviation of 0.48 and 0.22 Å, respectively). To further assess the folded state and conformation of the cpEGFP module in GCaMP2 and in isolation in an unbiased way, distance difference analyses based on corresponding Cα positions were performed (Figure 5.2.D; see Supplemental Data for details). The comparison of cpEGFP (lower part of the matrix) and GCaMP2ΔRSET (upper half of the matrix) with the structure of GFP-S65T (Ormo et al., 1996) produced featureless matrices indicating that in both structures, cpEGFP is properly folded, indistinguishable in ternary structure to intact GFP (blue: distances showing minimal difference between two structures). Minor conformational changes were restricted to a single loop (residues 66-72 in GCaMP2, corresponding to residues 154-160 in GFP) and the N-terminal three residues of the first β-sheet of the C-EGFP (Figure 5.2) that connect to the M13 fragment (residues 62-64 in GCaMP2/cpEGFP) (Figure5.2). The latter region contributed to an intact β-barrel in cpEGFP, but is peeled away when the M13 helix was bound to the CaM domain in GCaMP2ΔSET•Ca^{2+}, producing a hole in the cpEGFP cage (Figure 5.2.B-D; see below).

Figure 5.2. Crystal structures of GCaMP2•Ca^{2+} and cpEGFP. (A) Domain organization of GCaMP2 and truncated derivatives. A schematic presentation of the GCaMP2 fusion protein is shown. The color scheme introduced here is maintained throughout the manuscript. Residue numbering for circularly permutated EGFP (cpEGFP) and GCaMP2ΔRSET follows the sequence of GCaMP2. (B) Crystal structure of the isolated cpEGFP moiety. The C-terminal fragment of C-EGFP is colored in light green, the N-terminal fragment is colored in dark green. Two orthogonal views are shown. (C) Crystal structure of monomeric GCaMP2ΔRSET in its Ca^{2+}-bound state. Crystals were grown in the presence of 1 mM Ca^{2+}. Two orthogonal views are shown. The M13 helix is shown in blue, and the calmodulin (CaM) domain is shown in red. The cpEFGP is colored as described in (B). (D) Comparison of crystal structures of GCaMP2, cpEGFP and GFP-S65T. Distance difference matrices based on Cα positions were used to compare the conformation of cpEGFP in isolation (bottom-right triangle) and as part of GCaMP2 (top-left triangle) with the structure of GFP-S65T (PDB code: 1EMA; see Supplemental Data for details). Difference matrices were regularized using a Z-score analysis and color-coded accordingly. Each entry in the matrix depicts the difference in distance between corresponding Cα atoms in the two structures. Distances that show little change are blue. Red entries represent distances that are significantly different in the two structures.

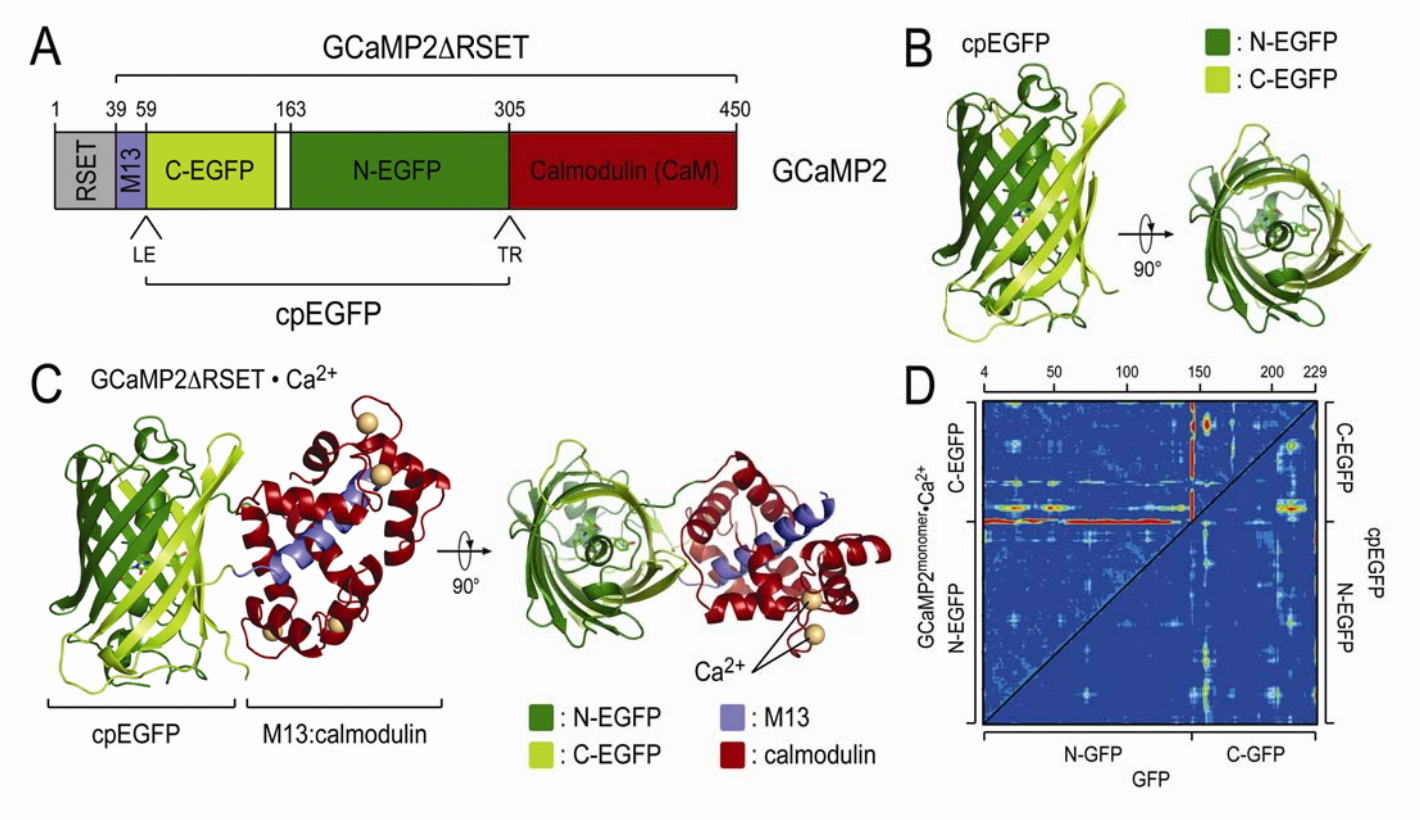

A
GCaMP2ΔRSET
1
39
59
163
305
450
RSET
M13
C-EGFP
N-EGFP
Calmodulin (CaM)
GCaMP2
LE
TR
cpEGFP
B
cpEGFP
: N-EGFP
: C-EGFP
90°
C
GCaMP2ΔRSET • Ca2+
90°
Ca2+
cpEGFP
M13:calmodulin
: N-EGFP
: C-EGFP
: M13
: calmodulin
D
4
50
100
150
200
229
GCaMP2monomer•Ca2+
C-EGFP
N-EGFP
C-EGFP
N-EGFP
cpEGFP
N-GFP
C-GFP
GFP

Table 5.1. Data collection and refinement statistics

	cpEGFP	GCaMP2ΔRSET Ca^{2+}-bound	GCaMP2 dimer Ca^{2+}-bound (1)	GCaMP2 dimer Ca^{2+}-bound (2)
Data collection[a, b]				
Space group	$P2_12_12_1$	$P4_12_12$	C2	C2
Cell dimensions				
a, *b*, *c* (Å)	51.1, 62.2, 69.5	122.0, 122.0, 97.8	127.9, 47.3, 69.6	127.5, 47.0, 67.9
. .. .. . (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 100.7, 90.0	90.0, 100.9, 90.0
Resolution (Å)	50-1.45 (1.50-1.45)	50-2.00 (2.07-2.00)	50-1.75 (1.81-1.75)	50-2.60 (2.69-2.60)
R_{sym}	4.2 (13.2)	12.7 (71.0)	4.0 (21.1)	14.2 (44.1)
I / *I*	43.3 (9.9)	23.6 (3.7)	35.5 (5.0)	8.0 (2.3)
Completeness (%)	99.3 (93.2)	100.0 (100.0)	100.0 (99.9)	98.0 (85.5)
Measured reflections	302,513	881,335	188,674	41,030
Unique reflections	39,659	50,531	41,444	12,458
Redundancy	7.6 (4.1)	17.4 (15.2)	4.6 (4.3)	3.3 (2.6)
Refinement[c]				
Resolution (Å)	50-1.45	50-2.00	50-1.75	50-2.60
R_{work} / R_{free}[d]	13.2/17.5	16.3/18.7	16.8/21.4	22.4/27.0
Free R test set size (#/%)	3171 / 8.1	4821 / 10.0	4076 / 10.1	1167 /10.1
No. atoms				
Protein	1851	3194	3158	3054
Ligand/ion	-	4 (Ca^{2+})	4 (Ca^{2+})	4 (Ca^{2+})
Water	383	453	518	55
B-factors				
Protein	13.40	28.19	23.95	28.91
Ligand/ion	n.a.	28.93	16.29	44.65
Water	34.70	34.11	38.78	24.18
R.m.s deviations				
Bond lengths (Å)	0.011	0.008	0.019	0.003
Bond angles (°)	1.510	1.126	1.664	0.739

[a] Values as defined in SCALEPACK (Otwinowski and Minor, 1997).
[b] Highest resolution shell is shown in parenthesis.
[c] Values as defined in PHENIX (Adams et al., 2002). All structures were solved by Molecular Replacement.
[d] No sigma cutoffs.

Fluorophore coordination in GCaMP2 and cpEGFP

Structural and spectroscopic studies have identified hallmarks of the distinct states of GFP manifested in the coordination of the fluorophore (Brejc et al., 1997; Yang et al., 1996). In the protonated state characterized by a major absorbance peak at 395 nm, hydrogen bonds between Glu-222, Ser-205 and the fluorophore contribute to the neutralization of its phenolic oxygen. In the state with an absorbance maximum at 489 nm (e.g. GFP-S65T or EGFP), an anionic phenolate is stabilized by interactions with Thr-203 and His-148 of the β-barrel (Brejc et al., 1997; Ormo et al., 1996) (Figure 5.3.A, bottom).

The interaction between the fluorophore and the threonine residue is preserved in the structure of Ca^{2+}-bound GCaMP2ΔRSET (Thr-116 in GCaMP2ΔRSET; Figure 5.3A, top). Additional hydrogen bonds with two water molecules complete the sp3 hybridization of the phenolic oxygen, indicative of an ionic, bright fluorophore. In contrast, the structure of cpEGFP resembles features of a protonated fluorophore, as evidenced by the loss of a hydrogen bond between Thr-116 and the phenolic oxygen of the fluorophore, and the presence of a hydrogen bond between Glu-135 and Ser-118 (Glu-222 and Ser-205 in GFP-S65T) (Figure 5.3.A, middle). The fluorophore itself and other residues involved in coordination and maturation of the fluorophore (residues Gln-252 and Arg-254 in GCaMP2; residues 94 and 96 in GFP-S65T) adopt very similar conformations in all the structures (Figure 5.4.A) (Ormo et al., 1996).

The structural analysis is consistent with emission and excitation properties of the proteins. GCaMP2ΔRSET behaves similarly to the parent protein GCaMP2. In the absence of Ca^{2+}, its fluorescence intensity is low with an absorbance maximum at 399 nm, indicative of a neutral fluorophore (Figure 5.3.B). Upon Ca^{2+} binding, the absorbance maximum shifts to 488 nm, approximately a 5-fold increase at this wavelength compared to the apo-state, accompanied with a 7.5-fold increase in

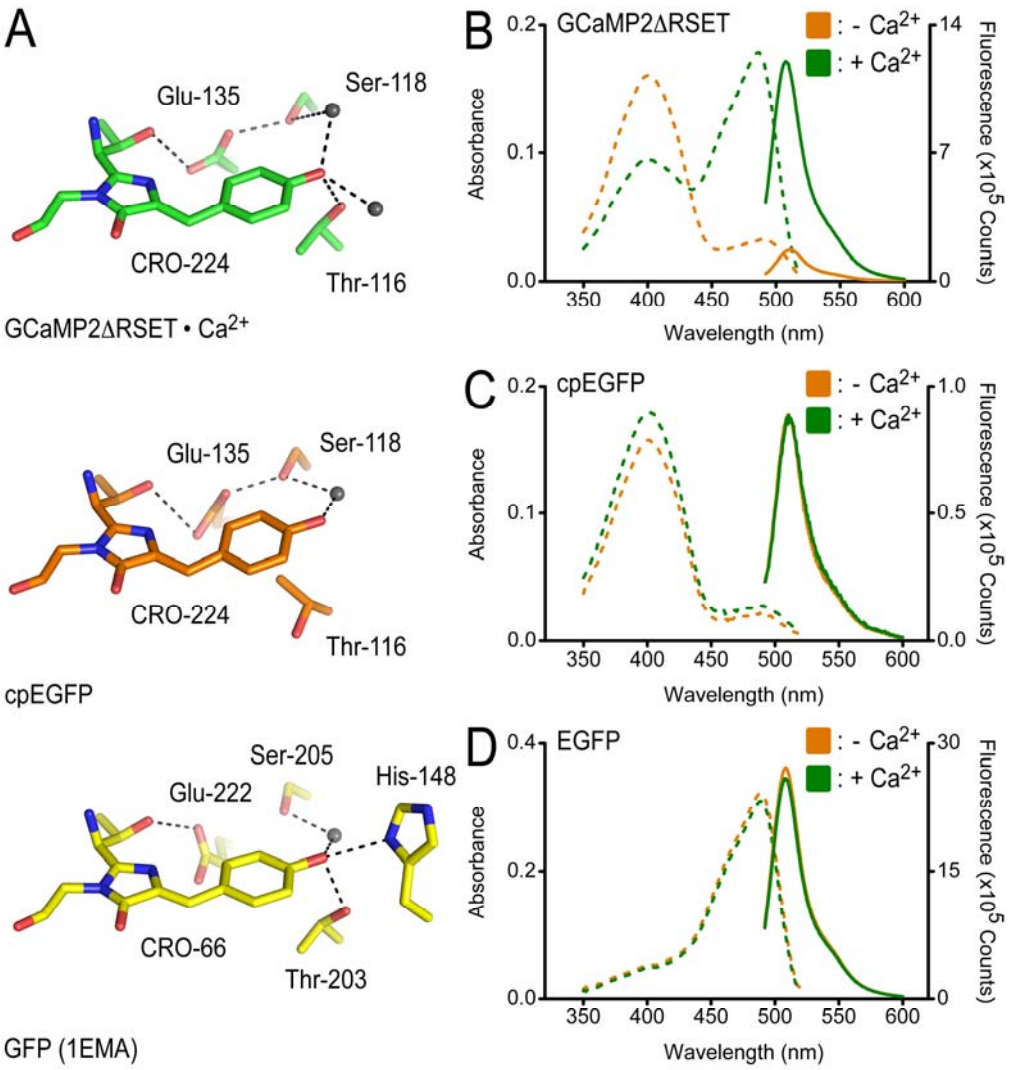


Figure 5.3. Coordination of the fluorophore in structures of GCaMP2 and cpEGFP. (A) Close-up view of the fluorophore-interacting residues in GCaMP2ΔRSET•Ca^{2+} (top), cpEGFP (middle), and GFP-S65T (bottom). Residue labeling is according to GCaMP2 numbering, except labeling of residues in the structure of GFP-S65T (PDB code: 1EMA) (Ormo et al., 1996). Hydrogen bonds are shown as dashed lines. (B) Spectroscopic properties of GCaMP2ΔRSET. Absorbance (dashed lines) and fluorescence emission (solid lines) spectra were measured at 25ºC in Ca^{2+}-free (10 mM EGTA; orange) buffer or in the presence of Ca^{2+} (40 μM; green). See Material and Methods for experimental details. (C) Spectroscopic properties of cpEGFP. Experimental conditions were identical to (B). (D) Spectroscopic properties of EGFP. Experimental conditions were identical to (B).

fluorescence intensity measured at 507 nm (Figure 5.3.B). The absolute increase in fluorescence is slightly higher than the one reported previously for GCaMP2 (Hendel et al., 2008; Tallini et al., 2006), possibly due to the usage of a homogenously monomeric sample (see below). In contrast, cpEGFP displays Ca^{2+}-independent absorbance and emission properties and resided in a constitutively protonated state with an absorbance maximum at 399 nm and low fluorescence intensity (Tsien, 1998) (Figure 5.3.C). Consistent with its high fluorescence and deprotonated fluorophore observed in the crystal structure, EGFP absorbed maximally at a peak wavelength of 488 nm both in the presence and absence of Ca^{2+} (Figure 5.3.D). These results suggest that the Ca^{2+}-dependent change in fluorescence derives from the fact that the destabilized fluorophore in cpEGFP is protonated at physiological pH, and that Ca^{2+} binding to calmodulin, and the attendant interaction with the M13 peptide, restores the deprotonated fluorophore structure, resulting in a marked increase in GFP fluorescence.

Direct contacts between cpEGFP and CaM determine the bright state

We next examined the molecular basis for the fluorophore stabilization described above. The Ca^{2+}-bound CaM module has a tight grip on the M13 helix, bringing the CaM and cpEGFP moieties into close proximity with a buried surface area of 1581 $Å^2$ (4287 $Å^2$ including M13) at the interface (Figure 5.5.A). It is interesting to note that the CaM domain forms a ring-shaped seal around a hole in the cpEGFP cage that is present due to the circular permutation and partially exposes the fluorophore to solvent (Figure 5.5.A), a positioning that would enable protection of the exposed residues.

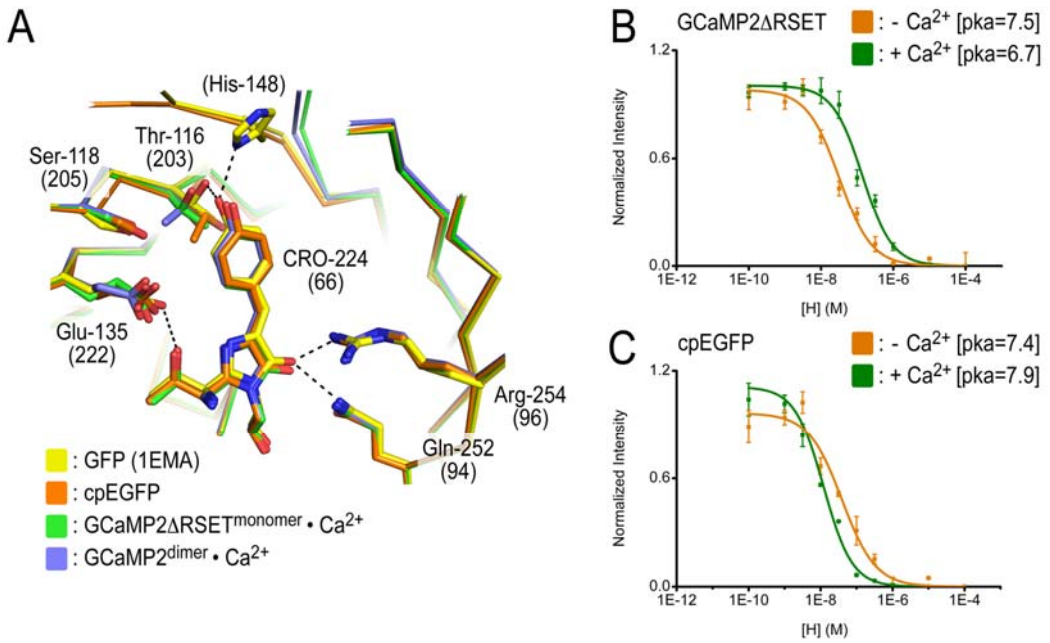
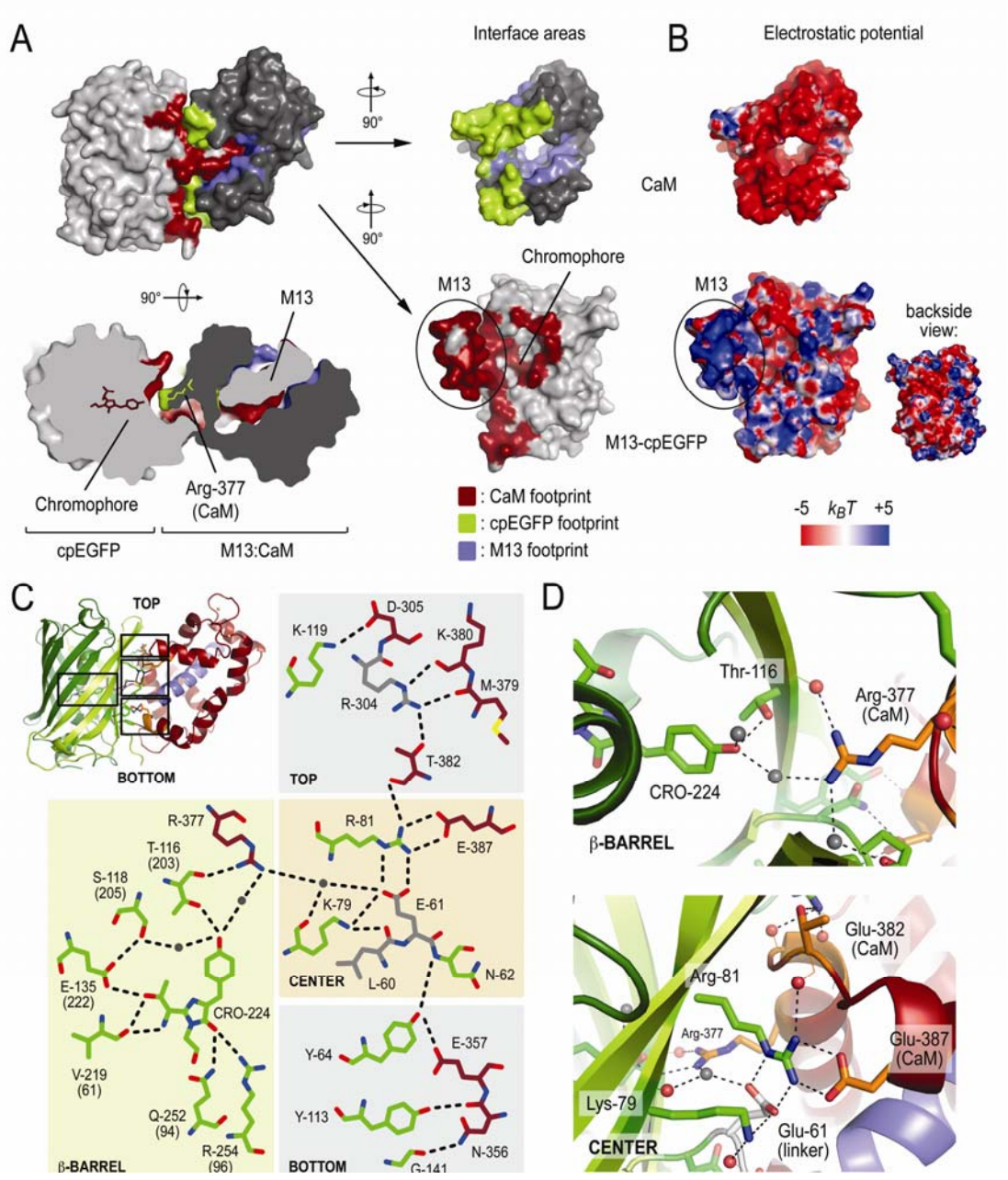


Figure 5.4. Structural comparison of chromophore coordination and pH dependence of GCaMP2 and cpEGFP. (A) Comparison of the conformation of the chromophore and chromophore-coordinating residues. Structure of GFP-S65T (PDB code: 1EMA; yellow carbon atoms), cpEGFP (orange carbon atoms), monomeric GCaMP2•Ca^{2+} (green carbon atoms) and dimeric GCaMP2•Ca^{2+} (blue carbon atoms) were superimposed on the respective cpEGFP (or GFP) moieties. A close-up view of the fluorescent moiety is shown. Residues discussed in the text are shown in stick presentation. Residue numbers follow GCaMP2 numbering. Corresponding residue numbers in GFP are shown in brackets. (B) pH sensitivity of GCaMP2ΔRSET in the absence (orange) and presence (green) of Ca^{2+}. Sensitivity to changes in pH was measured at a protein concentration of 10 μM using buffers containing sodium acetate (pH4.5-5.0), MES (pH5-6.5), HEPES (pH6.5-8), or CHES (pH8.5-10). pKa values were calculated by fitting the experimental data to the Hill equation. (C) pH sensitivity of cpEGFP in the absence (orange) and presence (green) of Ca^{2+}.

The interaction between CaM and cpEGFP is facilitated by a concentration of positively charged residues at the surface of cpEGFP that can form favorable contacts with an overall negatively charged CaM domain (Figure 5.5.B). In addition, there is a considerable amount of shape complementarity in the interfacial region (Figure 5.5.A) that allows a continuous hydrogen bond network to form along the β-barrel of cpEGFP and the point of permutation (Figure 5.5.C).

Figure 5.5. Intramolecular interfaces in monomeric GCaMP2•Ca^{2+}. (A) Interfaces between the cpEGFP, M13 and CaM modules in the structure of monomeric GCaMP2ΔRSET•Ca^{2+}. Residues of the M13-cpEGFP module interacting with CaM are colored red. Interfacial residues on CaM are colored in green and blue for contacts with cpEGFP and the M13 helix, respectively. A top view, rotated 90º around the horizontal axis with respect to the view shown above, is shown as a cutaway rendition of the surface (bottom-left). The fluorophore of cpEGFP and Arg-377 of CaM are shown in stick presentation. Surface presentation of the isolated CaM domain and M13-cpEGFP unit were rotated by +90° and -90°, respectively, with respect to the view of the assembled structure (top-left). (B) Electrostatic potential of the M13-cpEGFP module and CaM mapped onto its molecular surface. Views are identical to (A). Red represents negative and blue represents positive potential (-5 to +5 k_BT). (C) Schematic diagram of the fluorophore environment and the hydrogen bond network between cpEGFP and CaM. The numbering scheme for GCaMP2 was used. Corresponding residue numbers in GFP are shown in brackets. Carbon atoms of residues in cpEGFP, CaM and linker segments are shown in green, dark red, and grey, respectively. Hydrogen bonds shown in the figure are between 2.7 and 3.3Å (not drawn to scale). (D) Close-up views of the interfacial regions in GCaMP2ΔRSET•Ca^{2+}. Water-mediated interaction between the fluorophore and Arg-377 of the CaM domain (top) and cpEGFP:CaM interfacial residues (bottom) are shown.


A
Interface areas
90°
90°
90°
M13
Chromophore
Arg-377
(CaM)
cpEGFP
M13:CaM
Chromophore
M13
M13-cpEGFP
B
Electrostatic potential
CaM
M13
backside
view:
: CaM footprint
: cpEGFP footprint
: M13 footprint
-5 k_BT +5
C
TOP
BOTTOM
D-305
K-119
K-380
M-379
R-304
T-382
TOP
R-377
T-116
(203)
S-118
(205)
R-81
E-387
K-79
E-61
CENTER
L-60
N-62
E-135
(222)
CRO-224
V-219
(61)
Q-252
(94)
R-254
(96)
β-BARREL
Y-64
E-357
Y-113
N-356
BOTTOM
D
Thr-116
Arg-377
(CaM)
CRO-224
β-BARREL
Glu-382
(CaM)
Arg-81
Arg-377
Glu-387
(CaM)
Lys-79
Glu-61
(linker)
CENTER

Instead, Arg-377 forms a water-mediated interaction with the fluorophore, suggesting that Arg-377 is functionally equivalent to His-148 (Figure 5.5.C-D). Significant design efforts yielded an optimized linker sequence connecting the functional moieties of GCaMP2 (Nakai et al., 2001). Most of these regions were resolved in the crystal structure of GCaMP2•Ca^{2+} (except for a 15-residue fragment connecting the two parts of EGFP). Linkers were found to be peripherally involved in the packing of the domains and as such might restrict inter-domain flexibility. From the multitude of interactions at the interface, Arg-81 appears to be a central coordinator (Figure 5.5.C-D). It is hydrogen-bonded to Glu-61 (and Glu-387), one of the two residues linking the M13 fragment to cpEGFP (Figure 5.2.A and 5.3.C). In addition, it is indirectly linked to Arg-304 located in the linker connecting the cpEGFP and CaM moieties (via Thr-382; Figure 5.5.C). The linker length appears to be important as well (Nakai et al., 2001), constraining the conformational freedom of the system even further (see below).

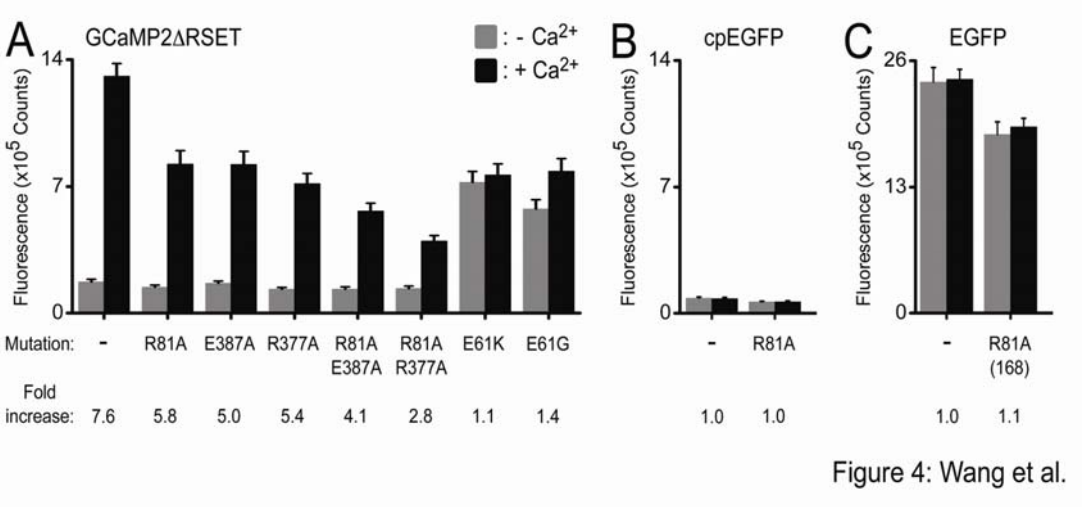


Figure 5.6. Mutational analysis of GCaMP2•Ca^{2+}. (A) Mutations in GCaMP2ΔRSET. (B) Fluorescence emission of corresponding mutations in the isolated cpEGFP module. (C) Fluorescence emission of corresponding mutations in EGFP. The GCaMP2 numbering scheme was used. Arg-81 is equivalent to Arg-168 in GFP and EGFP.

Mutating Arg-377, Arg-81 or its binding partner Glu-387 (to alanine) resulted in a dimished fluorescence signal upon Ca^{2+} binding without affecting the emission intensity in the absence of Ca^{2+} (Figure 5.6.A), while equivalent mutations of Arg-81 in cpEGFP and EGFP affected their fluorescence intensity only slightly (Figure 5.6.B -C). A double-mutant Arg-81-Ala/Glu-387-Ala, removing both side chains of the interacting residues, showed only moderate affects on Ca^{2+}-dependent fluorescence increase compared to the single mutation, consistent with either modification disrupting the bright state conformation. In contrast, the double-mutant Arg-81-Ala/Arg-377-Ala exhibited markedly reduced Ca^{2+} sensitivity indicative of additive contributions of Arg-81 and Arg-377 to the interdomain interface and overall mechanism. Mutating the linker residue Glu-61 to either glycine or lysine generated a defective probe with intermediate fluorescence intensity but lacked Ca^{2+} responsiveness (Figure 5.6.A). Therefore, the negative charge of Glu-61 appears to stabilize the neutral state of the fluorophore under Ca^{2+}-free conditions reducing its background fluorescence. Taken together, the mutagenesis data suggests that contributions from several residues collectively account for the performance of GCaMP2.

Consistent with a shift in fluorophore protonization as the general mechanism of Ca^{2+}-dependent fluorescence, the interactions between cpEGFP and CaM described above contributed to a decrease in the apparent acidity of the fluorophore in Ca^{2+}-bound GCaMP2ΔREST (Nakai et al., 2001). The apparent pK_a value for the high-fluorescent state was 6.9, ~0.8 units lower than under Ca^{2+}-free conditions, consistent with data reported for GCaMP1 (Nakai et al., 2001). In contrast, cpEGFP was less sensitive to pH, with pK_a values of 7.4 in the absence of Ca^{2+} and 7.9 in the presence of Ca^{2+}. These values were approximately 1.5 units higher than the one determined for EGFP (Patterson et al., 1997), further confirming that increased solvent accessibility due to circular permutation is the major determinant for the neutral low-fluorescent state (Baird et al., 1999).

Taken together, Ca^{2+} sensing by GCaMP2 can be described by a switching from a class-1 towards a class-2 GFP by restricting solvent access of the fluorophore and stabilizing it in an ionized form (Tsien, 1998).

Structures of the Ca^{2+}-bound and Ca^{2+}-free state of GCaMP2 in solution

Analysis of GCaMP2 by size exclusion chromatography suggested that Ca^{2+}-bound GCaMP2 adopted a similar but not identical conformation as the Ca^{2+}-free state (Figure 5.7A), with the Ca^{2+}-bound state being slightly more compact based on its retention time. To further explore the conformational changes occurring upon Ca^{2+} binding, we carried out small angle X-ray scattering experiments, yielding molecular geometry parameters such as the radius of gyration (R_g) and maximum diameter of a molecule or assembly (D_{max}) (Table 5.2). The scattering profiles for the Ca^{2+}-free and Ca^{2+}-bound state were very similar at low angles (corresponding to the range of momentum transfer of $0.019 < S < 0.13$ $Å^{-1}$) (Figure 5B). Radii of gyration and maximum diameters of the two states calculated from the scattering profiles were comparable ($R_g[+Ca^{2+}]= 23.1\pm0.5$ *vs* $R_g[-Ca^{2+}]= 23.7\pm0.5$; $D_{max}[+Ca^{2+}]= 70.0\pm2.0$ *vs* $D_{max}[-Ca^{2+}]= 68.0\pm2.0$), indicating an overall similar shape. Distance distribution functions [P(r)] corroborated this finding but also revealed a conformational change between GCaMP2 in the presence or absence of Ca^{2+}. In the Ca^{2+}-free state, the peak of the P(r) function shifted about 5Å towards larger dimension compared to the P(r) function for GCaMP2•Ca^{2+} or the corresponding profile calculated from the coordinates of the crystal structure (Figure 5.7.C). This apparent compaction of GCaMP2 upon Ca^{2+} binding is consistent with the gel filtration experiments (see above).

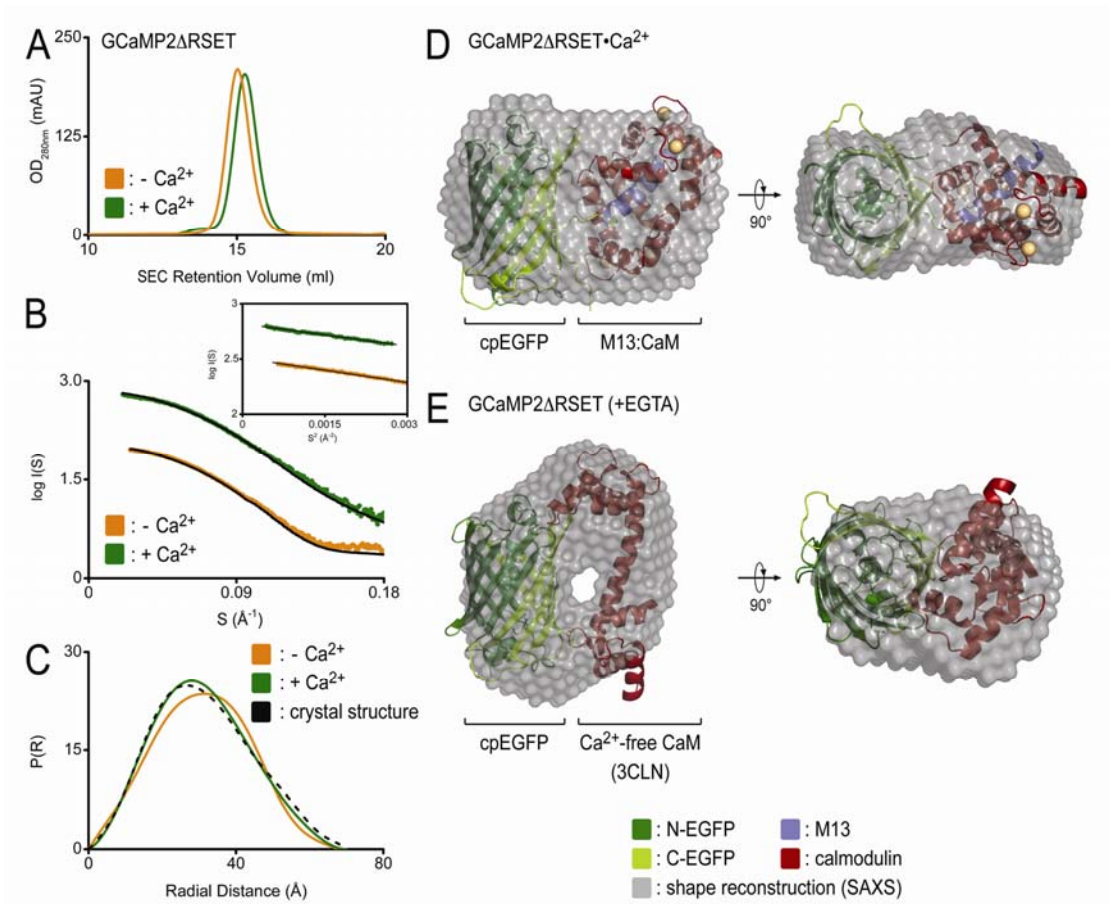


Figure 5.7. SAXS-based shape reconstruction of GCaMP2 in the presence or absence of Ca^{2+}. (A) Size exclusion chromatographic (SEC) analysis of GCaMP2 in the presence and absence of Ca^{2+}. Proteins were analyzed on a S200 gel filtration column (10/300; GE Healthcare) in gel filtration buffer containing EGTA (10 mM; orange trace) or Ca^{2+} (1 mM; green trace). (B) Solution scattering data for GCaMP2ΔRSET in its Ca^{2+}-bound and Ca^{2+}-free state. Small-angle X-ray scattering curves of GCaMP2ΔRSET•Ca^{2+} (green) and its EGTA-treated form (orange) are shown after averaging and solvent-subtraction. Theoretical scattering profiles calculated from the *ab initio* models with the lowest χ values are shown (black line). The inset shows Guinier plots (including linear fits) at the low angle region (S_{max}*R_g<1.3). (C) Distance distribution [P(r)] functions for GCaMP2ΔRSET. P(r) curves of GCaMP2ΔRSET•Ca^{2+} (green) and its EGTA-treated form (orange) were calculated from SAXS data shown in (B) or from the crystal structure of GCaMP2ΔRSET•Ca^{2+} (dashed line). (D) SAXS-based shape reconstruction of GCaMP2ΔRSET•Ca^{2+}. The overall volume from shape reconstructions after averaging (grey envelope) was calculated from 40 independent models. The crystal structure of monomeric GCaMP2ΔRSET•Ca^{2+} was docked into the envelope manually. Two orthogonal views are shown. (E) SAXS-based shape reconstruction of Ca^{2+}-free GCaMP2ΔRSET. The crystal structures of cpEGFP and Ca^{2+}-free CaM (PDB code: 3CLN) (Babu et al., 1988) were docked into the envelope manually.

Table 5.2. SAXS data statistics

	R_g (Å)[a]	D_{max} (Å)[b]	MM_{seq} (kDa)[c]	MM_{exp} (kDa)[d]	V_p (nm^3)[e]	$I_p^{norm}/I_{std}^{norm}$[f]	NSD[g]
10 mM EGTA	23.7±0.5	68.0±2.0	50.7	45.9±5	79±8	0.91±0.10	0.63±0.02
40 M Ca^{2+}	23.1±0.5	70.0±2.0	50.7	46.2±5	66±7	0.92±0.10	0.54±0.01

[a] R_g determined from Grunier equation and data that satisfies $Q_{max}xR_g<1.3$; similar values were obtained using the program GNOM (Svergun, 1992).
[b] D_{max} was determined using the program GNOM (Svergun, 1992).
[c] MM_{seq} is the molecular mass calculated from the primary sequence.
[d] MM_{exp} is the molecular mass calculated from the scattering data:

$$MM_{protein} = MM_{lysozyme} \times \frac{I^o{}_{protein} \times C_{lysozyme}}{I^o{}_{lysozyme} \times C_{protein}}$$

where $MM_{lysozyme}$ is the molecular mass of lysozyme**,** I^0 is the scattering intensity and C (mg/ml) is the protein concentration (Petoukhov and Svergun, 2006).
[e] V_p is the excluded volume (Porod volume) calculated using PRIMUS (Konarev et al., 2003). Data with S>0.25 were excluded from the calculation (Petoukhov and Svergun, 2006; Porod, 1982).
[f] Normalized intensities for proteins (p) $I^p{}_{norm}$ were calculated by

$$I_p{}^{norm} = (\frac{I^o}{C}) / MM_{seq}$$

where I^0 is the intensity calculated using the Guinier equation and C (mg/ml) is the protein concentration determined from its absorbance at 280 nm. Lysozyme served as a mono-dispersed protein standard (std). Ideally, the ratio $I^p{}_{norm}/I^{std}{}_{norm}$ equals 1.00 (Jeffries et al., 2008; Taraban et al., 2008).
[g] After superimposing the independently modeled envelopes, pair-wise normalized spatial discrepancy (NSD) values (Volkov and Svergun, 2003) were calculated as part of the DAMAVER routine. NSD values close to unity indicate good agreement between individual models.

To investigate the origin of these differences, we calculated low-resolution *ab initio* shape reconstructions (see Material and Methods for details) (Figure 5.7.D and E). Envelopes (shown in grey) represent the overall shape of GCaMP2 at a maximum resolution of ~30Å. The shape reconstruction of the Ca^{2+}-bound state agreed very well with the crystal structure obtained for GCaMP2ΔRSET•Ca^{2+} (Figure 5.7.D). In contrast, Ca^{2+}-free GCaMP2 adopted a donut-shaped conformation with a central hole in the center of the envelopes. While one side of the envelope could accommodate the structure of cpEGFP, the other side had dimensions that fit the structure of Ca^{2+}-free apo-CaM very well (Babu et al., 1988) (Figure 5.7.E).

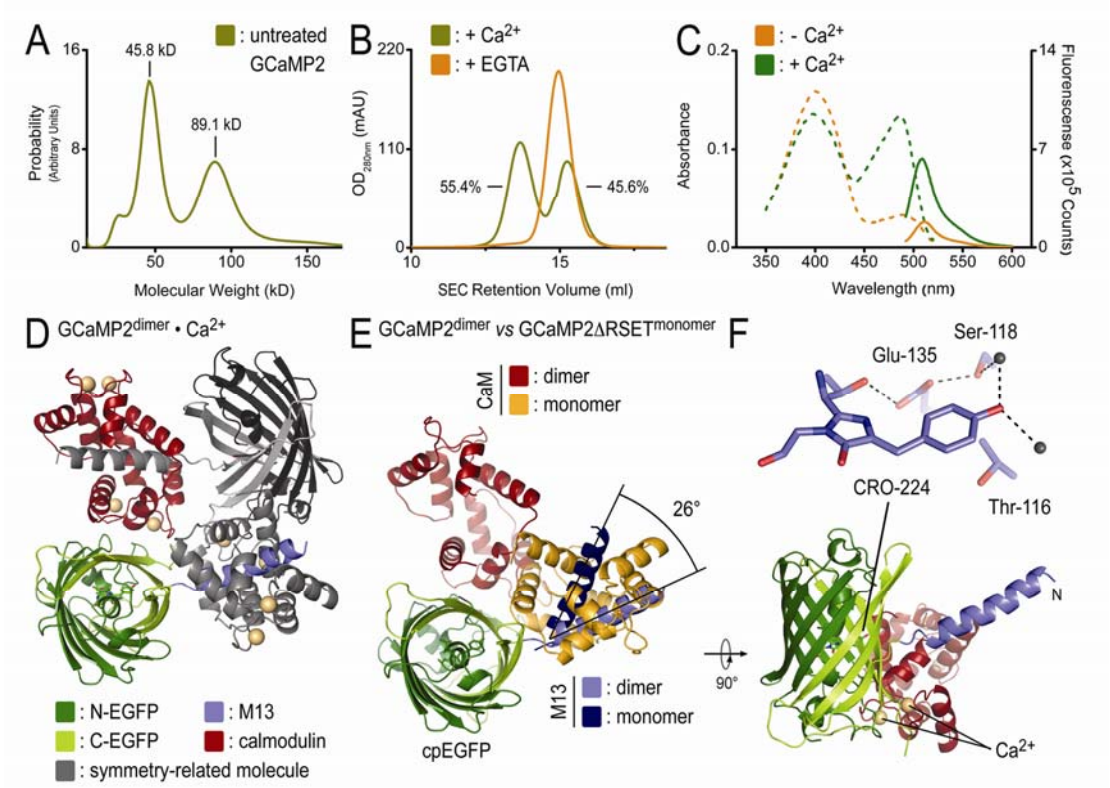


Figure 5.8. Characterization and crystal structure of dimeric GCaMP2•Ca^{2+}. (A) Identification of a dimeric GCaMP2 by analytical ultracentrifugation. Sedimentation velocity experiments using GCaMP2 purified from *E. coli* yielded bimodal molecular weight distributions with masses (45.8 kD and 89.1 kD), close to the theoretical mass of monomeric (50.7 kD) and dimeric GCaMP2 (101.4 kD), respectively. (B) Monomerization of GCaMP2 by EGTA-treatment. Size exclusion chromatograms for initial preparation of GCaMP2 showed bimodal distribution of the protein. Both peaks were sensitive to EGTA-treatment, converging to a single peak corresponding to Ca^{2+}-free, monomeric protein. EGTA treatment produced protein that remained monomeric in the presence or absence of Ca^{2+} (see Figure 5A). (C) Spectroscopic properties of partially dimeric GCaMP2. Absorbance (dashed lines) and fluorescence emission (solid lines) spectra were measured at 25ºC in Ca^{2+}-free (10 mM EGTA; orange) buffer or in the presence of Ca^{2+} (40 μM; green). (D) Crystal structure of the dimeric assembly. Crystals were grown from partially dimeric Ca^{2+}-bound GCaMP2. A crystallographic dimer of GCaMP2 is shown with protomer colored according to the scheme introduced in Figure 1A, and a crystal symmetry-related molecule shown in grey. Ca^{2+} is shown as yellow spheres. (C) Superposition of monomeric and dimeric GCaMP2. A single protomer of dimeric GCaMP2, colored as in (D), was superimposed on the cpEGFP domain of monomeric GCaMP2ΔRSET•Ca^{2+}, with its M13 helix and CaM module colored in dark blue and orange, respectively. (F) Fluorophore coordination in dimeric GCaMP2•Ca^{2+}. A close-up view of the fluorophore-interacting residues in dimeric GCaMP2•Ca^{2+} is shown. Residue labeling is according to GCaMP2 numbering.

The overall similar dimensions of Ca^{2+}-free and Ca^{2+}-bound GCaMP2 and the shape reconstructions suggest that apo-GCaMP2 exists in a pre-docked state, likely to be stabilized by electrostatic interactions and the short nature of the linker segments connecting the domains in GCaMP2 (see above). Such a conformation might explain the rapid kinetics of Ca^{2+}-induced changes in GCaMP2 that happen at milli-second timescale (Nakai et al., 2001; Tallini et al., 2006).

Identification of a dimeric, low-fluorescent state of GCaMP2•Ca^{2+}

During the purification of the proteins from *E. coli*, we noticed a bimodal distribution of GCaMP2 and GCaMP2ΔRSET in size exclusion chromatography (Figure 5.8.A and B). Analysis by analytical ultracentrifugation revealed that GCaMP2 coexisted in a monomeric and a dimeric state (Figure 5.8.A), with the monomer peak eluting at 15.3 ml and the dimer peak, accounting for about 55% of the sample, eluting at 13.7 ml from a size exclusion column (Figure 5.8.B). The dimer was sensitive to treatment with EGTA (Figure 5.8.B). The EGTA-treated protein eluted with a slightly shorter retention time than the monomer peak of the input protein, indicating that both the dimeric and monomeric species purified from bacterial lysates were bound to Ca^{2+} (Figure 5.8.B). Absorbance and emission profiles of the Ca^{2+}-bound monomer-dimer mixture indicated a half-maximal efficiency with regard to fluorescence intensity compared to the EGTA-treated, monomeric protein (Figure 5.8C and 5.8B). The partial response could be attributed to the monomeric fraction in the sample that accounted for 45% of the sample (Figure 5.8A-C). The spectroscopic analysis suggested that the dimeric state of GCaMP2 is a Ca^{2+}-bound, low-fluorescence state.

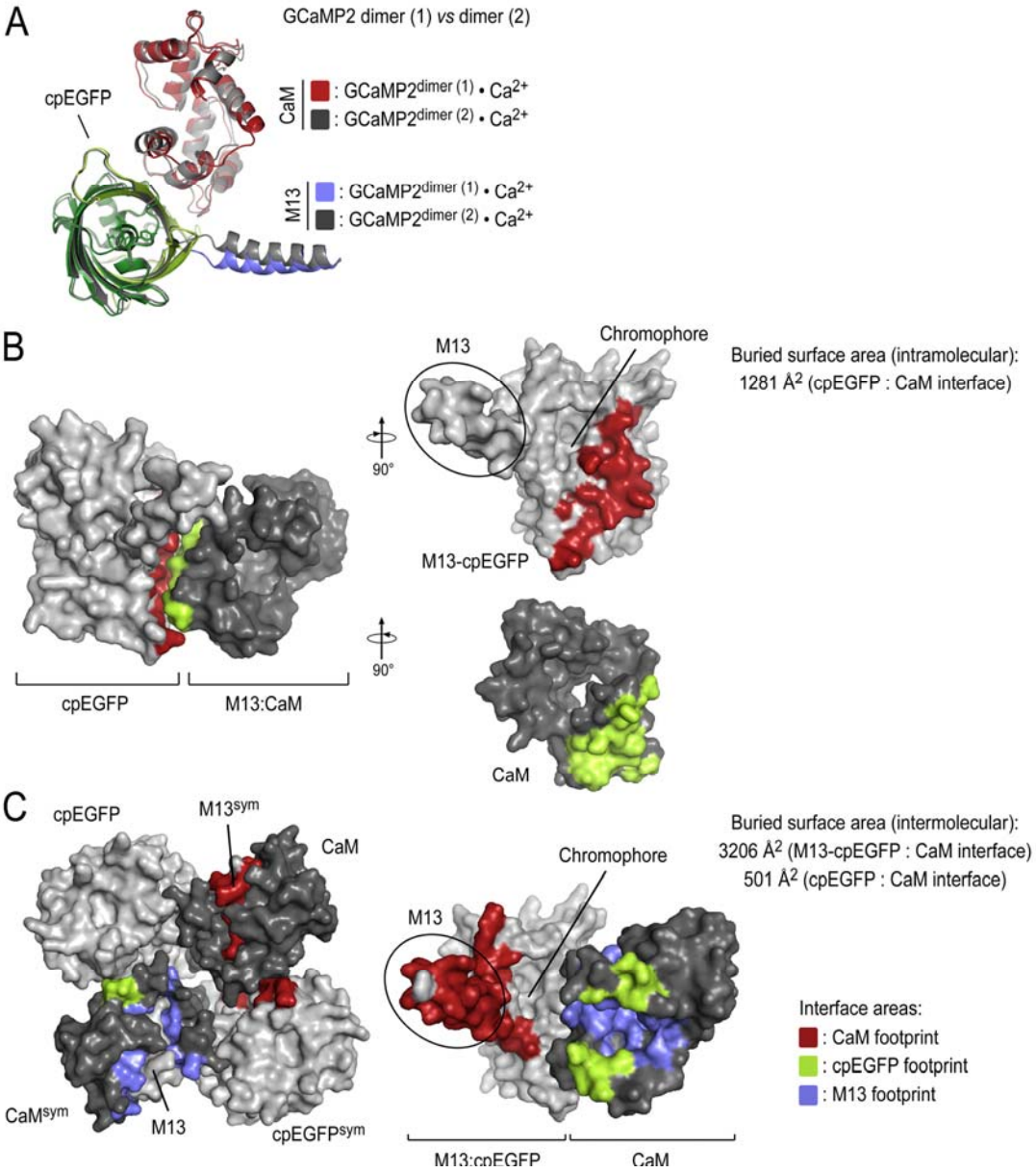
The results were consistent with structural studies on the basis of crystal structures of the Ca^{2+}-bound dimeric state (Figure 5.8.D). Crystals were obtained in two independent trials (Figure 5.9.A), starting from a partially dimeric protein (Figure

5.8.A) or from a sample that was purified in a monomeric state (Figure 5.9.A). In solution, the latter protein remained monomeric upon Ca^{2+} addition suggesting that the high protein concentration and crystallization condition contributed to dimer formation. The structure showed a domain-swapped assembly, in which the M13 helix of one molecule was bound to the CaM moiety of a crystal symmetry-related molecule (Figure 5.8.D). Comparison of the dimeric and monomeric states showed that the M13 helices and CaM domains were separated by a rotation of about 26º and 90º, respectively, relative to cpEGFP. The CaM domain of a symmetry-related molecule in the dimer occupied the approximate position of CaM in the monomeric structure (Figure 5.8.D and E).

Yet, interfacial areas in the dimer were overall less extensive and overlapped only partially with areas that were buried in the monomeric, Ca^{2+}-bound conformation (Figure 5.9.B and C, Figure 5.5.A). Most notably, interactions between cpEGFP and CaM did not occlude the hole in the β-barrel cage introduced by circular permutation, leaving the fluorophore solvent-exposed in such a conformation (Figure 5.9.B and C).

The fluorophore environment in the dimeric structure reflected the spectroscopic properties discussed above. The phenolic oxygen of the fluorophore formed hydrogen bonds with two water molecules. Thr-116 (Thr-203 in GFP-S65T) adopted a rotamer conformation incompetent of chromphore stabilization, characteristic of a low-fluorescent, protonated state also observed in the structure of cpEGFP (Figure 5.8.F). The structural studies corroborated the notion based on the spectroscopic analysis that the dimeric state of GCaMP2 is Ca^{2+}-insensitive.

Figure 5.9. Intra- and intermolecular interfaces in dimeric GCaMP2•Ca^{2+}. (A) Comparison of two independent dimeric GCaMP2•Ca^{2+} structures. The structure of dimer (1) was obtained from protein that purified as a monomer-dimer mixture. Protein used to grow crystals yielding dimer structure (2) was EGTA-treated during purification producing monomeric protein. Dimerization occurred during crystallization. (B) Intramolecular interfaces between cpEGFP, M13 and CaM in the structure of dimer GCaMP2•Ca^{2+}. Buried surface areas were calculated using the program CNS (Brunger et al., 1998). Residues of the M13-cpEGFP module interacting with CaM colored red. Interfacial residues on CaM are colored in green. Surface presentation of the isolated CaM domain and M13-cpEGFP unit were rotated by +90° and -90°, respectively, with respect to the view of the assembled structure (top-left). (C) Intermolecular interfaces between cpEGFP, M13 and CaM in the structure of dimer GCaMP2•Ca^{2+}. Residues of the M13-cpEGFP module interacting with CaM colored red. Interfacial residues on CaM are colored in green and blue for contacts with cpEGFP and the M13 helix, respectively. A dimeric assembly (left panel) and one of its monomers (right panel) are shown.


A
GCaMP2 dimer (1) vs dimer (2)
cpEGFP
CaM
: GCaMP2dimer (1) • Ca2+
: GCaMP2dimer (2) • Ca2+
M13
: GCaMP2dimer (1) • Ca2+
: GCaMP2dimer (2) • Ca2+
B
M13
Chromophore
Buried surface area (intramolecular):
1281 Å2 (cpEGFP : CaM interface)
90°
M13-cpEGFP
90°
cpEGFP
M13:CaM
CaM
C
cpEGFP
M13sym
CaM
Buried surface area (intermolecular):
3206 Å2 (M13-cpEGFP : CaM interface)
501 Å2 (cpEGFP : CaM interface)
Chromophore
M13
CaMsym
M13
cpEGFPsym
M13:cpEGFP
CaM
Interface areas:
: CaM footprint
: cpEGFP footprint
: M13 footprint

CONCLUSIONS

Here, we described the molecular mechanism underlying Ca^{2+}-dependent fluorescence of the high signal, genetically encoded Ca^{2+} sensor GCaMP2 based on high-resolution crystal structures and solution scattering studies. Many aspects of the particular design of GCaMP2 are likely to contribute to the success of this Ca^{2+} indicator, including the particular point of permutation allowing for reversible solvent exposure of the fluorophore, the complementarity between interfaces, and the functional conservation of fluorophore-coordinating residues. Our results indicate that circular permutation of EGFP destabilizes the deprotonated state of the fluorophore, and that this state is re-established in a Ca^{2+}-dependent manner through intensive contacts between the condensed CaM:M13 structure and EGFP moieties. A pre-docked state of CaM in the absence of Ca^{2+} might contribute to the rapid transition between the two fluorophore states.

The complex and extensive nature of the interactions underlying this mechanism carries important implications for the design of GECIs and other molecular sensors based on the cpEGFP platform. First, achieving the reversible destabilization of a bright state with other fluorescent proteins or fusion peptides by random means will likely require extensive mutagenesis and screening, as occurred with the development of GCaMP2. Rather, forward design of structures to achieve this goal would seem more likely to yield effective signaling molecules. Second, the surface properties such as charge and shape complementarity of the fluorescent module and the fusion peptides should be considered so as to optimize the kinetics of state transition and to minimize interactions with other binding partners. Alternatively, the incorporation of binding sites (e.g. a calmodulin binding site) could also confer a pre-docked state. The results may suggest that it is unlikely, however, that marked improvement in kinetics will be achieved using current cpEGFP-based signaling

proteins. While this may limit the utility of these GECIs for the most rapid cell signaling events (Hendel et al., 2008; Mao et al., 2008; Tallini et al., 2006), GCaMP2 transitions occur in less than 100 ms at physiological temperatures, sufficient for many signaling processes. Third, the discovery of a dimeric dim state indicates the importance of preventing intermolecular binding, as these proteins contribute to background fluorescence and thereby may markedly limit the dynamic range of the signaling system. In summary, our studies provide a structural basis for future improvements in the rational design of protein-based signaling molecules and identify key limitations to this strategy.

REFERENCES

Adams P., Grosse-Kunstleve R., Hung L., Ioerger T., McCoy A., Moriarty N., Read R., Sacchettini J., Sauter N., Terwilliger, T. (2002) PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. D Biol. Crystallogr. *58*, 1948-1954.

Babu Y.S., Bugg C.E., Cook W.J. (1988) Structure of calmodulin refined at 2.2 A resolution. J. Mol. Biol. *204*, 191-204.

Baird G.S., Zacharias D.A., Tsien, R.Y. (1999) Circular permutation and receptor insertion within green fluorescent proteins. Proc. Natl. Acad. Sci. U S A *96*, 11241-11246.

Baker N.A., Sept D., Joseph S., Holst M.J., McCammon J.A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci U S A *98*, 10037-10041.

Brejc K., Sixma T.K., Kitts P.A., Kain S.R., Tsien R.Y., Ormo M., Remington S.J. (1997) Structural basis for dual excitation and photoisomerization of the Aequorea victoria green fluorescent protein. Proc Natl Acad Sci U S A *94*, 2306-2311.

Brunger A.T., Adams P.D., Clore G.M., DeLano W.L., Gros P., Grosse-Kunstleve R.W., Jiang J.S., Kuszewski J., Nilges M., Pannu N.S., *et al.* (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr *54 (Pt 5)*, 905-921.

Clapperton J.A., Martin S.R., Smerdon S.J., Gamblin S.J., Bayley P.M. (2002) Structure of the complex of calmodulin with the target sequence of calmodulin-dependent protein kinase I: studies of the kinase activation mechanism. Biochemistry *41*, 14669-14679.

Emsley P., Cowtan K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr *60*, 2126-2132.

Hendel T., Mank M., Schnell B., Griesbeck O., Borst A., Reiff D.F. (2008) Fluorescence changes of genetic calcium indicators and OGB-1 correlated with neural activity and calcium in vivo and in vitro. J Neurosci *28*, 7399-7411.

Jeffries C.M., Whitten A.E., Harris S.P., Trewhella J. (2008). Small-angle X-ray scattering reveals the N-terminal domain organization of cardiac myosin binding protein C. J Mol Biol *377*, 1186-1199.

Konarev P.V., Volkov V.V., Sokolova A.V., Koch M.H.J., Svergun D.I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. Journal of Applied Crystallography *36*, 1277-1282.

Kotlikoff M.I. (2007). Genetically encoded Ca2+ indicators: using genetics and molecular design to understand complex physiology. J Physiol *578*, 55-67.

Ledoux J., Taylor M.S., Bonev A.D., Hannah R.M., Solodushko V., Shui B., Tallini Y., Kotlikoff M.I., Nelson, M.T. (2008). Functional architecture of inositol 1,4,5-trisphosphate signaling in restricted spaces of myoendothelial projections. Proc Natl Acad Sci U S A *105*, 9627-9632.

Mank M., Griesbeck, O. (2008). Genetically encoded calcium indicators. Chem Rev *108*, 1550-1564.

Mao T., O'Connor D.H., Scheuss V., Nakai J., Svoboda K. (2008) Characterization and subcellular targeting of GCaMP-type genetically-encoded calcium indicators. PLoS ONE *3*, e1796.

Miyawaki A., Griesbeck O., Heim R., Tsien R.Y. (1999) Dynamic and quantitative Ca2+ measurements using improved cameleons. Proc Natl Acad Sci U S A *96*, 2135-2140.

Miyawaki A., Llopis J., Heim R., McCaffery J.M., Adams J.A., Ikura M., Tsien R.Y. (1997) Fluorescent indicators for Ca2+ based on green fluorescent proteins and calmodulin. Nature *388*, 882-887.

Nagai T., Sawano A., Park E.S., Miyawaki, A. (2001) Circularly permuted green fluorescent proteins engineered to sense Ca2+. Proc Natl Acad Sci U S A *98*, 3197-3202.

Nagai T., Yamada S., Tominaga T., Ichikawa M., Miyawaki A. (2004) Expanded dynamic range of fluorescent indicators for Ca(2+) by circularly permuted yellow fluorescent proteins. Proc Natl Acad Sci U S A *101*, 10554-10559.

Nakai J., Ohkura M., Imoto, K. (2001) A high signal-to-noise Ca(2+) probe composed of a single green fluorescent protein. Nat Biotechnol *19*, 137-141.

Nausch L.W., Ledoux J., Bonev A.D., Nelson M.T., Dostmann W.R. (2008) Differential patterning of cGMP in vascular smooth muscle cells revealed by single GFP-linked biosensors. Proc Natl Acad Sci U S A *105*, 365-370.

Ormo M., Cubitt A.B., Kallio K., Gross L.A., Tsien R.Y., Remington, S.J. (1996) Crystal structure of the Aequorea victoria green fluorescent protein. Science *273*, 1392-1395.

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

Palmer A.E., Tsien, R.Y. (2006) Measuring calcium signaling using genetically targetable fluorescent indicators. Nat Protoc *1*, 1057-1065.

Patterson G.H., Knobel S.M., Sharif W.D., Kain S.R., Piston D.W. (1997) Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. Biophys J *73*, 2782-2790.

Petoukhov M.V., Svergun D.I. (2006). Joint use of small-angle X-ray and neutron scattering to study biological macromolecules in solution. Eur Biophys J *35*, 567-576.

Pologruto T.A., Yasuda R., Svoboda, K. (2004) Monitoring neural activity and [Ca2+] with genetically encoded Ca2+ indicators. J Neurosci *24*, 9572-9579.

Porod G. (1982). General Theory. In Small-Angle X-Ray Scattering, O. Kratky, ed. (Academic Press, London), pp. 17-51.

Reiff D.F., Ihring A., Guerrero G., Isacoff E.Y., Joesch M., Nakai J., Borst A. (2005) In vivo performance of genetically encoded indicators of neural activity in flies. J Neurosci *25*, 4766-4778.

Rodriguez Guilbe M.M., Alfaro Malave E.C., Akerboom J., Marvin J.S., Looger L.L., Schreiter, E.R. (2008) Crystallization and preliminary X-ray characterization of the genetically encoded fluorescent calcium indicator protein GCaMP2. Acta Crystallogr Sect F Struct Biol Cryst Commun *64*, 629-631.

Roell W., Lewalter T., Sasse P., Tallini Y.N., Choi B.R., Breitbach M., Doran R., Becher U.M., Hwang S.M., Bostani T.*, et al.* (2007). Engraftment of connexin 43-expressing cells prevents post-infarct arrhythmia. Nature *450*, 819-824.

Souslova E.A., Belousov V.V., Lock J.G., Stromblad S., Kasparov S., Bolshakov A.P., Pinelis V.G., Labas Y.A., Lukyanov S., Mayr L.M., Chudakov D.M. (2007) Single fluorescent protein-based Ca^{2+} sensors with increased dynamic range. BMC Biotechnol. *7*, 37.

Svergun D., Barberato C., Koch M.H.J. (1995) CRYSOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. Journal of Applied Crystallography *28*, 768-773.

Svergun D.I. (1992) Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria. Journal of Applied Crystallography *25*, 495-503.

Svergun D.I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys J *76*, 2879-2886.

Tallini Y.N., Brekke J.F., Shui B., Doran R., Hwang S.M., Nakai J., Salama G., Segal S.S., Kotlikoff M.I. (2007) Propagated endothelial Ca2+ waves and arteriolar dilation in vivo: measurements in Cx40BAC GCaMP2 transgenic mice. Circ Res *101*, 1300-1309.

Tallini Y.N., Ohkura M., Choi B.R., Ji G., Imoto K., Doran R., Lee J., Plan P., Wilson J., Xin H.B., *et al.* (2006) Imaging cellular signals in the heart in vivo: Cardiac expression of the high-signal Ca2+ indicator GCaMP2. Proc Natl Acad Sci U S A *103*, 4753-4758.

Taraban M., Zhan H., Whitten A.E., Langley D.B., Matthews K.S., Swint-Kruse L., Trewhella J. (2008) Ligand-induced conformational changes and conformational dynamics in the solution structure of the lactose repressor protein. J Mol Biol *376*, 466-481.

Tsien R.Y. (1998) The green fluorescent protein. Annu Rev Biochem 67, 509-544.

Volkov V.V., Svergun D.I. (2003) Uniqueness of ab initio shape determination in small-angle scattering. Journal of Applied Crystallography *36*, 860-864.

Yang F., Moss L.G., Phillips G.N., Jr. (1996) The molecular structure of green fluorescent protein. Nat. Biotechnol. *14*, 1246-1251.

CHAPTER 6

MOLECULAR MECHANSIM OF COLOR SWITCHABLE FLUORESENT PROTEIN GMKATE

ABSTRACT

Fluorescent proteins that can switch between distinct colors have made significant contributions to modern biomedical imaging technologies and fundamental molecular cell biology. Here we report the identification and biochemistry analysis of a green-color mKate variant GmKate. GmKate can undergo reversible green-to-red color conversion. At physiological pH, GmKate absorbs blue light (445 nm) and emits green fluorescence (525 nm), resembling GFP; at pH above 9.0, GmKate absorbs 600 nm light and emits 648 nm far-red fluoresence, similar to its parental molecule mKate. Based on the spectrum analysis and crystal structures of GmKate at green and red states, the reversible color transition is attributed to the different protonation states of *cis*-chromophore. This is different from the prevailing color transition mechanisms that are due to either the covalent modification in or near the chromophore region, or the chromophore isomorhporization. The *cis*-chromophore protonation states are mediated by a hydrogen network observed in the crystal structures. This study established GmKate as a prototype of genetically encoded pH sensor for biological studies, and provides novel insights aiding the development of other green-to-red color convertible fluorescent sensors.

INTRODUCTION

The color-switchable fluorescent proteins have facilitated major breakthroughs in modern molecular cell biology. They are widely used in studying biological events, such as cellular trafficking and endocytosis, and contributed to the development of novel innovations such as super-resolution microscopy (PALM, FPALM and STROM; Patterson and Lippincott-Schwartz, 2002; Nienhaus et al., 2006; Subach et al., 2010; Kruhlak et al., 2009; Andresen et al., 2008; Adam et al., 2010). The development of novel color switchable sensors, especially ones that can respond to physiological triggers such as H^+, calcium and cycli-di-GMPs will be of great importance for future biomedical studies.

The most well-understood color switching mechanisms are based on studies of several photo-activatable proteins. The irreversible photoconversion process usually involves chemical modifications in or near the chromophore upon light radiation and can be classified into two types. The type I mechanism involves a β-elimination reaction as observed for Kaede (Ryoko et al., 2002), EosFP (Wiedenmann et al., 2004), IrisFP (green to red transition. Adam et al., 2008), Dendra (Gurskaya et al., 2006) and KikGR (Tsutsui et al., 2009). The crystallographic studies indicate that the bond between the amide nitrogen and His^{62}-C_α is cleaved, and a double bond between His^{62}-C_α and -C_β is formed to further extend the π-conjugation system, resulting in a red fluorescent chromophore. The type II mechanism involves a decarboxylation reaction of a glutamic acid near chromophore. This mechanism has been observed in PA-GFP (Hendreson et al., 2009; Patterson and Lippincott-Schwartz, 2002) and PAmCherry (Subach et al., 2009). In PA-GFP, the decarboxylation is coupled with a chromphore isomerization process, which renders the protein from the dim state to a green fluoresent state (Hendreson et al., 2009). In PAmCherry, the decarboxylation is

coupled with the oxidation of Tyr^{63}-C_α and Tyr^{63}-C_β single bond that extends the π electron conjugation and turns on the chromphore.

The well-characterized reversible photoactivable proteins utilize chromophore isomerization mechanism, which sometimes is coupled with chromophore protonation or deprotonation, as observed for IrisFP (dark to green or dark to red transition; Adam et al., 2008), asFP595 (Tretyakova et al., 2007; Schafer et al., 2008) and mTFP0.7 (Henderson et al., 2007). In the IrisFP (dark to green transition) case, the light excitation promotes the isomerization of the chromophore. The original neutral chromophore turns into an anionic high-fluorescence green state. While the light provides the energy for the isomerization, the residues surrounding the chromophore also contribute to stabilization of an anionic chromophore, presumably through a hydrogen bond network (Adam et al., 2008).

mKate was originally engineered from the red fluorescent protein eqFP578, a protein cloned from the sea anemone *Entacmaea quadricolor* (Shcherbo et al., 2007). Wild-type mKate is a bright monomeric fluorescent protein with excitation and emission peaks at 585nm and 635nm, respectively. Structural characterization revealed that mKate chromophore can adopt two different conformations (Pletnev et al., 2008). The *trans*-conformation is prodominated under low pH (2-4) conditions and is not fluorescent. The *cis*-conformation dominates at higher pH (7-9) and is highly fluorescent. At physiological pH 7.4, more than 80% of the chromophore is in the *cis*-conformation. Excitation of mKate with green light results in a 10-20% increase in brightness, which can be explained by photo-induced *trans-cis* isomerization (Pletnev et al., 2008). Based on the crystal structures and random mutagenesis, an improved version, mKate2, has been reported recently (Shcherbo et al., 2009). mKate2 contains three mutations. Mutation $S^{158}A$ destabilizes the *trans*-conformation and substantially enhances mKate's brightness. Mutation $V^{48}A$ and $K^{238}A$ were found to accelerate

protein's maturation rate. Another independent optimization procedure based on wild-type mKate yielded mNeptune with a peak excitation at 600nm (Lin et al., 2009). mNeptune is an auto-fluorescent protein with the most far-red-shifted excitation and proven great performance in deep tissue imaging in living mice (Lin et al.,2009). The bathematic shift is attributed to a collective effect from mutations $S^{158}C$, $M^{41}G$, $S^{61}C$ and $Y^{197}F$. $M^{41}G$ provides extra room for a water molecule that forms a hydrogen bond with the terminal oxygen of chromophore. $S^{158}C$ enhances mNeptune brightness following a similar mechanism as S158A in mKate2 (Pletnev et al., 2008; Shcherbo et al., 2009), but with much simplified photobleaching kinetics (Lin et al., 2009).

Interestingly, mKate and its variant ($S^{158}A$ mutant, $S^{158}C$ muant and mNeptune) all contain some residual green fluorescent component when excited at 470nm (Lin et al., 2009). Recent studies on several orange and red fluorescent proteins suggested that mKate and Katushka can undergo red-to-green photoconversion upon 405 nm and 561 nm excitation (Kremers et al., 2009). These observations suggest that the green emission is an intrinsic property of mKate-like proteins. However, the nature of the green component is largely unknown. In this article, we report the identification of a green-colored mKate variant GmKate. In solutions of physiological pH, GmKate predominately absorbs at 445 nm and emits green light peak at 525 nm. Interestingly, this optical property alters with pH. We found that the absorbance and emission properties of GmKate are largely depending on the pH of the solution. GmKate is barely fluorescent when subject to very low pH (2.0-4.0), with a faint yellow color at a high protein concentration. At intermediate pH (5.0-9.0), GmKate emits efficiently green fluorescence with absorbance and emission peaks at 450 nm and 525 nm, respectively. At very high pH (9.0-10.6), GmKate appears blue under ambient light, with absorbance and emission maxima shifted to 600 nm and 648 nm, respectively. This color conversion process is fast and reversible, without a requirement of external

energy. Our biochemical and structural analysis on the GmKate reveals a novel mechanism for this reversible green-to-color conversion process, independent of chromophore isomerization and maturation.

MATERIALS AND METHODS

Protein Expression and Purification

The coding regions corresponding to wide type mKate were amplified by standard PCR and were cloned into the pRSET expression plasmid. GmKate and mKate^{S158A} plasmid were generated using the QuikChange XL Mutagenesis Kit (Stratagene) following the manufacturer's instructions. GmKate used for structural study was cloned into a modified pET28a expression plasmid (Novagen) yielding N-terminally hexahistidine SUMO fusion proteins. The hexahistidine-tagged SUMO-moiety was cleavable using the protease from Ulp-1 from *S. cerevisiae*.

Transformed *E.coli* cells BL21 (DE3) (Novagen) were grown in TB medium supplemented with 50 mg/l antibiotics at 37°C. At a cell density corresponding to an absorbance of 1.0 at 600 nm the temperature was reduced to 18°C, and protein production was induced with 1 mM IPTG. Proteins were expressed for 12–16 hr. Cells were collected by centrifugation, resuspended in NiNTA buffer A (25 mM Tris-Cl pH 8.2, 500 mM NaCl and 20 mM imidazole). After cell lysis by sonication, cell debris was removed by centrifugation at 40,000 x g for 1 hr at 4°C. Clear lysates were loaded onto HisTrap NiNTA columns (GE Healthcare) equilibrated in NiNTA buffer A. The resin was washed with 20 column volumes NiNTA buffer A, and proteins were eluted in a single step with NiNTA buffer A supplemented with 500 mM imidazole. GmKate was incubated with SUMO protease ULP-1 at 4°C overnight for removal of the hexahistidine-SUMO tag, and the cleaved protein was collected in the flow-through during NiNTA affinity chromatography. Proteins were further subjected to size

exclusion chromatography on a Superdex200 column (16/60; GE Healthcare) equilibrated in gel filtration buffer (150 mM NaCl, 25 mM HEPES pH 7.4).

Crystallization, X-ray Data Collection, and Structure Solution

Crystals were obtained by hanging drop vapor diffusion by mixing equal volumes of protein (~20-50 mg/ml) and reservoir solution followed by incubation at 20°C. The green colored Gmkate crystal is obtained with the reservoir solution containing 0.1 M Tris pH7.0, 22% PEG3350, 0.1M $MgCl_2$. Blue colored Gmkate crystals grew in 20% PEG3350, 0.1 M glycine pH10.0, and 0.2 M $MgCl_2$. The reservoir solution for crystallization of red-colored mKateS^{158A} consisted of 22% PEG3350, 0.1M Tris-HCl pH 7.5. All crystals were cryo-protected using crystallization solutions supplemented with 20% xylitol, frozen in liquid nitrogen, and kept at 100 K during data collection.

Data sets were collected using synchrotron radiation at the Cornell High Energy Synchrotron Source (CHESS, Ithaca, beamline A1, wavelength 0.977Å). Data reduction was carried out with the software package HKL2000 (Otwinowski et al., 1997). Phases were obtained from molecular replacement using the software package PHENIX (Adams et al., 2002) with 1.8Å structures of mKate at pH 2.0 (PDB code: 3BX9) (Ormo et al., 1995) as search models. Manual refinement in COOT (Emsley et al., 2004) and minimization using PHENIX (Adams et al., 2002) yielded the final models with good geometry (with all residues being in allowed regions of the respective Ramachandran plots). Illustrations were made in Pymol (DeLano Scientific).

Absorbance and fluorescence spectroscopy

Absorbance spectra (260-800 nm; 10 mm path length) of purified proteins (20 μM) were recorded in duplicates on a DU730 UV/Vis spectrophotomer (Beckman Coulter) at 25ºC. Emission spectra were recorded in triplicates on a fluorescence spectrophotomer (Photon Technology International) at a protein concentration of 100 nM and an excitation wavelength of 445 nm and 600 nm. Spectrums to changes in pH was measured using buffers containing sodium acetate (pH4.5-5.0), MES (pH5-6.5), HEPES (pH6.5-8), or CHES (pH8.5-10).

Quantum yield calculation.

The quantum yields were calculated using Rhdopsin and EGFP as optical standards. The absorbance maximum and the corresponding integrated emission intensity were measured for five different protein concentrations as described before. Although GmKate has dual excitation/emissions, only major absorbance at each pH (445 for green state, 600 for red state) is considered. The linear fitting of the spots result in the slope value that is proportional to the quantum yield.

Acrylimine hydrolysis assay

Proteins (2mg/ml) were incubated in 15uL gel filtration buffer, acidic buffer (0.1M HCl) and alkaline buffer (0.1M NaOH) at room temperature for 10 mins. Samples were then boiled in 95 degree heat block for 10mins before loading on a 12% SDS-PAGE. The gels were stained with 1% Coomassie-blue R250 solution for 10mins and de-stained in 10% Acetate acid solution overnight.

RESULTS AND DISCUSSIONS

GmKate emits green light at physiological pH

At pH 7.0, GmKate absorbance spectrum contains two peaks. The major peak locates at 445 nm, which excites the 525nm emission maxima. The 600nm minor peak intensity is about 10% of the major peak, and barely emits any fluorescent upon excitation (Figure 6.1.A-B). The quantum yield of 445nm/525nm is 0.03, comparable to the one of mPlum (Wang et al., 2004; Shanar et al., 2005). The green color emission is readily visible in ambient light even at a very low protein concentration. In contrast, the mKate^{S158A} that lacks S^{143}C mutation has predominately one absorbance peak at 588nm and emits strongly at 625 nm (Figure 6.1.A-B). Residue Serine^{143} is proposed to stabilize the chromophore in the *cis*-conformation (Pletnev et al., 2008). The replace of Serine with Cystine is predicted to inhibit its function as a hydrogen bond donor and/or acceptor, thus destabilizes the *cis*-conformer. Indeed, mKate^{S143C} is of no fluorescence at physiological pH, similar to the behavior of wide type mKate at lower pH condition (Pletnev et al., 2008). It is less clear which chromophore isomer state is favored in GmKate, since both S158 and S143 are mutated to hydrophobic residues that cannot stabilize chromophore. The other chromophore surrounding residues most likely play important roles in determining its isomer state and the optical properties.

GmKate contains the acrylimine bond as in wide type mKate

To understand the nature of green fluorescence, we first studied the biochemistry of Gmkate chromophore. The Dsred variant K83R also emits green light at physiological pH, due to the fact that the acrylimine bond is not formed in this mutant and it was trapped in a GFP-chromophore like conformation (Gross et al., 2000). We therefore test if this explains the green component in Gmkate. Here we show that the green fluorescence of GmKate is not due to an incomplete chromophore

as the one observed in DsRed^{K83R}. In the acrylimine hydrolysis assay, GmKate showed two fragment bands of apparent masses 15 and 22 kDa, corresponding to the N-terminal 1-63 residue fragment and C-terminal 64-223 fragment. The fragmentation rate gets much faster in the acidic buffer and the bands get smeared due to the unspecific hydrolysis of peptides. All these behaviors are essentially identical to the wide type mKate. In contrast, EGFP remains one major band in all of the experiment conditions. This experiment suggests that the acrylimine bond in the GmKate is already formed, and the chromophore is similar to the one in mKate.

GmKate emits far-red light at high pH

Indeed, GmKate has similar optical properties at high pH as mKate has. At pH above 9.0, GmKate solution turns blue under the ambient light as mNeptune does. The blue color even maintains in the protein crystals grown in the high pH conditions. The 600nm peak increases about 13 folds, and the 445 nm absorbance peak decreases to more than a half of the one at pH 7.0. Excitation at 600nm results in an emission peak around 648nm, about 13nm red-shifted from wide type mKate (Shcherbo et al., 2007) and very close to the 650nm reported for mNeptune (Lin et al., 2009), indicating that the chromophore resembles the one in mNeptune(Figure 6.1E-F). In contrast, mKate^{S158A} 588nm/625nm remains the same level at the pH range 7.0-10.0 (Figure 6.1E-F). This green-to-red transition is reversible and happens at the millisecond time scale.

The pH dependent color transition can be caused by several different mechanisms including isomorphorization, deprotonation or both. The dim-to-red conversion of mKate is due to the chromophore isomorphorization from the low quantum yield, non-planar *trans*-conformer to the high quantum yield, planar *cis*-conformer (Pletnev et al., 2008). The dim-to-green conversion of GFP is instead due

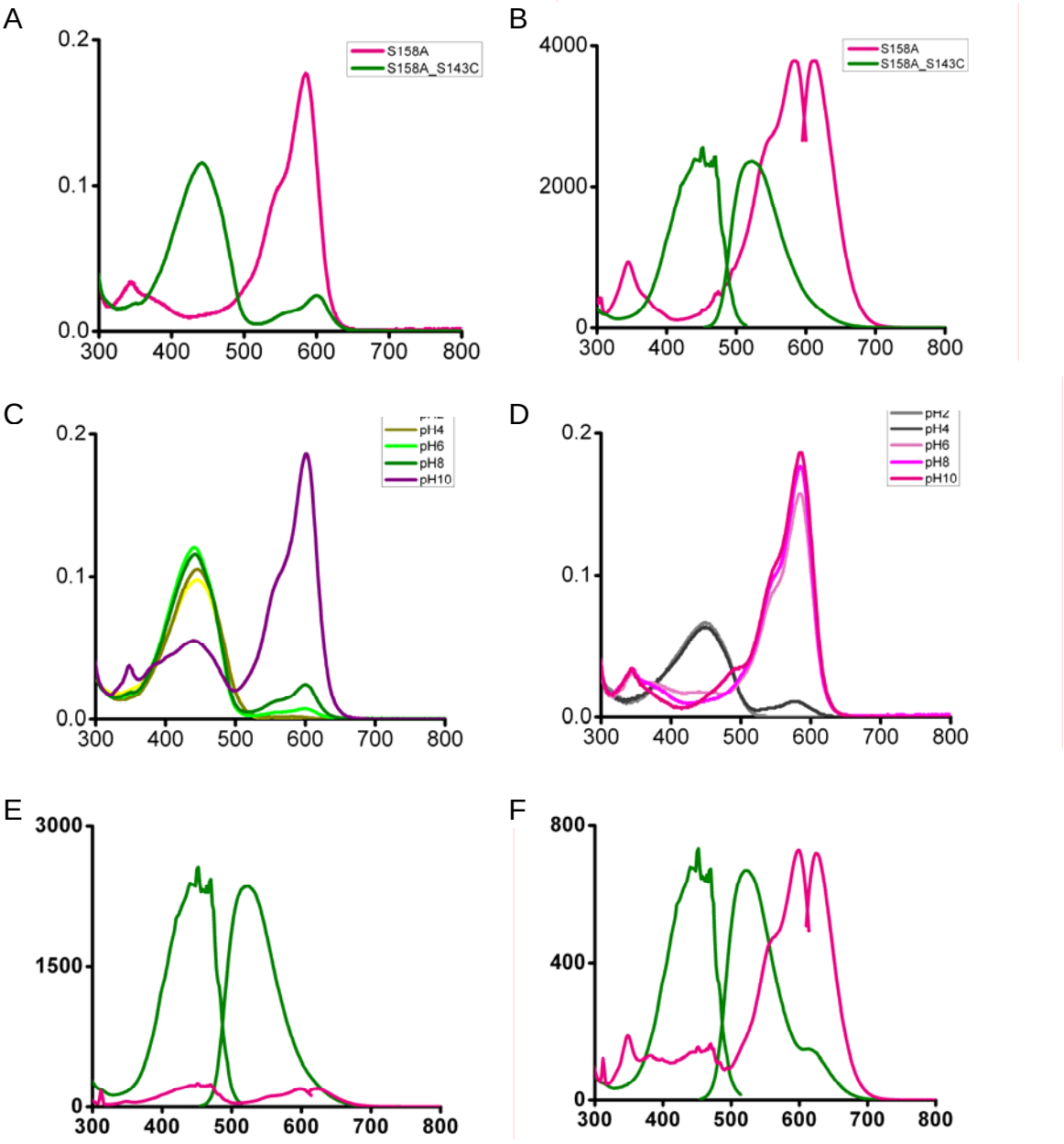


Figure 6.1: Absorbance, emission and excitation spectrums of GmKate and mKate^{S158A} (A) GmKate (green) absorbs primarily at 445nm, with a minor peak at 600nm, mKate^{S158A} (red) has a major absorbance peak at 588nm. The measurements were done with 20uM proteins. (B) Upon 445nm excitation, GmKate (green) emits at 525nm with a quantum yield around 0.03. mKate^{S158A} (red) emits 625nm far-red light upon 588nm excitation. Protein concentration is 10nM for mKate^{158A} and 50nM for GmKate. (C) Absorbance profiles of GmKate (left panel) and mKate^{S158A} (right panel) depend on the pH. The transition from 445nm peak to 600nm peak occurs when the pH is higher than 9.0 for GmKate and 5.0 for mKate^{158A}. (D) Dual excitation and emission spectra of Gmkate at pH 7.0 (left panel) and pH 10.0 (right panel) the 445nm/525nm is presented in green color, and 600nm/648nm is colored red.

to the protonation status change. GFP *cis*-chromophore absorbs at 398nm with low quantum yield in its protonated state and 488 nm with high quantum yield in its deprotonated state (Ormo et al., 1996; Zhang et al., 1996; Brejc et al., 1997; Tsien 1998). For some reversible photoconveritable proteins such as mTFP0.7 (Henderson et al., 2007), IrisFP (Adam et al., 2008) and asFP595 (Andresen et al., 2005), the deprotonation is coupled with the isomorphorization process. To further understand the molecule mechanism of GmKate color conversion, we exam the potential protonation status of GmKate at different pH.

Protonation States of GmKate

At pH 10.0, GmKate chromphore most likely stays in the deprotonated state, as the one proposed for mNeptune (Lin et al., 2009). pH titration suggests that the apparent pKa of green-to-red transition is 9.5(Figure 6.1.C). When pH is below 9.0, the 600nm absorbance is much inhibited, and 445nm/525nm is dominating. This indicates that 445nm/525nm is corresponding to a protonated state. At pH 6.0-9.5, Gmkate emits green fluorescence. Further decreasing pH will only diminish the brightness, but not the absorbance profile. In the pH below 5.5, GmKate no longer emits green fluorescence. This optical behavior resembles the dim state of fully matured IrisFP, which contains a neutral chromophore absorbing at 450nm (Adam et al., 2008). Above observations suggest that the protonated GmKate chromophore can stay in two different states with distinct quantum yields. We propose that the non-fluorescent state at very low pH is the non-planar neutral *trans*-conformation as observed for mKate at similar pH range (Pletnev et al., 2008). The high-fluorescent green state is presumably the planar *cis*-conformer in a protonated state. Interestingly, mKate^{S158A} absorbs at 450nm only when the pH is lower than 5.0, which has no detectable fluorescence (Figure 6.1.D). This suggests that the isomorphorization in mKate^{S158A} may take place before the protonation occurs. We did not observe a

second 450nm absorbance state which emits green light for mKate^{S158A} at pH above 6.0, which is consistent with the previously reported data. Our hypothesis is further supported by the crystal structures of GmKate at green and red states.

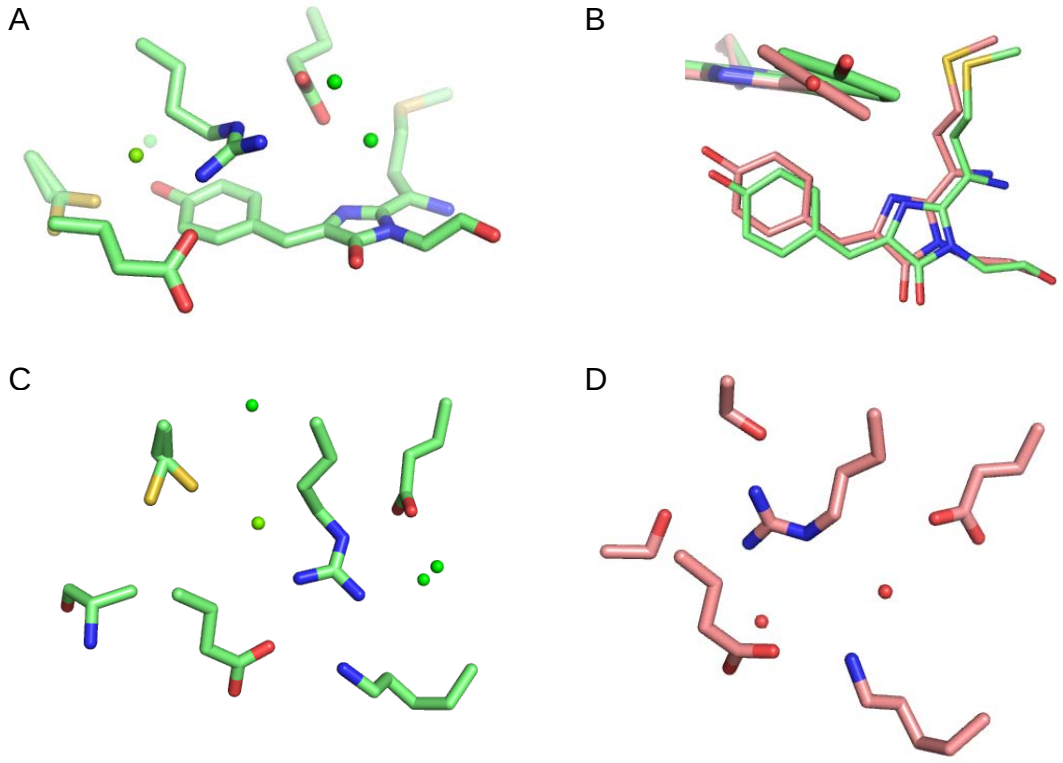


Figure 6.2 : Crystal structures of GmKate in the green state. (A) Gmkate *cis*-chromophore coordination. The small populated trans-conformer is not shown here for the illustration purpose. The crystallization water of low B-factors are colored in green. Cys143 adopts two conformations with equal possibility in the structure. (B) Translational movement of Chromophore in GmKate (green) and mKate (red). The alignment is performed with the entire structure selected and uses LSS (Least Square Superposition) method, Gmkate phenol ring is more coplanar with imdazole ring than the one in the mKate as shown in the inset picture. (C) Hydrogen bond coordination of R197 in GmKate. (D) Hydrogen bond coordination of R197 in mKate.

Overall structure of GmKate

The crystal structures of GmKate in the green and far-red states were refined at 2.0 Å and 1.8 Å resolutions, respectively. No planarity restriction is applied to the refinement. In both structures, the asymmetric unit contains four proteins that

resemble a tetrameric complex. The overall fold of protein is largely identical between GmKate and mKate, with 0.218 Å (green) and 0.221 Å (blue) deviated away from mKate (Pletnev et al., 2008). This is comparable to the 0.213 Å deviation between different asymmetric units observed in the mKate crystal cell.

Chromophore structure of GmKate in green state

Despite the fact that Cys143 can not act as a hydrogen bond donor, green GmKate possess predominately a *cis*-chromophore (Figure 6.2.A). Crystallographic refinement indicates that *cis*-conformation occupies 75±5% of the chromophore region, which is comparable to the value 80% as observed in the mKate structure at pH7.0. This confirms our prediction that the green GmKate chromophore is in a *cis*-conformation. The presence of Cys143 further creates a hydrophobic environment that inhibits the chromphore from deprotonation. In GmKate, chromophore slightly translated 0.7 Å from the position in mKate, as measured by the distance between two phenolic oxygen atoms.(Figure 6.2.B) This is most likely because the original position cannot be retained due to the lose of hydrogen bond to Ser143. Moreover, Gmkate chromphore adopts a more planar conformation, with an 11° dihedral angle between imdazole ring and phenol ring, much smaller than 31° in the mKate (Figure 6.2.B). The surrounding residue R197, E215 and the coordinating water molecules have major conformation rearrangements. In Gmkate, R197 adopts a completed different conformation and form hydrogen bonds with Glu215 (3.0 Å), Glu145 (2.8 Å) and a water molecule (2.8 Å) that is not present in mKate structure (Figure 6.2.C-D). The unique conformation posits the Arginine guanidinium group right on top of the Tyrosine ^{62} -C_α= Tyrosine ^{62}-C_β double bond, presumably altering the conjugated π electron density distribution on the chromophore via a cation-π interaction, as proposed for mNeptune (Lin et al.,2009). Such conformation may contribute to the increased planarity of the chromophore. In mKate, R197 is instead hydrogen bonding

with S143 and Ser158. Two residues locked Arginine in a position that the geometry of cation-π interaction is much disrupted. The resulting cavity is filled by a water molecule (Figure 6.2.D). The GmKate chromophore and its surrounding environment largely resemble the one in the mNeptune structure, despite the fact that two proteins have completely different optical properties at the same pH. This difference is most likely caused by the distinct protonation states of the Chromophore. In mKate and mNeptune, phenolic oxygen forms a hydrogen bond with Ser^{158}, presumably in the deprotonated state (Figure 6.2.D). This anionic *cis*-Chromophore absorbs red light and emits far-red fluorescence. In GmKate, this hydrogen bond can no longer be formed; therefore the chromphore preferably stays protonated. We propose that this neutral *cis*-chromophore is responsible for the green emission of GmKate.

Chromophore structure of GmKate in red state

Red GmKate chromophore completely adopts *cis*-conformation (Figure 6.3.A). Since green GmKate is already dominated by the *cis*-conformation, simple increase of its population cannot explain the dramatic change of the chromophore optical properties. Comparison between red state and green state shows that R197 adopts a slightly different side chain conformation in the red GmKate. In one protomer, χ and ψ angle swiped almost 180°, making the guandiol group more parrell to the chromophore plane. The dihedral angle between guandiol plan and phenol ring is 31°, 15° smaller than the one observed in the green state (Figure 6.3.A-C). We propose that such parallel configuration will further stabilize the chromophore in an anionic state via a π^+ - π stacking mechanism. R197 is proposed to interact with chromophore via cation-π interaction in mNeptune (Lin et al., 2009). Cation- π interaction is usually referring to a non covalent molecular interaction between an electron-rich π system and an adjacent monopole cation within the distance of 6 Å. In addition to

being a cation, Arg guandiol group contains a delocalized π_4^6 electron conjugation system, with the central Carbon atom in the planar sp^2 hybrid state. In the red GmKate, this electron-deficient conjugation system lies on the top of the four sp^2 carbon atoms with similar triangular geometry in the chromophore, forming a π^+ - π interaction. In this stacking configuration, R197 servers as an electron acceptor and thus can stabilize the chromophore in the anionic state. π^+ - π interaction is much stronger than the π - π interaction due to the removal of electron repulsion, and usually happens when the donor and acceptor comes closer than 4 Å. The parrell face-to-face stacking is favored as a low energy conformation for π^+ - π system. Based on the above analysis, we propose that the red GmKate contains an anionic *cis*-chromophore, which is stabilized by R197. This role of R197 is consistent with the one that has been reported for mNeptnue (Lin et al.,2009). The side chain conformation largely resembles the ones in mNeptune.(Figure 6.3B) In the crystal structure, R197 forms hydrogen bonds with E145 (2.6 Å), E215 (2.9 Å) and K67 (2.8 Å), in addition to two water molecules (2.7 Å and 3.1 Å). E215 hydrogen bonds with one water molecule (3.0 Å) that interacts with Chromophore imdazole ring (3.2 Å). E145 and K67 also interact with Try178 (2.8 Å and 3.2 Å) (Figure 6.3D). The overall network maintains in the green GmKate, although the individual distance varies slightly. This surrounding hydrogen bond network may serve as a buffer system that tightly regulates the level of protonation and electron delocalization on the R197, thus influence its ability to stabilize the chromophore at anionic state. It is less clear how the external pH may alter the hydrogen network's behavior, since the hydrogen atoms are invisible in the crystal structure.

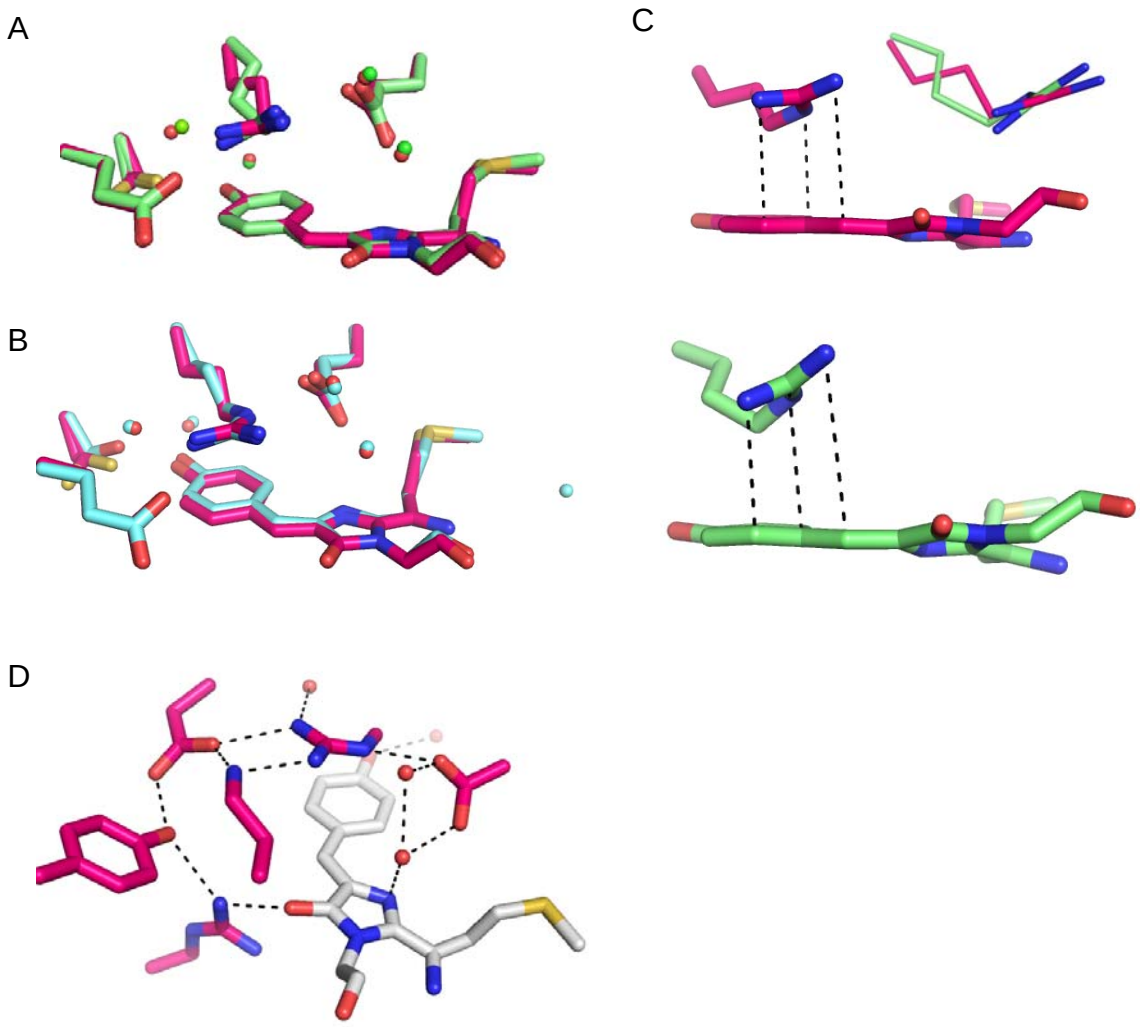


Figure 6.3: Crystal structure of GmKate in the red state. (A) Structural comparison of Gmkate pH 10.0 (red) and pH 7.0 (green). The chromophore regions are quite similar in two structures, except for the arginine 197. (B) Structural comparison of red GmKate with mNeptune (blue). The chromophore and surrounding residues are essentially identical. The water that interacts with chromophore terminal oxygen in mNeptune is missing in GmKate. Both alignments are performed with the entire structures selected and use LSS (Least Square Superposition) method (C) π^{+} - π interaction between R197 and chromophore in red GmKate (upper panel) and green GmKate (lower panel). The guanidinium group is less parallel to the chromophore in the green GmKate. The conformational different of R197 is high lighted in the inset graph. (D) The entire hydrogen bond network around chromophore region. Only the side chain of surrounding residues are shown.

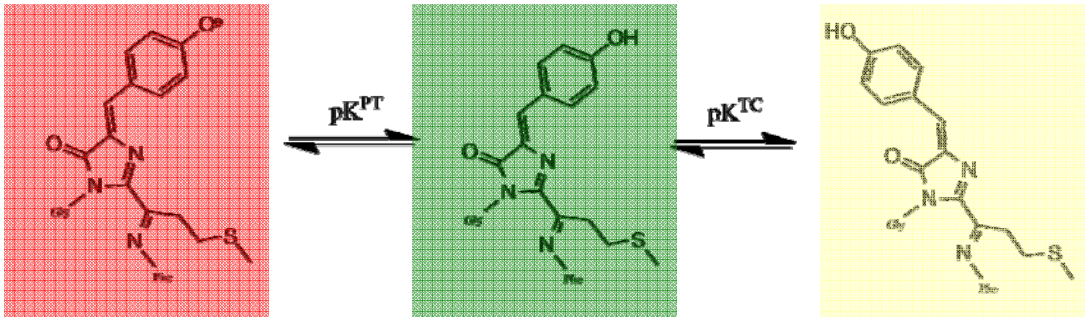


Figure 6.4. Coupled photon-transfer isomorhporization model for GmKate. Red panel indicates the planar anionic *cis*-state, green panel is the planar neutral *cis*-state and the light yellow panel is the non-planar neutral trans-state. pK^{PT} and pK^{TC} is termined as 9.5 and 5.5 for GmKate, respectively.

Protonation and deprotonation in GmKate

Based on the crystal structures of GmKate in green and red states, the molecular mechanism of reversible green-to-red transition can now be understood at an atomic level. Here we show that the well-characterized pH dependence isomorhporization is least likely responsible for this color transition process. At pH above 7.0, the majority of chromophore resides in *cis*-conformation, and it is the protonation states that determine the color of GmKate. In the pH range below the transition pKa, the chromophore favors the protonated state, with the absorbance and emission peaks at 445nm/525nm. In the pH range above the transition pKa, chromophore can be stabilized in the anionic state that absorbs 600nm light. Given the fact that GmKate has dual absorbance peaks at 445nm and 600nm at pH 10.0, it is mostly likely that the chromophore is in equilibrium between protonated the state and the deprotonated state. Our data also suggests that the *trans*-conformation is mostly like in a neutral state, since the absorbance maximum remains at 445nm at very low pH. The loose of fluorescence is most likely due to the non-planar nature of the *trans*-conformation. The pH dependence of GmKate involves two coupled processes,

proton transfer (with pKa^{PT}) and *trans*-*cis* isomorhporization (with pKa^{TC}) (Figure 6.4).

The optical behavior is thus determined by the value of pKa^{PT} and pKa^{TC}. When pKa^{PT} is much larger the pKa^{TC}, chromophore can stay in the planar anionic cis-conformer ($pH > pKa^{PT}$), planar neutral cis conformer ($pKa^{TC} < pH < pKa^{PT}$) and the non-planar neutral trans-conformer ($pH < pKa^{TC}$). For GmKate, the pKa^{PT} is 9.5±0.3 and the pKa^{TC} is 5.5. The pKa^{TC} is determined from the pH 5.5 GmKate structure which contains equal amount of *trans*- and *cis*- conformers. Therefore pH titration of GmKate results in red, green and non-fluorescent species; When the pKa^{PT} is smaller than pKa^{TC}, chromophore can stay in the planar anionic cis-conformer ($pH > pKa^{PT}$), non-planar anionic *trans*-conformer ($pKa^{TC} < pH < pKa^{PT}$) and the non-planar neutral *trans*-conformer ($pH < pKa^{TC}$). This scenario most likely explains the optical behavior of $mKate^{S158A}$. $mKate^{S158A}$ has absorbance peaks at 445nm at very low pH. The apparent pKa^{PT} is 4.6±0.3, as determined from the 445nm absorbance pH titration experiment. The pKa^{TC} for $mKate^{S158A}$ is 5.5±0.3. At pH above 7.0, the entire chromophore resides in the anionic cis-conformation as seen in crystal structure, and emits red light. Further pH decrease will switch chromophore into the trans-conformation that is non-fluorescent, regardless its protonation status. We refer this model as coupled proton-transfer isomphorization model.

CONCLUSION

Here we report the identification and biophysical analysis of an mKate variant that can undergo reversible green-to-red color switching. The transition happens at the millisecond time scale, much faster than other green-to-red photo-convertible fluorescent proteins. We further described the molecular mechanism underlying pH-dependent color switch based on high-resolution crystal structures and spectrum

studies. The intriguing reversible color switch of GmKate is modulated by the chromophore surrounding residues that determine its protonation status. The green-colored chromophore stays in a neutral conformation and the red-colored chromophore stays in the anionic state, which is stabilized by the R197 through cation-pi interaction and pi-pi stacking interaction. Other chromophore surrounding residues may also contribute to this process. The proton transition is proposed to be regulated by a water mediated hydrogen network in the chromophore region. Our results indicate that in addition to the isomorhporization, chromophore proton transition also plays an important role in determining the optical properties for the mKate related fluorescent proteins. Our coupled proton transfer-isomphorphoration model very well explains the different optical properties of Gmkate and mKate^{S158A} in a wide pH range.

The complex nature of the chromophore chemical configurations underlying this mechanism carries important implications for the design of genetically encoded pH sensor and other molecular sensors based on the mKate platform. In particular, our work suggests that the transition between red emission and green emission is not necessarily evolving covalent modification or external energy such as laser activation. This potentially provides great conveniences to change a red-colored sensor to a green colored sensor through genetic modifications, instead of designing from scratch. The direct application of GmKate as a pH sensor is currently hindered by two major challenges. First, the pKa^{PT} is way too high than the expected physiology range; second, the quantum yield of both green state and red state are relatively low when compared with EGFP and mKate. However, neither problem is unsolvable. Further attempts to overcome these problems include evolutionary mutegenesis to improve its *in vivo* brightness and structural based protein engineering to alter its pKa^{PT} and pKa^{TC}.

REFERENCES

Adam V., Carpentier P., Violot S., Lelimousin M., Darnault C., Nienhaus G.U., Bourgeois D. (2009) Structural basis of X-ray-induced transient photobleaching in a photoactivatable green fluorescent protein. J. Am. Chem. Soc. 131(50):18063-5.

Ando R., Hama H., Yamamoto-Hino M., Mizuno H., Miyawaki A. (2002) An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein Proc. Natl. Acad. Sci. USA 99: 12651–12656.

Andresen M., Wahl M.C., Stiel A.C., Gräter F., Schäfer L.V., Trowitzsch S., Weber G., Eggeling C., Grubmüller H., Hell S.W., Jakobs S.(2005) Structure and mechanism of the reversible photoswitch of a fluorescent protein. Proc. Natl. Acad. Sci. USA.13 ;102 (37):13070-4.

Amat P., Granucci G., Buda F., Persico M., Tozzini V. (2006) The chromophore of asFP595: a theoretical study. J. Phys. Chem. B. 110(18):9348-53.

Bravaya K.B., Bochenkova A.V., Granovsky A.A., Savitsky A.P., Nemukhin A.V.(2008) Modeling photoabsorption of the asFP595 chromophore. J. Phys. Chem. A. 112 (37):8804-10.

Brejc K., Sixma T.K., Kitts P.A., Kain S.R., Tsien R.Y., Ormö M., Remington S.J. (1997) Structural basis for dual excitation and photoisomerization of the Aequorea victoria green fluorescent protein. Proc. Natl. Acad. Sci. 94: 2306-2311.

Grigorenko B., Savitsky A., Topol I., Burt S., Nemukhin A.(2006) Ground-state structures and vertical excitations for the kindling fluorescent protein asFP595. J. Phys. Chem. B. 110(37):18635-40.

Giepmans B.N.G., Adams S.R., Ellisman M.H., Tsien R.Y. (2006) The fluorescent toolbox for assessing protein location and function. Science. 312: 217-224.

Gurskaya N.G., Verkhusha V.V., Shcheglov A.S., Staroverov D.B., Chepurnykh T.V., Fradkov A.F., Lukyanov S., Lukyanov K.A. (2006) Engineering of a monomeric green-to-red photoactivatable fluorescent protein induced by blue light. Nat. Biotechnol. 24(4):461-5.

Lin M.Z., et al., (2009) Autofluorescent proteins with excitation in the optical window for intravital imaging in mammals. Chem. Biol. 6(11):1169-79.
Merzlyak E.M. et al., Bright far-red fluorescent protein for whole-body imaging. (2007) Nat. Methods. 4(9):741-6.

Miyawaki A., Tsien RY. (2000) Monitoring protein conformations and interactions by fluorescence resonance energy transfer between mutants of green fluorescent protein. Methods. Enzymol. 327: 472-500

Nienhaus G.U. et al. (2006) Photoconvertible fluorescent protein EosFP: biophysical properties and cell biology applications. Photochem. Photobiol. 2006 Mar-Apr;82(2):351-8.

Nienhaus K., Nar H., Heilker R., Wiedenmann J., Nienhaus G.U. (2008) Trans-cis isomerization is responsible for the red-shifted fluorescence in variants of the red fluorescent protein eqFP611. J. Am. Chem. Soc. 130(38):12578-9.

Nienhaus K., Nienhaus G.U., Wiedenmann J., Nar H.(2005) Structural basis for photo-induced protein cleavage and green-to-red conversion of fluorescent protein EosFP.

Ormö M., Cubitt A.B., Kallio K., Gross L.A., Tsien R.Y., Remington S.J. (1996) Crystal structure of Aequorea victoria green fluorescent protein. Science. 273: 1392-1395.

Pletnev S., et al., (2008) A crystallographic study of bright far-red fluorescent protein mKate reveals pH-induced cis-trans isomerization of the chromophore. J. Biol. Chem. 24; 283(43):28980-7

Quillin M.L., Anstrom D.M., Shu X., O'Leary S., Kallio K., Chudakov D.M., Remington S.J. (2005) Kindling fluorescent protein from Anemonia sulcata: dark-state structure at 1.38 A resolution. Biochemistry. 44(15):5774-87.

Schäfer L.V., Groenhof G., Klingen A.R., Ullmann G.M., Boggio-Pasqua M., Robb M.A, Grubmüller H.(2007) Photoswitching of the fluorescent protein asFP595: mechanism, proton pathways, and absorption spectra. Angew. Chem. Int. Ed. Engl. 46(4):530-6.

Shaner N.C., Steinbach P.A., Tsien R.Y. (2005) A guide to choosing fluorescent proteins, Nature Methods 2: 905 – 909

Shcherbo D. et al. (2009) Far-red fluorescent tags for protein imaging in living tissues. Biochem. J. 418(3):567-74.

Stiel A.C., Trowitzsch S., Weber G., Andresen M., Eggeling C., Hell S.W., Jakobs S., Wahl M.C. (2007) 1.8 Å bright-state structure of the reversibly switchable fluorescent protein Dronpa guides the generation of fast switching variants. Biochem. J. 402(1):35-42.

Tsien R.Y. (1998) Annu. Rev. Biochem.1998, 67, 510−544

Wang L., Jackson W.C., Steinbach P.A., Tsien R.Y. (2004) Evolution of new nonantibody proteins via iterative somatic hypermutation. Proc. Natl. Acad. Sci. USA 101, 16745-16749 (2004).

Wiedenmann J., Ivanchenko S., Oswald F., Schmitt F., Röcker C., Salih A., Spindler K.D., Nienhaus G.U. (2004) EosFP, a fluorescent marker protein with UV-inducible green-to-red fluorescence conversion. Proc. Natl. Acad. Sci. U S A. 101(45):15905-10.

Yampolsky I.V., Remington S.J., Martynov V.I., Potapov V.K., Lukyanov S., Lukyanov K.A.(2005) Synthesis and properties of the chromophore of the asFP595 chromoprotein from Anemonia sulcata. Biochemistry. 44(15):5788-93.

Yang F, Moss L.G., Phillips GN.(1996) Structure and fluorescence mechanism of GFP. Nature Biotech. 14, 1246-1251

CHAPTER 7

CONCLUSIONS AND FUTURE DIRECTIONS

BAR SUPER-FAMILY PROTEINS

Protein-mediated membrane deformation presents a general mechanism that cells use to control membrane curvatures in dynamic membrane remodeling processes. this thesis systematically studied how two players, proteins and membranes, contribute to the membrane deformation process during CME. Particularly, this thesis establishes that the interplay between proteins and membranes are bi-directional and of high complexity. Here we show that BAR domain-containing protein Pacsin can deform liposomes into diverse membrane shapes including small vesicles and tubes of variable diameters. The unique protein molecular structure, featured by a rigid membrane interacting surface and an amphipathic peptide region, is the driving force of the membrane curvature change.

The intra- and inter- molecule interactions appear to accurately regulate the membrane deformation activity of Pacsin. The presence of the C-terminal SH3 domain significantly impairs its membrane sculpture activity through an autoinhibitory mechanism. The inhibited activity can be recovered by the non-catalytic PRD domain from dynamin, through interactions with Pacsin SH3 domain. Here we show that Pacsin-dynamin^{PRD} complex has fission activity *in vitro*. Prior to our study, The tubulation activity defines BAR domain-containing proteins as curvature initiators and tubulators (Peter et al., 2004; Shimata et al., 2007; Henne et al., 2007; Frost et al., 2008; Henne et al., 2010). Recent studies suggest that Pacsin-dynamin complex plays an important role in regulating the synaptic vesicle recycling process when the neurons are subject to intense stimulations (Andersson et al., 2008; Anggono et al., 2006). Based on our *in vitro* study, we propose that this may be due to the unique fission activity of the complex that increases the vesicle generation rates.

The biophysical studies on membranes indicate that the remodeling process is sensitive to the membrane physical-chemical properties, including membrane composition, curvature, phase and rigidity. Environmental factors such as temperature and incubation time also play roles in membrane remodeling processes. The decoupling of individual contribution is not impossible, but proves to be challenging given the complexity of membrane systems. Our work suggests that membrane properties are indispensible in analyzing membrane remodeling processes, and we provide a systematic scheme on how to analysis different aspects of membrane contributions.

The assembling and disassembling mechanisms of BAR domain proteins remain mystery. In the assembling process, BAR domain proteins are recruited from the bulk solution on to a membrane surface and form a regular array. The nucleation step may be achieved by the spontaneous fusion of several randomly distributed BAR domains protein patches on the membrane surface (Dommersnes and Fournier, 2002) or through the specific lipid-directed membrane localization if BAR domains are linked with certain lipid interacting domains. Once the nucleation is achieved, other BAR domains may be further recruited through a curvature-driven positive feedback (McMahon and Gallop, 2005). The *in vivo* disassembling mechanism is much less understood. *In vitro*, deformed membrane tubules are still coated by the protein arrays, suggesting that the disassembling is least likely spontanouse. Some other cellcular components must facilicate disassembling process. Studies from other membrane coating proteins shed light on this problem. Dynamin disassembles from membranes due to a conformational change upon GTP hydrolysis (Bashkirov et al., 2009; Stowell et al., 1999; Roux et al., 2006). Some BAR domains are linked with GTPase domains and may use a similar mechanism for their disassembling; the binding of other cellular components onto BAR domains is another way to trigger disassembling, such as the

process of Clathrin uncoating via the local conformation distort induced by Auxilin/Hsc70 (Fotin et al., 2004). In addition, the phosphorylation of unique tyrosine residues at BAR domain dimerization interface can also disassemble the dimer and trigger the disassembling.

Despite great progress in understanding BAR-domain mediated membrane deformation *in vitro*, *in vivo* functions of BAR domains are largely unknown. Studies from FCHo2 suggest that it involves in defining the position of endocytic site, in addition to stabilizing the budding neck structures (Henne et al., 2010). Our study indicates that Pacsin may be involved in the membrane fission of dynamin dependent endocytosis. For other BAR superfamily members, there appears to be no direct correlation between their *in vitro* activities and *in vivo* functions. F-BAR domain from Pacsin 2 generates both tubular invaginations and micro-spikes when expressing in cells (Shimata et al., 2010). These reports suggest that the insights from isolated BAR domains may only partially reflect their functions *in vivo*, where BAR domains are dynamically interacting with other proteins. Characterizations of BAR domain function at the presence of other cellular components will be a substantial step for revealing their cellular functions.

It is also worth mentioning that BAR superfamily proteins are not the only peripheral membrane proteins that can drive membrane curvature change. Indeed, the tubulation activities seem to be very redundant for the protein machineries involved in CME and different BAR and F-BAR domain proteins may be present in the same process. This apparent functional redundancy indicates that the molecular mechanism of membrane remodeling is highly sophisticated. The functional and regulatory discrepancies among different membrane remodelers define their unique roles in modulating membrane curvatures, which is largely unknown. *In vitro* reconstitution of CME process appears to be a very promising way to decouple the complexity of

membrane deformation and reveal the mystery that underlines the membrane remodeling process.

GENETICALLY ENCODED CALCIUM INDICATORS

The single-fluorophore based genetically encoded calcium indicators (GECIs) possess great advantages to detect calcium signals in various cellular organelles, thus are very promising for the clinical use. Previous studies have made great progress towards designing the high performance GECIs of robust signal response and fast kinetics (Tallini et al 2006; Tian et al., 2009). Some well established sensors such as GCAMPs and Camgaroos have already been successfully used to capture the calcium flux transition from diastoles to systole state in spontaneously beating heart of the anaesthetized and ventilated mouse (Tallini et al., 2006) and directly monitor neuronal activity when expressed in the motor cortex of the mouse (Tian et al., 2009). The further optimization of current sensors and the development of novel sensors are hindered by the fact that little is known about the molecular mechanism of GECIs illumination.

Our major result of this part is the study of GCAMP sensors illumination mechanism based on the crystal structure information. From our biophysical characterizations of GCAMP2 in the distinct states, we found the C-terminal calmodulin domain directly participates regulating the chromophore brightness, by covering an artificially engineering hole on the β-barrel of GFP. We refer this mechanism as "cap-the-hole" mechanism. Many aspects of the particular design of GCaMP2 are likely to contribute to the success of this Ca^{2+} indicator, including the particular point of permutation allowing for reversible solvent exposure of the fluorophore, the complementarity between interfaces, and the functional conservation of fluorophore-coordinating residues. Our results indicate that circular permutation of

EGFP destabilizes the deprotonated state of the fluorophore, and that this state is re-established in a Ca^{2+}-dependent manner through intensive contacts between the condensed CaM:M13 structure and EGFP moieties. A pre-docked state of CaM in the absence of Ca^{2+} might contribute to the rapid transition between the two fluorophore states.

Structure-based fluorescent protein design has made significant contributions to the recent fluorescent protein optimization and sensor development. The enhanced red fluorescent protein mKate2 contains a critical mutation $S^{158}A$ that has been proposed to destabilize the non-fluorescent *trans*-state based on the crystal structures of mKate (Shcherbo et al., 2009). The most far-red excitation sensor mNeptune also contains the $S^{158}C$ mutation (Lin et al., 2009). GCAMP3 was developed primarily based on the mutation T03V that has been identified from GCAMP2 crystal structure that resulted in a sensor with more robust signal response. Obtaining molecular mechanisms of fluorescent proteins based on the crystal structures have greatly increases the sensor development efficiency.

The sensor kinetics is mainly limited by the intra-molecule domain motion kenetics and the chromophore response kenetics. The intra-molecular motion involves large protein domain rearrangement and usually happens at the millisecond to second time scale. In the case of GCAMPs, this becomes the rate limiting step, since the chromophore protonation/deprotonation can be achieved at the sub-millisecond time scale. For some other fluorescent proteins, especially these of complicated chromophore maturation steps, the chromophore response can become the rate limiting step and significantly impair sensors' performance.

Developing novel GECIs with red emission colors will provide great convenience for the practical clinical use. The chromophore illumination mechanisms of Red Fluoresent Proteins are in general less understood than these of Green

Fluoresent Proteins, and this adds more difficulties in engineering the red colored proteins. Nevertheless, "cap-the-hole" mechanism together with other GECI strategies presents promising ways for developing red-colored emission calcium sensors. Another major direction of sensor development is to design sensors that can detect other cellular signaling messengers such as H^+, Zn^{2+} and di-GMP. "cap-the-hole" mechanism has high potential to be applied for this category sensor development, since mechanism is not restricted to the M13: Calmodulin moiety but can be extended to other protein modules that can response to other signals (Figure 7.1).

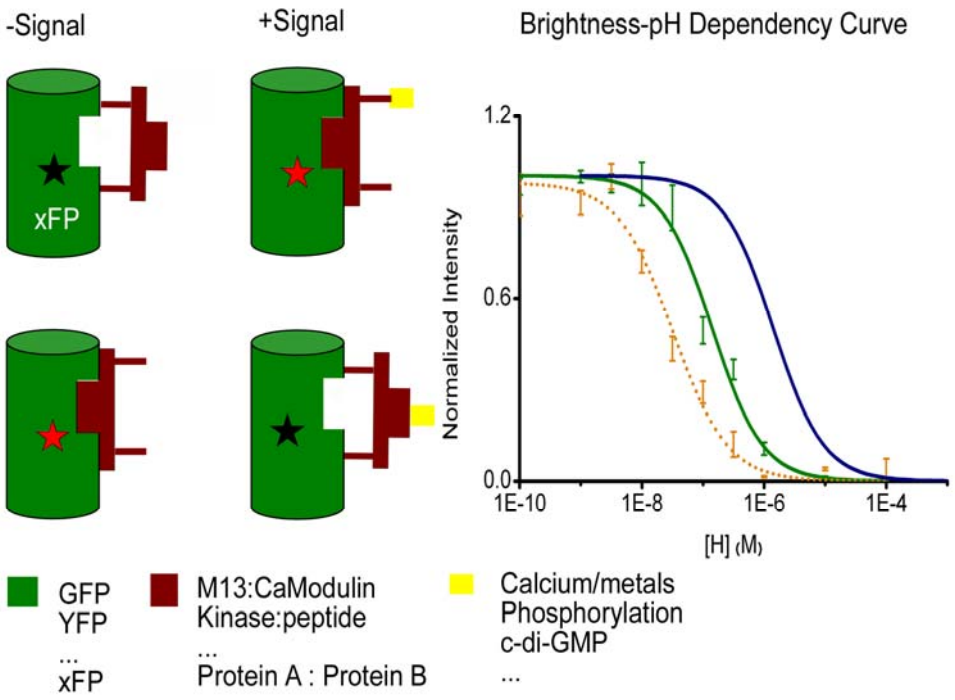


Figure 7.1. "cap-the-hole" sensor design principle is based on the protonation status change of GFP chromophore. Left panel: the design principle of signal response; right panel: The differences of apo-sensor (orange) and signal bound sensor (green) in response to the external pH. The gap between two curves determins the dynamic range of sensors and their performances. The blue curve shows the most promising status for a signal bound sensor.

The second major contribution of this thesis is the mechanistic study of GmKate, the fluorescent protein that can undergo reversible pH dependent green-to-red color conversion. Our study unexpectedly reveals that the green-to-red color transition does not necessarily involve chemical modifications in or near chromophore region, as observed from photo-convertible fluorescent proteins. We have demonstrated that the color switch can be achieved by regulating the protonation status of chromophore, which provides a fast and efficient way to achieve the color transition. This mechanism in general can be used for developing other types of color switchable proteins that requires fast transition kinetics.

REFERENCES

Andersson F., Jakobsson J., Löw P., Shupliakov O., Brodin L. (2008) Perturbation of syndapin/PACSIN impairs synaptic vesicle recycling evoked by intense stimulation. J. Neurosci. 28:3925-3933.

Anggono V., et al. (2006) Syndapin I is the phosphorylation-regulated dynamin I partner in synaptic vesicle endocytosis. Nat. Neurosci. 9:752-760.

Bashkirov P.V., et al. (2008) GTPase cycle of dynamin is coupled to membrane squeeze and release, leading to spontaneous fission. Cell 135:1276-1286.

Braun A., et al. (2005) EHD proteins associate with syndapin I and II and such interactions play a crucial role in endosomal recycling. Mol. Biol. Cell. 16:3642-3658.

Campelo F., McMahon H.T., Kozlov M.M .(2008) The hydrophobic insertion mechanism of membrane curvature generation by proteins. Biophys. J. 95:2325-2339.

DiProspero N.A., et al. (2004) Early changes in Huntington's disease patient brains involve alterations in cytoskeletal and synaptic elements. J. Neurocytol. 33:517-533.

Dommersnes P.G., Fournier J.B. (2002) The many-body problem for anisotropic membrane inclusions and the self-assembly of "saddle" defects into an "egg carton". Biophys. J. 83:2898-2905.

Frost A., et al. (2008) Structural basis of membrane invagination by F-BAR domains. Cell 132:807-817.

Fotin A, et al. (2004) Molecular model for a complete clathrin lattice from electron cryomicroscopy. Nature 432:573-579.

Henne W.M., et al. (2007) Structure and analysis of FCHo2 F-BAR domain: a dimerizing and membrane recruitment module that effects membrane curvature. Structure 15:839-852.

Itoh T., et al. (2005) Dynamin and the actin cytoskeleton cooperatively regulate plasma membrane invagination by BAR and F-BAR proteins. Dev. Cell 9:791-804.

Peter B.J., Kent H.M., Mills I.G., Vallis Y., Butler P.J., Evans P.R., McMahon H.T. (2004) BAR domains as sensors of membrane curvature: the amphiphysin BAR structure. Science. 303(5657):495-9.

Pucadyil T.J., Schmid S.L. (2008) Real-time visualization of dynamin-catalyzed membrane fission and vesicle release. Cell 135:1263-1275.

Tallini Y.N., Brekke J.F., Shui B., Doran R., Hwang S.M., Nakai J., Salama G., Segal S.S., Kotlikoff M.I. (2007). Propagated endothelial Ca2+ waves and arteriolar dilation in vivo: measurements in Cx40BAC GCaMP2 transgenic mice. Circ. Res. 101, 1300-1309.

Tallini Y.N., Ohkura M., Choi B.R., Ji G., Imoto K., Doran R., Lee J., Plan P., Wilson J., Xin H.B., et al. (2006). Imaging cellular signals in the heart in vivo: Cardiac expression of the high-signal Ca2+ indicator GCaMP2. Proc. Natl. Acad. Sci. U S A 103, 4753-4758.

Tsien R.Y. (1998). The green fluorescent protein. Annu Rev Biochem 67, 509-544.

Wang Q., Kaan H.Y., Hooda R.N., Goh S.L., Sondermann H. (2008) Structure and plasticity of Endophilin and Sorting Nexin 9. Structure 16:1574-87.

Zimmerberg J., Kozlov M. (2006) How proteins produce cellular membrane curvature. Nat. Rev. Mol. Cell Biol. 7:9-19.