Biochemical and Structural Studies of Membrane Proteins

Citation

Wang, Ruiqi Rachel. 2012. Biochemical and Structural Studies of Membrane Proteins. Doctoral dissertation, Harvard University.

Dissertation Advisor: Thomas Walz Ruiqi Rachel Wang

Biochemical and Structural Studies of Membrane Proteins

Abstract

Membrane proteins live at the interface between a cell and its environment; hence, they play a variety of important physiological roles such as transmembrane transport, signal transduction, and cell adhesion. The importance of membrane proteins in biology and medicine requires that we understand their structure and function on the atomic level. In this thesis, I studied members of two different membrane protein families, namely the neuronal and keratinocyte TRPV ion channels that sense temperature changes and MP20, a member of the PMP22/EMP/MP20/claudin superfamily. Using a variety of biochemical, X-ray crystallographic and electrophysiological techniques, I addressed mechanistic questions pertaining to the regulation of thermosensitive TRPV channels by ATP and calmodulin in neurons and keratinocytes. For MP20, a protein specific for the lens of the mammalian eye, I used a vesicle assay in combination with electron microscopy (EM) to study its function, ruling out the possibility that MP20 is involved in the formation of membrane junctions. Furthermore, I made progress in expressing and crystallizing MP20 for X-ray diffraction studies.

In a separate effort, I also worked on improving and expanding the use of monolayer purification and Affinity Grids, recently introduced techniques to prepare specimens for single-particle EM based on the recruitment of His-tagged proteins to nickel lipid-containing lipid monolayers. I extended the use of these techniques by synthesizing a glutathione lipid that can be used to recruit GST-tagged proteins. A major hurdle in the use of monolayer purification techniques, however, is the extent of non-specific protein

binding to the lipid monolayer. I found that incorporating PEG lipids in the monolayer appears to reduce the problem of non-specific protein binding. While it remains to be seen whether these techniques can be developed to a point at which it will be possible to recruit exclusively tagged proteins out of cell lysates, my goal is to continue to improve and expand the use of the monolayer purification and Affinity Grid techniques in hope to make single-particle EM more easily amenable to biochemists and cell biologists.

Table of Contents

Table of Figures

List of Tables

Acknowledgements

This thesis will not have been possible without the help and support of the people around me.

First and foremost, I want to thank Dr Thomas Walz for his excellent mentorship and guidance. He has been a great source of support that I could always rely on. He has been exceedingly patient with me and has guided and supported me in all my scientific endeavors. Always, he has been a great cheerleader, and my time in his lab has taught me to be independent and strong. There is so much that I have learned from him and so much more I desire to learn from him.

Next, I would like to thank everyone who has served on my dissertation advisory committees and preliminary qualifying examination committee. My heartfelt gratitude goes to Dr Alan Saghatelian who has served on all my committees. I would like to thank him for all his wonderful advice and guidance, and especially for his unwavering and continuous support when I switched labs in the middle of my PhD. In addition, I have also benefitted greatly from the breadth and depth of his knowledge in chemistry and biology. I would also like to thank Dr Stephen Harrison for his very insightful advice on all my committee meetings and beyond. His participation on my committee meetings has sharpened my mind greatly and widened my horizons infinitely. Indeed, I always felt I learned something new, walking out of every single committee meeting. Also, I would like to thank Dr David Paul for being an encouraging member of my dissertation advisory committee. His encouragement gave me a lot of optimism towards my projects. Finally, I would like to thank Dr Daniel Goodenough for being a valuable member on my committee.

Next, I would like to thank Dr Danijela Dukovski and Dr Zongli Li for help within the lab. Thanks for assisting me with my experiments and teaching me how to use the electron microscope. My projects would not have been possible without you. I would also like to thank Dr Andreas Schenk, Dr Hui-Ting Chou and Cecilie Lin for their friendship and very useful scientific discussions. They have enriched my experience in the Walz lab infinitely. Finally, I would like to thank all other members in the Walz lab for companionship, support and general help within the lab.

From outside the lab, there are many people I would like to thank for their generosity and kind assistance with my work. This includes Yui Vinayavekhin, Tejia Zhang and Edwin Homan from the Saghatelian lab; Dr Aaron Schmidt, Dr Florian Schmitzberger, Dr Marco Morelli, Dr Uhn-Soo Cho, Dr Marcelo Berardi, Dr Kevin Corbett, Dr Junhua Pan and Dr Simon Jenni from the Harrison lab; Dr Mohd Altaf Bhat from the Moazed lab; Eugene Drokhlyansky from the Goldberg lab. Prior to switching labs, I learned much and benefitted greatly from working with Dr Christopher Phelps and Dr Wihelm Weihofen. They were instrumental in building up my foundations in biochemistry and molecular biology in general. To this day, Dr Wihelm Weihofen continues to be a great friend and an invaluable source of scientific knowledge and advice. I would also like to thank the

Chemical Biology Program for allowing me to do a PhD in the company of all these great people.

Last but not least, I would like to thank my fiancé, Shyh-Chang Ng. He has been a great source of support and encouragement ever since we met in my second year of graduate school. He has always been there, sharing all my moments, including my happiest and darkest. I am eternally grateful to him for standing by me all the time, for lifting me up when I am sad and for celebrating every little moment of triumph I have. Finally, he has been an excellent role model, my source of pride and happiness and a pillar of strength that I can reliably rest upon. I would also love to thank Simon Ong and Carol Lo for being such wonderful parents. I am extremely grateful to them for supporting me in my decision to do a PhD in chemical biology and enduring another 6 to 8 years of me being away from Singapore. I also thank them for constantly keeping me in their thoughts, constantly planning ever since the day I was born, that I would always have a head start in life. They have been amazing parents, without whom I would not be where I am today.

Chapter 1: Introduction to membrane proteins and methods used to study the structure of membrane proteins

The importance of membrane proteins

Biological membranes define the physical boundaries of a cell and its organelles. They are formed by a bilayer of lipids that creates a semi-permeable barrier. Embedded in these lipid bilayers are membrane proteins, which are responsible for most of the functions of biological membranes and mediate all the interactions of cells and organelles with their environment. Some membrane proteins form selective and often regulated pores, channels or transporters that are responsible for the passive or active (energy-dependent) transport of ions, metabolites and other molecules across the membrane. For instance, TRPV channels allow diffusion of Ca^{2+} in response to specific temperature ranges (Clapham, 2007), and mechanosensitive channels conduct ions and open in response to pressure on the membranes (Haswell *et al.,* 2011). Other membrane proteins function as receptors that sense changes in the environment. For example, rhodopsins convert light signals into electrical signals that can be detected by the brain (Palczewski, 2012). In addition, every signaling pathway starts with a membrane protein. For example, the Notch signaling pathway is activated by the interaction of a Notch receptor with other Notch receptors or ligands on a neighboring cell (Artavanis-Tsakonas *et al.,* 1995). Wnt signaling is initiated by binding of ligands to the membrane receptor Frizzled, which then sets off a downstream signaling cascade (Malbon *et al.,* 2001). In Hedgehog signaling, binding of Hedgehog to the transmembrane receptors Patched and Smoothened triggers downstream events in the pathway (Lum and Beachy, 2004), and phosphorylation of the TGF-β membrane receptor upon ligand binding is the initial event in TGF signaling. Yet other membrane proteins are involved in energy production, such as the complexes in respiration (Karp, 2008) and photosynthesis (Blankenship, 2008), and the ATP synthase, which couples the proton gradient across cell membranes to the generation of ATP, the energy unit used in most cellular reactions (Boyer, 1995). Membrane proteins also catalyze a wide variety of enzymatic reactions. It is thus not surprising that membrane proteins constitute approximately one third of the cell proteome (Wallin and von Heijne, 1998).

Since they perform many important physiological roles and since they are the starting point of every known signaling cascade, membrane proteins are important drug targets. Plasma membrane proteins are particularly attractive drug targets, because they are amenable to drugs that do not permeate membranes. For instance, instead of being limited to small molecule drugs that are membrane-permeable but often lack specificity, plasma membrane proteins can be targeted with more specific drug types such as antibodies. Many drugs currently on the market target membrane proteins, in particular G protein-coupled receptors and ion channels, and many pharmaceutical companies have ongoing large-scale screening efforts aimed at identifying drugs that target disease-relevant membrane proteins.

Challenges of working with membrane proteins

Because of their chemical properties, studying the structure and function of membrane proteins is challenging. Soluble proteins have mostly hydrophilic surfaces and are thus stable in solution. Membrane proteins also have hydrophilic surfaces but in addition feature a hydrophobic belt that anchors them in the lipid bilayer. Because of this amphiphilic nature, membrane proteins are not stable in solution. Studying membrane proteins in solution thus requires the use of lipid mimetics that cover their hydrophobic belts. Detergents are most commonly used for this purpose, but detergents are an imperfect substitution for a lipid bilayer and harsh detergents can denature fragile membrane proteins. Alternatives to detergents have thus been introduced, including nanodiscs (Leitz *et al.*, 2006; Kijac *et al.*, 2007; Raschle *et al.*, 2009), bicelles (Vold *et al.*, 1997; Prosser *et al.*, 2006), and amphipols (Tribet *et al.*, 1996; Bowie, 2001). However, the functions of membrane proteins often depend on the physical presence of a lipid bilayer. Thus, to study activities such as transmembrane transport, membrane proteins have to be reconstituted into lipid bilayers, complicating *in vitro* assays.

Compared with soluble proteins, it is also more difficult to express membrane proteins in high quantities. Soluble proteins are expressed and fold in the cytoplasm of a cell or are secreted into the environment, making it possible to express them in large amounts. In contrast, membrane proteins need to be inserted into cell membranes, where they adopt their native fold. Since the available membrane area and the number of membrane insertion machineries are limited, the amount of membrane protein that can be expressed in a cell is much smaller than for soluble proteins. The often low expression levels have severely impeded structural and functional studies of membrane proteins.

In summary, membrane proteins have unique characteristics that make them challenging to study *in vitro*, both for structural and functional studies.

Importance of knowing the structure of proteins

Genetic and biochemical studies are powerful approaches to establish the biological functions of proteins, but they are limited in understanding how proteins perform their function. Structural information can provide direct mechanistic insights into the workings of proteins, and structures at atomic resolution reveal the chemical basis that allows proteins to perform their function. Such information is extremely valuable in understanding how drugs affect protein function and also allows drugs to be designed to be more effective and/or more specific for the target protein. The three most commonly used techniques to determine protein structures are x-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, and electron microscopy (EM).

X-ray crystallography

X-ray crystallography involves growing 3D crystals of proteins that can be analyzed by x-ray diffraction. Protein crystals are usually produced by adding a precipitant to a concentrated protein solution to induce crystal formation. The growth of protein crystals is mostly a trial and error procedure, and often requires thousands of crystallization conditions to be screened to identify conditions that promote growth of sufficiently large and well-ordered protein crystals for analysis. Suitable crystals are then used to collect x-ray diffraction patterns. After phasing the diffraction data set using one of several established methods, an electron density map is calculated that can be used to build and refine an atomic model of the crystallized protein. X-ray crystallography is the most widely used structural biology approach and yielded structures of a wide variety of proteins.

NMR spectroscopy

NMR spectroscopy measures the resonance of atomic nuclei such as ^{13}C and ^{1}H of proteins placed in a magnetic field (Keeler, 2005). The protein spectrum reveals how the nuclear spins of its atoms are affected by the neighboring atoms, thus providing information on the chemical environment of each atom (Keeler, 2005). By mapping out information about the environment of each atom, NMR allows one to derive structural knowledge about the protein (Keeler, 2005). Unlike in crystallographic methods, in which proteins are typically constrained in a particular conformation, NMR data are recorded of unconstrained proteins in solution. NMR spectra thus also contain valuable information about the dynamics of the proteins. Due to current technical limitations, however, NMR can only be used routinely to determine the structure of proteins that are smaller than ~30 kDa in size. This limitation excludes many proteins from structure determination by NMR spectroscopy.

Electron microscopy

In transmission EM, an electron beam is used to record a magnified image of the specimen. Due to the very short wavelength of electrons, it is theoretically possible to obtain images that contain structural information at atomic resolution. Three distinct approaches are used in molecular EM: single-particle EM, electron crystallography and electron tomography, but only the former two can provide structural information at high resolution.

Electron crystallography requires 2D crystalline specimens, which can be produced for both soluble and membrane proteins. The most common method for 2D crystallization of soluble proteins is monolayer crystallization, in which soluble proteins are adsorbed to a lipid monolayer at an air/water interface to induce them to form 2D arrays (Uzgiris and Kornberg, 1983; Ribi *et al*., 1987; Darst *et al.,* 1988; Kubalek *et al.,* 1994; Thess *et al.,* 2002). Membrane proteins are reconstituted with lipids at a low-lipid-to-protein ratio to form 2D arrays (Hite *et al.*, 2007; Raunser and Walz, 2009). Electron crystallography is especially useful for membrane proteins, because it allows their structure to be analyzed in their native environment, a lipid bilayer (Hite *et al.*, 2007; Raunser and Walz, 2009). As with 3D crystallization, the growth of 2D crystals is largely a trial and error procedure. Once suitable 2D crystals have been obtained, data collection typically entails recording high-resolution images to obtain phase information and electron diffraction patterns to obtain more accurate amplitude values than those that can be extracted from images (Raunser and Walz, 2009). As in x-ray crystallography, the resulting density map can be used to model and refine the structure of the crystallized protein. Unlike x-ray crystallography, however, specimen preparation, data collection and data analysis are arduous and time-consuming (Hite *et al.*, 2010), and compared with x-ray crystallography, only few structures have been determined by electron crystallography.

Single-particle EM does not depend on crystalline specimens. Structural information is obtained by first imaging many copies of a protein in different orientations and then computationally combining the different projections to reconstruct a 3D density map of the imaged protein. To prepare proteins for the vacuum of the electron microscope, they are either negatively stained, i.e., embedded in a layer of heavy metal crystals, or vitrified, i.e., embedded in a layer of amorphous ice (Cheng and Walz, 2008). Negative staining has the advantage that it creates additional contrast, because heavy metals are better electron scatterers than the light atoms that make up biological specimens. At the same time, the heavy metal crystals limit the resolution that can be obtained with negatively stained specimens to about 20 Å, and the drying of the specimen introduces flattening artifacts. Conversely, vitrification does not add contrast, and therefore proteins prepared by this method currently have to be bigger than ~250 kDa to be visible in the ice layer. At the same time, the resolution that can be obtained with vitrified specimens is not limited by the preparation method, and vitrification does not introduce any preparation artifacts, preserving the proteins in a near-native environment. Single-particle EM reconstructions usually do not provide near-atomic resolution information, but the method has other advantages. Compared with X-ray crystallography and NMR spectroscopy, much less material is needed and at a much lower concentration. In addition, single-particle EM can be used to study structurally heterogeneous samples and thus allows, for example, visualization of a protein in different conformations (Cheng and Walz, 2008).

Challenges of determining structures of membrane proteins

Given their biological and medical importance, knowing the structure of membrane proteins is of crucial importance. However, as outlined above, structure determination of membrane proteins is complicated by their amphiphilic nature. Several new approaches have been developed to address the challenges posed by membrane proteins.

Obtaining suitable crystals of membrane proteins for analysis by x-ray crystallography, often requires screening of multiple detergents to identify one that can stabilize the protein but does not interfere with the formation of crystal contacts. In addition, detergents can occasionally interfere with the native conformation of a membrane protein. To overcome these problems, new strategies have been developed such as crystallization in lipidic cubic phases (Landau and Rosenbusch 1996; Pebay-Peyroula *et al.*, 1997), lipid bicelles (Faham and Bowie, 2002), and inclusion of lipids in the purification and crystallization of the protein (Long *et al.*, 2007).

Structure determination by NMR is typically limited to proteins smaller than 30 kDa. This size limitation is particularly serious for membrane proteins, because their size is further increased by the presence of associated detergent or other ampiphile molecules needed to stabilize them in solution. Approaches used to circumvent this problem include solid state NMR (Hong *et al.,* 2012) and fragment based coupling, used in conjunction with a partially aligned medium (Berardi *et al.,* 2011). Unlike for solution NMR, there is no theoretical molecular weight limit for solid state NMR. Solid state NMR is usually ideal for molecules that exhibit anisotropic mobility, which is the case for membrane proteins reconstituted into lipid bilayers (Hong *et al.,* 2012). The second approach uses solution NMR in a partially aligned medium (Hong *et al.,* 2012). In this case, NMR data of the protein of interest are collected, for example, in the presence of DNA nanotubes, which partially align the proteins (Berardi *et al.,* 2011). The measured data are then matched to values derived from a database constructed from known protein structures. This method was first employed to solve the structure of UCP2, a mitochondrial carrier protein (Berardi *et al.,* 2011).

For single-particle EM, specimens are prepared either by negative staining or vitrification. Many membrane proteins are too small to be seen in a vitrified ice layer, and negative staining limits the resolution that can be obtained. Furthermore, detergents needed to stabilize membrane proteins in solution reduce the surface tension of the buffer, making it challenging to prepare thin ice layers. In addition, detergents also increase the density of the buffer and thus reduce the contrast the proteins generate. Ways that are currently explored to address these problems include the addition of Fabs to increase the size of the proteins, the development of phase plates to increase the image contrast, and direct electron counter cameras for more sensitive detection of signals.

Electron crystallography is in theory ideal to study the structure of membrane proteins, because the proteins are reconstituted into a lipid bilayer, which best mimics their natural environment. However, as in x-ray crystallography, it is difficult to obtain the large and well-ordered crystals needed for high-resolution data collection. In addition, the time, effort and equipment needed to prepare specimens, collect and analyze data make electron crystallography a technique that is not widely used. To increase the efficiency of producing 2D crystals, efforts are underway to automate the screening of 2D crystallization conditions (Iacovache *et al.*, 2010; Coudray *et al.,* 2011). In addition, new software is being developed to accelerate data processing (Gipson *et al.*, 2007a; Gipson *et al.*, 2007b; Philippsen *et al.*, 2003; Philippsen *et al.*, 2007; Schenk *et al.,* 2010). Nevertheless, much more effort will be needed to bring electron crystallography to a level at which it can be routinely used to determine membrane protein structures.

Chapter 2.1: Introduction to thermosensitive TRPV channels

Living organisms are able to detect temperature changes in the environment by converting local changes in temperature into chemical signals. This ability is extremely important for processes crucial for survival such as homeostasis and pain sensation. The Transient Receptor Potential (TRP) channel superfamily consists of six different families of ion channels that perform a vast variety of physiological functions, of which many pertain to our senses of smell, taste and touch (Venkatachalam and Montell, 2004). The TRPV subfamily consists of six family members, of which four are non-selective cation channels that activate upon sensing certain temperatures. TRPV1 is activated at temperatures above 43°C, and TRPV2 is activated at above 52°C (Caterina *et al.,* 1999). The temperatures sensed by these two channels are sufficiently high to cause pain in mammals, explaining why these two channels are mainly found in sensory neurons that are involved in pain sensation. TRPV3 and TRPV4 are activated at more ambient temperatures, between 32 and 39°C for TRPV3 (Peier *et al.,* 2002; Smith *et al.,* 2002; Xu *et al.,* 2002) and between 25 and 34°C for TRPV4 (Guler *et al.*, 2002; Watanabe *et al.,* 2002). Consistent with their function of sensing warm, innocuous temperatures, TRPV3 and TRPV4 are predominantly localized in keratinocytes and are potentially involved in regulating cutaneous thermal homeostasis (Lee and Caterina, 2005).

One outstanding question in this field is how TRPV channels are able to sense different temperature thresholds, such that organisms can mount different responses to the different temperatures. One clue comes from the manner by which TRP channels are regulated and sensitized. While all thermoTRPV channels are known to be non-selective cation channels, it is known that it is the flux of Ca^{2+} that signals to downstream components of the pathway. Indeed, the activities of TRPV channels are modulated intracellularly by Ca^{2+}, calmodulin (CaM), nucleotides, and phosphoinositides (Prescott and Julius, 2003). Desensitization of TRPV1 increases with increasing levels of both intracellular and extracellular Ca^{2+} levels (Koplas *et al.,* 1997). Furthermore, it has been demonstrated that removal of intracellular ATP and GTP causes tachyphylaxis of TRPV1 in response to agonist (Koplas *et al.,* 1997). Similarly, TRPV4 is first potentiated and then inactivated by intracellular Ca^{2+}, likely mediated by CaM (Strotmann *et al.,* 2003). Both TRPV1 and TRPV4 are desensitized after repeated and prolonged sensitization (Güler *et al.,* 2002). While the sensitivity of TRPV3 also depends on Ca^{2+} and CaM, it experiences increased sensitization upon repeated stimulation (Peier *et al.,* 2002). The mechanisms underlying these similarities and differences in regulation patterns remained largely unknown. Structural data for these channels is needed to obtain mechanistic insights into the regulation of thermoTRPVs and to fully understand how these channels work to help the body sense different temperatures.

While little is known about the structure of TRPV channels, the available biochemical and structural evidence points towards the importance of the cytoplasmic domains for their function. TRPV channels being part of the TRP channel superfamily share a somewhat common topology with all the other family members. All TRP channels are homotetrameric and their transmembrane domain is predicted to be very similar to that of the Shaker K^+ channel. Each of the four subunits contains six transmembrane helices, and helices 5 and 6 assemble to form the non-selective cation pore. Unlike other members of the family, TRPV channels contain an N-terminal cytoplasmic domain consisting of six ankyrin repeats. Each ankyrin repeat consists of 33 amino acids that form a helix-turned-helix motif. Ankyrin repeats are found in many proteins and are commonly involved in mediating protein-ligand interactions (Sedgwick and Smerdon, 1999; Mosavi *et al.,* 2004). In addition, TRPV channels also contain a C-terminal cytoplasmic domain that is less well characterized. To date, with no high-resolution structure for any TRPV channel, the only available structural data are low-resolution structures obtained by cryo-negative stain EM of TRPV1 and TRPV4 (Moiseenkova-Bell *et al.,* 2008; Shigematsu *et al.,* 2010). These EM reconstructions show a large density corresponding to the cytoplasmic domains of the TRPV channels (Moiseenkova-Bell *et al.,* 2008; Shigematsu *et al.,* 2010), suggesting their potential importance in regulating the function of the channels.

In recent years, multiple structures of the N-terminal ankyrin repeat domain (ARD) have been deteremined for three thermosensitive TRPV channels, namely TRPV1, TRPV2 and TRPV4 (Jin *et al.,* 2006; Lishko *et al.,* 2007). The crystal structure of the ARD of another member of the family, TRPV6, has also been made available (Phelps *et al.,* 2008). TRPV6 is not known to be involved in temperature sensing and forms a pore that is highly selective for Ca^{2+} ions. The ARD structures established that the N-terminal domain contains only six, rather than the previously predicted three ankyrin repeats. In addition, in the case of TRPV1, the crystal structure was fortuitously obtained in the presence of ATP (Lishko *et al.,* 2007). The crystal structure thus revealed that ATP binds to the ARD of TRPV1 through hydrogen bonds between ATP and several charged residues on the first three ankyrin repeats (Lishko *et al.,* 2007). Subsequent competition assays with other nucleotides demonstrated that the phosphate groups and purine headgroup of ATP were important for this interaction. These assays also confirmed that, unlike other protein-ATP interactions, the binding of ATP to the ARD does not require cations like Mg^{2+} and Ca^{2+} (Lishko *et al.,* 1997). The crystal structures of the TRPV2 and TRPV4 ARDs were obtained in the absence of ATP (Jin *et al,* 2006; Landouré *et al.,* 2010). While the TRPV2 ARD adopted a very similar fold to that of TRPV1 in the presence of ATP, the structure of the TRPV4 ARD showed a large conformational change. Unlike in the TRPV1 ARD crystal structure, helix 3 in the TRPV4 ARD structure points away from the ligand-binding interface. Out of the three thermosensitive TRPVs with known ARD structures, only TRPV1 and TRPV4 are predicted to bind ATP. Thus, the crystal structures suggest a possibly important physiological function for the role of ATP in regulating the activities of thermosensitive TRPVs. Electrophysiological experiments indeed demonstrated that addition of ATP results in increased sensitization of TRPV1 to stimulus (Lishko *et al.,* 2007). In addition, it was shown that Ca^{2+}-CaM interacts with TRPV1 through the same interface on TRPV1 to which ATP binds and that binding of Ca^{2+}-CaM desensitizes TRPV1 to stimulus. These findings led us to investigate whether this mode of regulation by ATP and Ca^{2+}-CaM is conserved among TRPV proteins.

We have characterized the interactions of other members of the TRPV family with ATP and Ca^{2+}-CaM. In the case of TRPV6, which is not a thermosensitive channel, the ARD may adopt the same fold as in the other TRPV channels but it does not interact with either ATP or Ca^{2+}-CaM (Chapter 2.2) (Phelps *et al.,* 2008). On the other hand, like TRPV1, the ARDs of TRPV3 and TRPV4 both bind ATP and Ca^{2+}-CaM (Chapter 2.3) (Phelps *et al.,* 2010). While both channels show desensitization to stimuli in the presence of Ca^{2+}-CaM, TRPV3 shows decreased sensitization in the presence of ATP, very much unlike TRPV1 and TRPV4 (Chapter 2.3) (Phelps *et al.,* 2010). The obtained data, which were in agreement with previous observations with thermoTRPV channels, may explain the temperature thresholds of the different channels and provide clues as to why different temperatures elicit different bodily responses.

Chapter 2.2: Structural analyses of the ankyrin repeat domain of TRPV6 and related TRPV ion channels

Introduction

The transient receptor potential (TRP) proteins are a superfamily of cation channels with diverse functions. All TRP channels share a similar six-transmembrane-segment ion-transport domain flanked by intracellular N- and C-terminal domains, and are expected to function as tetramers. TRP channels are divided into seven subfamilies based on sequence similarity in their cytosolic domains (TRPA−ANKTM1, TRPC−canonical, TRPM−melastatin, TRPN−NOMP-C, TRPV−vanilloid receptor, and the more distantly related TRPML−mucolipin and TRPP−polycystin) (Nilius and Voets, 2005). The N termini of the TRPA, TRPC, TRPN and TRPV channels contain between 3 and 31 ankyrin repeats (Owsianik *et al.*, 2006), which are 33-amino acid residue motifs often involved in protein−protein interactions (Gorina *et al.,* 1996).

TRPV proteins have been identified in eukaryotes ranging from invertebrates to humans, and many are believed to play a role in sensing the environment (Clapham 2003; Montell 2003; Vriens *et al.*, 2004). The mammalian TRPV subfamily is composed of six members that contain a six-repeat ankyrin repeat domain (ARD) in their N-terminal cytosolic domain (Figure 2.2.1A) (Jin *et al.*, 2006; McCleverty *et al.*, 2006; Lishko *et al.*, 2007). The TRPV channels can be further broken down into two subfamilies: TRPV1−4, all of which are thermosensitive, non-selective cation channels expressed primarily in sensory neurons and keratinocytes, and TRPV5 and TRPV6, both of which are highly selective for Ca^{2+} and are expressed primarily in epithelial tissue (Montell *et al.,* 2002; Den Dekker *et al.*, 2003). TRPV6 is the primary transporter of calcium in the intestinal epithelium, while TRPV5 is predominant in the kidney (Hoenderop *et al.*, 2003). Both TRPV5 and TRPV6 (previously known as CaT2 and CaT1, respectively (Clapham *et al.*, 2005)) are inward rectifying channels, constitutively active at low Ca^{2+} concentration and physiologic membrane potentials, and more than 100 times more selective for Ca^{2+} than Na^{+} (Vennekens *et al.*, 2000).

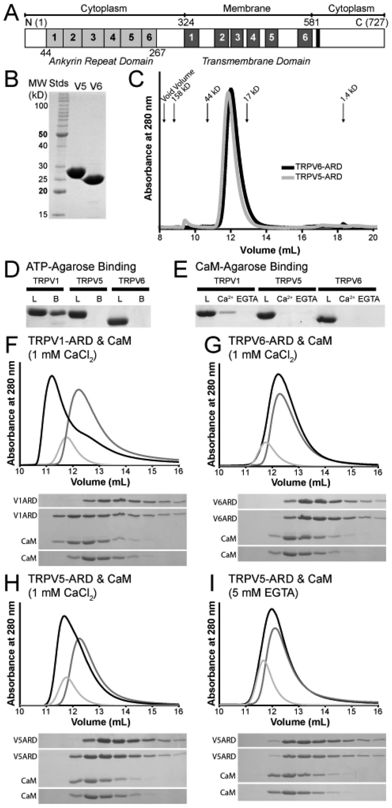


Figure 2.2.1: Biochemical properties of the TRPV5-ARD and TRPV6-ARD. A) Schematic representation of a TRPV protein primary structure and the location of each domain. Individual ankyrin repeats and predicted transmembrane segments are

Figure 2.2.1 (Continued) numbered and colored light gray and dark gray, respectively. The conserved TRP-box in the C terminus is colored black. The residue numbers correspond to those of mouse TRPV6. B) 15% SDS−PAGE analysis of the TRPV5-ARD and TRPV6-ARD purified from *E. coli*. C) Analytical Superdex 75 size exclusion chromatography of the TRPV5-ARD (gray) and TRPV6-ARD (black). Absorbance at 280 nm is plotted against elution volume. The void volume and elution volume of molecular weight standards are indicated. D) ATP-agarose pull-down assay of TRPV ARDs, with 15% SDS−PAGE analysis of the loaded (L) and ATP-agarose-bound (B) protein. E) CaM-agarose pull-down assay of TRPV ARDs, with 15% SDS−PAGE analysis of loaded protein (L) and protein bound in the presence (Ca^{2+}) and absence (EGTA) of calcium. TRPV1-ARD (*9*) is included as a positive control in panels D and E. F−I) Analysis of CaM/ARD complex formation by Superdex 75 size exclusion chromatography. Shown are representative traces from at least two experiments for TRPV1-ARD (F), TRPV6-ARD (G), and TRPV5-ARD with Ca^{2+} (H) or EGTA (I). Absorbance at 280 nm is plotted against elution volume with CaM alone in light gray, TRPV-ARD in gray, and CaM plus TRPV-ARD in black. Shifts in retention time were confirmed by 20% SDS−PAGE analysis of 0.5 ml fractions covering the elution volumes between 10 and 15 ml for each injection.

TRPV5 and TRPV6 share approximately 75% sequence similarity and are capable of associating with each other; heterotetramers show intermediate electrophysiologic properties, depending on the ratio of TRPV5 to TRPV6 (Hoenderop *et al.*, 2003). Homo- and heterotetramerization are dependent on residues in all three regions of the protein, both the N- and C-terminal cytosolic domains and the transmembrane domain (Hellwig *et al.,* 2005). Two ankyrin repeats of the TRPV6-ARD (repeats 3 and 5; residues 116−140 and 192−230, respectively) and ankyrin repeat 1 of TRPV5 (residues 64−76) have been implicated in channel assembly (Erler *et al.*, 2004; Chang *et al.*, 2004). C-terminal residues 596−601 are essential for TRPV5 channel assembly (Chang *et al.*, 2004). Furthermore, the C-terminal regions of both TRPV5 and TRPV6 are important for binding proteins involved in plasma membrane localization: the small GTPase Rab11a targets both TRPV5 and TRPV6 to the plasma membrane using residues 595−601 in TRPV5, and the corresponding residues (600−607) in TRPV6 (Van de Graaf *et al.*, 2006a); and residues 596−616 in TRPV5 (or 601−621 in TRPV6) are important for binding the PDZ domain-containing protein NHREF4, which may be involved in plasma membrane retention of the two channels (Van de Graaf *et al.*, 2006b). The C-terminal region also contains the TRP-domain, a short hydrophobic segment found in most TRP channels (Montell, 2001; Montell, 2005).

The activity of TRPV6, but not TRPV5, is inhibited by direct binding of regulator of G-protein signaling 2 (RGS2) to the TRPV6 N-terminal cytosolic domain (Meyer *et al.*, 2006; Meyer *et al.*, 2007). The Ca^{2+}-sensing protein calmodulin (CaM) also inhibits TRPV6 (*25*). CaM binding to and inhibition of TRPV6 requires both the N- and C-terminal cytosolic domains and the transmembrane domain (Nilius *et al.*, 2003). The activity of CaM on TRPV6 is opposed by protein kinase C (PKC), which is known to phosphorylate the C-terminal domain of TRPV6 (Niemeyer *et al.*, 2001).

In addition to its normal role as an epithelial calcium transporter, TRPV6 also plays a role in some types of prostate cancer (Wissenbach *et al.*, 2001; Fixemer *et al.*, 2003). The TRPV6-mediated increase in prostate cancer cell proliferation is mediated by stimulation of the Ca^{2+}-activated transcription factor NFAT, which inhibits apoptosis (Lehen'kyi *et al.*, 2007). This can be replicated by overexpression of TRPV6 in HEK-293 cells, which increases intracellular Ca^{2+} concentrations, resulting in increased cell proliferation (Schwarz *et al.*, 2006).

The TRPV6-ARD crystal structure was determined in order to better understand its role in channel assembly and regulation by other factors. Comparison with the crystal structures of the TRPV1 and TRPV2 ARDs and the sequences of other TRPV proteins leads to the identification of conserved structural features unique to the TRPV ARDs. We also show that the TRPV5 and TRPV6 ARDs are monomeric in solution, and unlike the TRPV1-ARD, they do not bind ATP or CaM to an appreciable extent. We discuss the implications of these structural and biochemical findings on the role of the ARD in the function of TRPV5 and TRPV6.

Materials and methods

Cloning, expression and purification

The human TRPV5 and mouse TRPV6 cDNAs were obtained from ATCC. The TRPV6-ARD (residues 42−266) and TRPV5-ARD (residues 11−267) were cloned between the *Nde*I and *Not*I restriction sites of pET21-C6H. The ARDs were expressed in *Escherichia coli* BL21(DE3) cells by induction with 0.4 mM isopropyl-β-D-thiogalactopyranoside

(IPTG) overnight at room temperature after the cells reached an OD_{600} of 0.4. Selenomethionine (Se-Met)-substituted TRPV6-ARD was expressed with feedback inhibition of methionine synthesis in M9 minimal medium supplemented with Se-Met under the same conditions. Cell pellets were resuspended in lysis buffer (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 20 mM imidazole) and 1 mM phenylmethylsulfonyl fluoride) supplemented with 0.1% Triton X-100, 0.2 mg/ml lysozyme, 50 μg/ml RNAse and 25 μg/ml DNAse, and lysed by sonication on ice. The cleared lysates were loaded onto Ni-NTA agarose (Qiagen) and eluted by a step gradient using lysis buffer containing 50, 100, 150 and 200 mM imidazole. Ten mM EDTA, pH 8.0, and 2.5 mM 2-mercaptoethanol were added to each fraction after elution. The ARDs were further purified on a Resource Q column (GE Healthcare) in 20 mM Tris-HCl, pH 8.0, 2 mM EDTA, 1 mM dithiothreitol (DTT) using a linear gradient of 0−1 M NaCl. Size exclusion chromatography on a Superdex 200 column (GE Healthcare) in 10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 5% (v/v) glycerol and 1 mM DTT was used as a final purification step before concentration to 30 mg/ml. The DTT concentration was raised to 10 mM in the anion exchange and size exclusion buffers during the purification of Se-Met-substituted TRPV6-ARD. Rat TRPV1-ARD and human CaM were purified using previously established protocols (Lishko *et al.*, 2007; Drum *et al.*, 2001).

Analytical size exclusion chromatography

All analyses were performed at 4°C. TRPV5- and TRPV6-ARD were diluted in Tris-buffered saline (TBS), pH 7.4, to 50 μM and injected on a Superdex 75 10/300 column (GE healthcare) in TBS, pH 7.4, with 1 mM DTT. Prior to injection the homogeneity of the samples was confirmed by 15% SDS−PAGE followed by Coomassie staining. Molecular weight standards (BioRad) were used to establish a standard curve to determine the apparent molecular weight of the eluted ARDs. The buffer used for CaM interaction assays was 20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM DTT with either 1 mM $CaCl_2$ or 5 mM EGTA. Samples containing 25 nmol of TRPV ARD, 25 nmol of CaM or both were incubated in the buffer for 45 min prior to injection.

ATP- and CaM-agarose pull-down assays

All binding assays were carried out at 4°C. For ATP-agarose pull-downs, 75 µl of a 50% slurry of agarose beads (11-atom spacer to ribose hydroxyls, Sigma) in 0.1 M Tris-HCl, pH 7.5, and 0.5 M NaCl was added to 12.5 µg of protein in 0.9 ml of binding buffer (10 mM Tris-HCl, pH 7.5, 50 mM NaCl, 1 mM DTT and 0.15% *n*-decyl-β-D-maltopyranoside). The samples were incubated for 1.5 h before washing three times with 0.9 ml of binding buffer. The samples were eluted by incubating 5 min at 95°C in 75 µl of 2x SDS sample buffer (100 mM Tris-HCl, pH 6.8, 4% SDS, 0.2% bromophenol blue, 20% glycerol and 200 mM DTT) and analyzed by 15% SDS−PAGE. For CaM-agarose pull-downs, 60 µg of protein was combined with 50 µl of a 50% CaM-agarose slurry (Sigma) in a final volume of 0.3 ml of binding buffer with either 2 mM $CaCl_2$ or 5 mM EGTA, pH 7.5, and incubated for 2 h before washing three times with 0.9 ml of the same buffer. Samples were eluted by incubating 5 min at 95°C in 50 µl of 2x SDS sample buffer and analyzed by 15% SDS−PAGE.

Crystallization of TRPV6-ARD

TRPV6-ARD crystals were grown by the hanging-drop vapor diffusion method at room temperature with a 1:1 protein to reservoir solution ratio, with 0.1 M NaHEPES, pH 7.5, 0.15 M K/Na tartrate and 5% glycerol in the reservoir. Se-Met TRPV6-ARD crystals were grown from the same solution plus 10 mM DTT. Crystals were cryoprotected in 0.1 M NaHEPES, pH 7.5, 0.2 M K/Na tartrate and 30% glycerol and flash frozen in liquid nitrogen.

Data collection, structure determination and analysis

X-ray data from native and Se-Met substituted crystals were collected at 100 K using an ADSC Q315 detector at the Advanced Photon Source ID24 beamline. Data were processed in HKL2000 (Otwinnoski *et al*., 1997), and data statistics are listed in Table 2.2.1. The TRPV6-ARD structure was determined to 1.7 Å by single isomorphous replacement with anomalous scattering (SIRAS). Initial phasing, solvent flattening and model building were carried out using autoSHARP (de La Fortelle *et al.*,1997; Vonrhein *et al*., 2006) and ARP/wARP (Perrakis *et al*., 1997; Perrakis *et al*., 1999; Morris *et al*., 2003). Model building was performed in COOT (Emsley and Cowtan, 2004), and refinement with Translation/Libration/Screw (TLS) was carried out in REFMAC5 (Murshodow *et al.,* 1997). Due to the lack of completeness in the low-resolution data, only reflections between 8.0 and 1.7 Å were used in the final refinement. Final refinement statistics are listed in Table 2.2.1. The coordinates have been deposited in the Protein Data Bank with the entry code 2RFA. Figures were generated with PyMOL (*Delano Scientific LLC*).

	Native	Se-Met
Data Collection		
space group	$P2_12_12_1$	
wavelength (Å)	0.97921	
cell dimens (*a*, *b*, *c*; Å)	30.76, 63.05, 116.14	31.18, 62.51, 116.40
resolution (Å)	30.0−1.7 (1.76−1.7)	30.0−2.4 (2.49−2.4)
R_{sym}	0.044 (0.445)	0.094 (0.526)
$I/\sigma(I)$	18.0 (3.4)	11.8 (3.6)
completeness (%)	98.8 (95.8)	100.0 (100.0)
redundancy	7.5 (6.4)	5.9 (5.9)
Refinement		
resolution (Å)	8.0−1.7	
number of reflections	25,172	
R_{work}/R_{free}	0.168/0.207	
molecules/asym unit	1	
residues in model	44−265	
number of atoms		
protein	1750	
water	227	
B-factors (Å^2)		
protein	35.4	
water	56.9	
rms deviations		
bond lengths (Å)	0.014	
bond angles (deg)	1.37	

Table 2.2.1: Data collection and refinement statistics[a]

Table 2.2.1 (Continued) [a]Values from the highest resolution shell are in parentheses.

Results

The TRVP5 and TRPV6 ARDs are monomeric in solution.

The ankyrin repeats of TRPV5 and TRPV6 play an important role in the tetramerization and assembly of their respective full-length channels (Erler *et al.*, 2004; Chang *et al.*, 2004). Therefore, we investigated whether the isolated ARDs could themselves self-assemble into tetramers. The ARDs of TRPV5 (residues 11−267) and TRPV6 (42−266) were overexpressed and purified from *E. coli* (Figure 2.2.1A and B) and analyzed by size exclusion chromatography to determine their oligomerization state. Both ARDs eluted at volumes consistent with a monomer (Figure 2.2.1C). The molecular weights calculated from the amino acid compositions are 29.9 and 26.1 kDa for TRPV5-ARD and TRPV6-ARD, respectively, and the apparent molecular weights determined from a standard curve from size exclusion were 29.7 and 27.4 kDa, respectively. Furthermore, the TRPV6-ARD behaved as a monomer at all concentrations tested, up to 0.4 mM (data not shown).

The isolated TRPV5 and TRPV6 ARDs do not bind ATP

TRPV1-ARD is known to bind to both ATP and Ca^{2+}-CaM (Lishko *et al.*, 2007), and it has been suggested that the ARD of TRPV6 binds CaM as well (Lamber *et al.*, 2004). We therefore tested whether the TRPV5 and TRPV6 ARDs could bind ATP or CaM in pull-down assays. Under conditions in which the TRPV1-ARD was pulled down efficiently by ATP-agarose, neither the TRPV5-ARD nor TRPV6-ARD interacted with ATP-agarose (Figure 2.2.1D). This finding was not unexpected, as two of the positively charged residues in the second ankyrin repeat of TRPV1 that contact the triphosphate moiety of ATP, K155 and K160, are substituted with negative or hydrophobic residues in TRPV5 and TRPV6 (E80 and I85 in both).

The ARD of TRPV5, but not TRPV6, binds CaM

We also identified conditions under which the TRPV1-ARD bound efficiently to CaM-agarose in a Ca^{2+}-dependent manner, as expected from previous size exclusion chromatography experiments showing that the two proteins form a 1:1 complex (Lishko *et al.*, 2007). Under those same conditions, the TRPV5-ARD and TRPV6-ARD were not pulled down by CaM-agarose in either the presence or absence of Ca^{2+} (Figure 2.2.1E).

We further investigated the interaction of CaM with TRPV5-ARD and TRPV6-ARD by size exclusion chromatography. As in the pull-down assays, TRPV1-ARD was used as a positive control (Figure 2.2.1F). Like in the CaM-agarose pull-down assays, no binding was observed between TRPV6-ARD and CaM in the presence of calcium (Figure 2.2.1G). In light of previously published serial deletion analysis data showing an interaction between CaM and ankyrin repeat 2 of TRPV6 (Lambers *et al.,* 2004), it is surprising that TRPV6-ARD did not readily bind to CaM-agarose. However, analysis of the structure suggests an explanation for this behavior (see the Discussion below).

A small but significant shift in the elution volume of TRPV5-ARD was observed in the presence of CaM, indicating that the two interact (Figure 2.2.1H). Like TRPV1-ARD, the interaction of TRPV5-ARD with CaM could be blocked by removal of calcium with 5 mM EGTA and is therefore calcium-dependent (Figure 2.2.1I). Like TRPV6, interactions between TRPV5 and CaM have been previously reported, but unlike TRPV6, Ca^{2+}-CaM had no effect on TRPV5 activity (Lambers *et al.,* 2004).

Overall structure of TRPV6-ARD

The crystal structure of TRPV6-ARD was determined using initial phases from single isomorphous replacement with anomalous scattering (SIRAS) and refined to a final R_{work}/R_{free} of 0.167/0.209 at a resolution of 1.7 Å (Table 1). In agreement with the monomeric state of the TRPV6-ARD observed by size exclusion chromatography, there was one monomer per asymmetric unit, with crystal packing incompatible with a tetrameric state.

Like the ARDs of other mammalian TRPV proteins (Jin *et al.*, 2006; McCleverty *et al.*, 2006; Lishko *et al.*, 2007), the TRPV6-ARD is composed of six ankyrin repeats (Figure 2.2.2). These repeats have the characteristic antiparallel inner and outer α-helices, with the helical layers linked together by finger loops (Figure 2.2.2). As in other ankyrin repeat-containing proteins, the hydrophobic core of the TRPV6-ARD is composed of conserved ankyrin-repeat consensus residues, while variable residues, which most likely specify binding partners, are found on the exposed faces of the repeat helices and at the tip of the finger loops. In addition, several residues conserved among TRPV ARDs line the base of the finger loops and form the short helices between repeat ANK2 and Finger 2 and ANK5 and Finger 5 (Figure 2.2.2B). As in the TRPV1 and TRPV2 ARDs, the TRPV6-ARD structure exhibits a large twist between ANK4 and ANK5. The molecular basis of this conserved twist is examined in detail below.

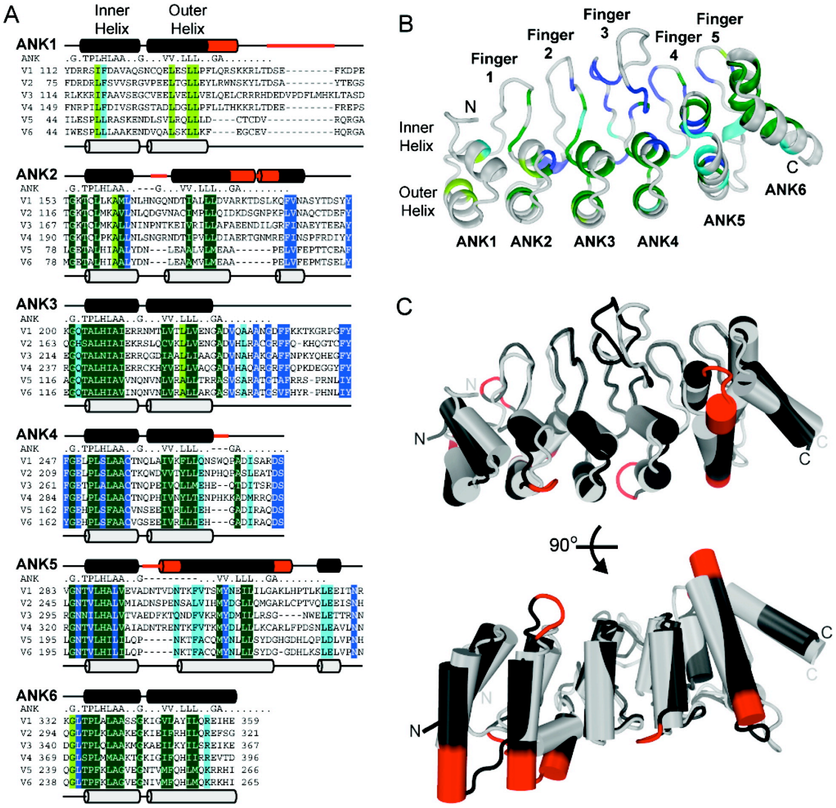


Figure 2.2.2: Structure of the TRPV6 ARD. A) Alignment of the ARDs of mammalian TRPV channels separated into individual repeats (ANK1−6). Residues that conform to the ankyrin repeat consensus (indicated above each repeat) in all species (from worms to humans; see Figure 2.2.5 are colored dark green, and ankyrin repeat residues conserved in just mammalian TRPVs are colored light green. Residues that are conserved in TRPV ARDs and either are not part of or deviate from the ankyrin repeat consensus are colored dark blue for residues conserved in all species and light blue for residues conserved in mammalian TRPV ARDs only. For ARDs for which a structure is yet to be determined (TRPV3, TRPV4 and TRPV5) the human sequences are used. For TRPV1, TRPV2 and TRPV6, the sequence of the crystallized species is used (rat TRPV1 and TRPV2, and mouse TRPV6). The secondary structure of TRPV1-ARD (dark gray; is

Figure 2.2.2 (Continued) indicated above each repeat, and the secondary structure of TRPV6-ARD (light gray) is shown below. Insertions in TRPV1−4 compared to TRPV5 and TRPV6 are shown in red. B) Sequence similarities mapped onto the overall structure of TRPV6-ARD. The backbone is colored according to the sequence similarity in the alignment shown in A. C) Location of the sequence insertions in TRPV1−4 relative to TRPV5 and TRPV6. The structures of the TRPV1 and TRPV6 ARDs were superimposed and colored according to the secondary structure in A. α-Helices are shown as cylinders. The view at the bottom corresponds to a 90° rotation around the horizontal axis to highlight the fact that most of the insertions map to the ends of outer helices.

Comparison of the TRPV6-ARD structure to that of TRPV1-ARD and TRPV2-ARD

As stated in the introduction, the mammalian TRPV proteins can be subdivided into two subgroups, the TRPV1−4 channels involved in thermosensation and the TRPV5 and TRPV6 channels, critical components of calcium homeostasis. This subdivision is also reflected in phylogenetic analyses (Liedtke and Kim, 2005) and in sequence similarity across the entire sequence, including within the ARDs (Phelps *et al.,* 2007). Correspondingly, although the overall structure of the TRPV6-ARD is similar to the TRPV1 and TRPV2 ARDs, there are notable differences. There are six sequence insertions within the ARDs of TRPV1−4, compared to TRPV5 and TRPV6. A comparison of the TRPV6-ARD structure to those of TRPV1-ARD (Lishko *et al.,* 2007) and TRPV2-ARD (Jin *et al.,* 2006; McCleverty *et al.*, 2006) reveals that four of the insertions result in extensions of the C-terminal end of outer helices and the other two extend the turn between helices in repeats ANK2 and ANK5 (Figures 2.2.2A and 2.2.2C).

None of the insertions perturb the orientation or length of the finger loops, and the tip of the fingers are remarkably similar in the structure of the TRPV1-ARD and the TRPV6-ARD (Figure 2.2.2C). One notable exception is the tip of the longest finger, Finger 3. The finger loops in TRPV6 and other TRPV proteins are longer than those in canonical ankyrin repeats with Finger 3 being the longest. Finger 3 is the only loop that adopts strikingly different conformations in the known TRPV ARD structures (Jin *et al.,* 2006; McCleverty *et al.*, 2006; Lishko *et al.,* 2007) (Figure 2.2.2C). The conformational variability is restricted to the tip of the finger, in the region corresponding to residues 153−159 in TRPV6. Notably, this region of Finger 3 is also the most variable in length in TRPV ARDs, varying from seven residues in TRPV2, TRPV5 and TRPV6 to 43 residues in the *Caenorhabditis elegans* TRPV OCR4 (Figure 2.2.5, Appendix).

Molecular basis for the unusual twist between ankyrin repeats 4 and 5 in the TRPV ARDs

As was previously observed in the TRPV1-ARD and TRPV2-ARD structures (Jin *et al.,* 2006; McCleverty *et al.*, 2006; Lishko *et al.,* 2007), there is a pronounced twist between TRPV6 repeats ANK1−4 and ANK5−6. This twist occurs in a region in which the ARD of the TRPVs have conserved substitutions deviating from the ankyrin repeat consensus (blue residues in Figures 2.2.2A and B). It had been suggested that the twist between repeats in the TRPV2-ARD were a result of the extended outer helices; however, TRPV6 shows the same twist despite having shorter outer helices (Figures 2.2.2A and C). Canonical ankyrin repeats, like repeats ANK2 through ANK4 in TRPV6, have a twist angle of approximately 6° between repeats and a distance of 12 Å between the final Cα of adjacent outer helices (see Figure 2.2.6, Appendix, for measurement details). In TRPV6, the twist angle between repeats ANK4 and ANK5 is 17° with a corresponding distance of 13 Å and the twist between repeats ANK5 and ANK6 is 22° with a distance of 17 Å. These increased twist angles are similar in the TRPV1- and TRPV2-ARDs (data not shown).

The twist between repeats ANK4 and ANK5 results from displacement of inner helix 5 from the regular packing of hydrophobic side chains observed between canonical ankyrin repeats. TRPV-conserved residues that deviate from ankyrin consensus in inner helix of repeat 5, V199 and L203 in TRPV6, replace the proline and alanine normally found in those positions (5 and 9, respectively) (Figure 2.2.3) (Mosavi *et al.,* 2002; Binz *et al.,* 2003). Furthermore, I202 on inner helix 5 is rotated out of the hydrophobic core between repeats ANK4 and ANK5 and instead packs against the Y161 and Y162 side chains at the base of Finger 3 and C172 on the outer face of inner helix 4 (Figure 2.2.3B).

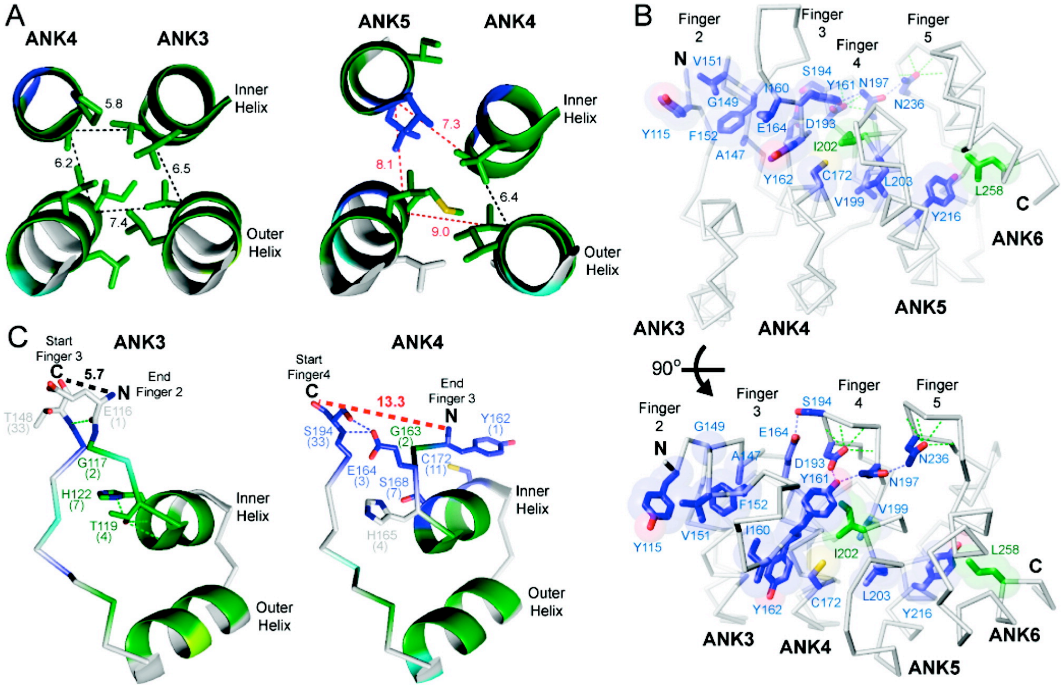


Figure 2.2.3: Conserved TRPV residues induce the unusually pronounced twist in the ARD and define the orientation of the finger loops. In all panels, green dashed lines indicate hydrogen bonds typically observed in canonical ankyrin repeats and blue dashed lines represent hydrogen bonds between TRPV conserved residues. Distances normally observed in ankyrin repeats are shown as dashed black lines, and significantly different distances in TRPV6-ARD are shown in red. All distances are in Ångstroms. All side chains are colored according to Figures 2.2.2A and B. Key residues are labeled, and numbers in parentheses correspond to the position of the residue within the ankyrin repeat consensus (see Figure 2.2.2A). A) The canonical ankyrin repeat packing pattern can be observed between ANK3 and ANK4 (left), while V199 and L203 on the inner helix of ANK5 break the regular stacking of the helices (right). In panel A, the regions illustrated are viewed approximately from the back of the orientation shown in Figure 2.2.3B. B) A network of hydrophobic packing and hydrogen bonds between TRPV-

Figure 2.2.3 (Continued) conserved residues stabilize the base of Finger 3, the displaced inner helix of ANK5 and the displaced outer helix of ANK6, generating the pronounced twist between repeats 4 and 5 observed in all TRPV ARDs. Only ankyrin repeats 3 through 6 are shown as a Cα trace. Residues involved in the conserved hydrophobic network are shown as sticks and transparent spheres (with coloring conforming to Figure 2.2.2A). C) TRPV substitutions disrupt the parallel orientation of the start and end of consecutive fingers. ANK3 (left) adopts a canonical ankyrin repeat fold with the threonine and histidine at positions 4 and 7, respectively, setting the orientation of the finger with respect to the helices. In contrast, the end of Finger 3 adopts a novel configuration as it lies against the inner helix of ANK4 (right).

Finally, a large twist within repeat ANK6 prevents a steric clash between the side chains of Y216 from repeat ANK5 and L258 from repeat ANK6 (Figure 2.2.3B). This results in a twist angle between the outer and inner helices of about 34°, approximately 15° more than seen in most ankyrin repeats. Meanwhile, Finger 5 and inner helix of repeat 6 are held in place by canonical ankyrin interactions.

Conserved TRPV residues determine the orientation of Finger 3 base

In all known TRPV ARD structures, including in TRPV6-ARD, Finger 3 breaks from the regular β-strand-like packing at the start and end of consecutive fingers. The position of an ankyrin repeat finger is usually fixed relative to the following inner helix by interactions between conserved threonine and histidine residues at positions 4 and 7, respectively, in the ankyrin repeat consensus, and the amide nitrogens of the first turn of the inner helix (Figure 2.2.3C left) (Mosavi *et al.,* 2004). In repeat ANK4 of TRPV

ARDs, position 4 is variable and position 7 is a conserved serine (Figure 2.2.2A). As a result, without the canonical constraint on backbone orientation, Finger 3 folds back along the face of inner helix 4 and the first ANK4 residue, Y162, packs against C172 (Figure 2.2.3C right). Instead of the canonical backbone β-strand hydrogen bonds observed between the end of Finger 2 and the start of Finger 3, the E164 side chain from Finger 3 bridges the gap and hydrogen bonds with S194 at the start of Finger 4 (Figure 2.2.3B and C). Y161 from Finger 3 also hydrogen bonds to D193 and N197 of Finger 4, further buttressing both loops in relation to each other. The base of Finger 3 is further stabilized by a conserved network of hydrophobic residues from repeats two through five that includes Y161, Y162 and C172 (Figure 2.2.3B).

Discussion

Biochemical and crystallographic analyses of the ARD of TRPV6 reveal that the six-repeat domain is monomeric. Like the TRPV1 and TRPV2 ARDs, the TRPV6-ARD shows a pronounced twist between repeats ANK1−4 and ANK5−6. The twist in the ARD is induced by conserved TRPV substitutions in the ankyrin repeat consensus and stabilized by conserved hydrophobic side chains at the bases of the unusually long Fingers 2 and 3. These hydrophobic residues also anchor the base of the most flexible and variable region in TRPV ARDs, the tip of Finger 3. The impact of these structural features on function and regulation of TRPV5 and TRPV6 is discussed below.

Role of the ARD in the assembly of tetrameric TRPV5 and TRPV6 channels

Previous reports have suggested that the ARDs of TRPV5 and TRPV6 are directly involved in tetramerization of the channels through interaction between the ARDs (Erler *et al.,* 2004; Chang *et al.,* 2004). The locations are mapped onto the TRPV6 ARD structure in Figure 2.2.4. Our size exclusion data clearly indicate that the TRPV5-ARD and TRPV6-ARD are monomeric in solution. Furthermore, the packing and symmetry in the TRPV6-ARD crystals are incompatible with tetrameric assembly of the ARD around a 4-fold symmetry axis. Residues that had previously been identified as critical for tetramerization are also at important conserved ankyrin repeat positions. Residues that are critical for ankyrin repeat stability, particularly the paired leucines in the outer helices of repeats 1 and 3 (Figure 2.2.4), are lost by either deletion of ANK1 in TRPV5 or mutation of ANK3 in TRPV6 in the tetramerization-deficient mutants (Erler *et al.,* 2004; Chang *et al.,* 2004). At positions in ankyrin repeats that are normally conserved, mutations to nonconsensus residues typically decrease the stability of ARDs, while conversely for ARDs that have nonconsensus residues at conserved positions, mutations back to ankyrin repeat consensus often increase ARD stability (Binz *et al.*, 2003; Zweifel *et al.*, 2003; Ferreiro *et al.*, 2007). Furthermore, the folding of ankyrin repeat structures is nucleated at specific repeats and propagated through the rest of the domain (Lowe *et al.*, 2007). Removal of ANK3 in TRPV6 (Erler *et al.*, 2004) or the introduction of destabilizing mutations in TRPV5 (Chang *et al.*, 2004) therefore likely disrupts the overall fold of the entire ARD, thereby preventing the assembly of tetrameric channels. Erler and colleagues also observed that two complementary segments of the TRPV6-ARD, ANK3−4 (residues 116−191) and ANK5−6 (residues 192−328), could interact in a two-hybrid assay (Erler *et al.*, 2004). This finding could be explained by the two segments stacking to form a single ARD in *trans*. Furthermore, interactions between the

N and C termini of both TRPV5 and TRPV6 have been observed (Erler *et al.,* 2004; Chang *et al.,* 2004), and this may be the reason that interactions were also observed between the TRPV6 N terminus and full-length protein, but not the N terminus with itself in bacterial two-hybrid assays (Erler *et al.*, 2004). It is also possible that TRPV5 and TRPV6 assembly is assisted by additional cellular factors that require the ARD and are unable to bind an ARD destabilized by mutation or partial deletion. In summary, while previous studies clearly demonstrate that the integrity of the ARD is important in channel assembly, our data indicate that it is not through self-tetramerization of the ARD. Rather, the TRPV ARDs likely regulate channel assembly through interactions with other factors or other regions within the channel.

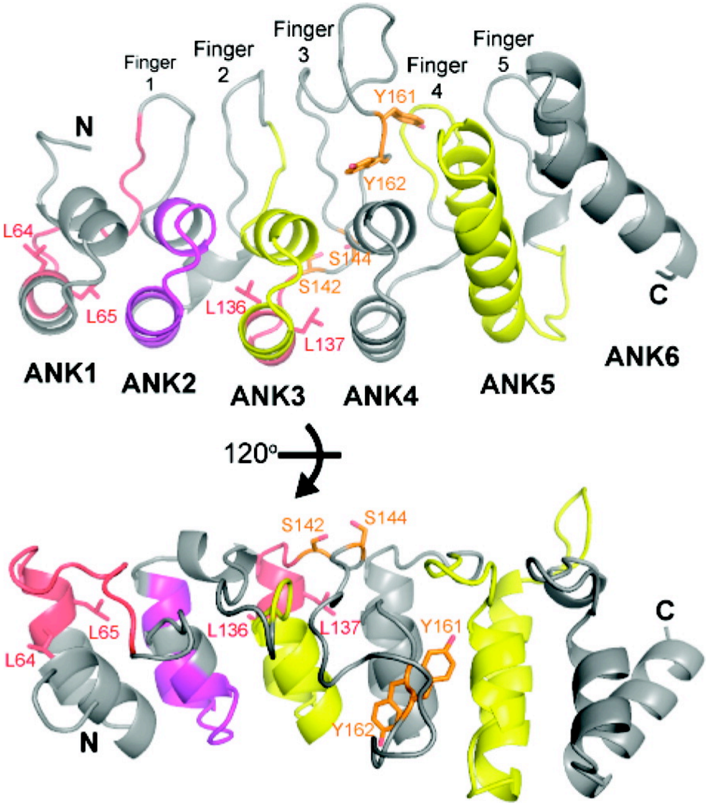


Figure 2.2.4: Previously identified functions of the TRPV6-ARD mapped onto the

Figure 2.2.4 (Continued) structure. Sections identified as critical for channel assembly/tetramerization (residues 64−76 in TRPV5 (Chang *et al.*, 2004) and residues 136−140 in TRPV6 (Erler *et al.*, 2004)) are colored red, and the regions of ankyrin repeats 3 and 5 assigned a role in channel assembly (residues 116−163 and 192−230 (Erler *et al.*, 2004)) are colored yellow. The region previously implicated in CaM binding (residues 85−100 (Lambers *et al.,* 2004)) is colored purple. Side chains whose phosphorylation may regulate TRPV6 activity are colored orange.

Regulation of TRPV5 and TRPV6 by CaM

Although in a previous study no effect of CaM on TRPV5 activity was observed (Lambers *et al.,* 2004), we observed an interaction between TRPV5-ARD and Ca^{2+}-CaM in the size exclusion assay, but not the more stringent CaM-agarose pull-down assay, indicating that the interaction is weaker than that observed between TRPV1-ARD and CaM. It should also be noted that the TRPV5-ARD construct used contained an additional 30 residues at its N terminus compared to the TRPV6-ARD construct used in this study. The TRPV5 construct was extended because the construct to most homologous to TRPV6-ARD could not be expressed in a soluble form. It is therefore possible that these extra residues contribute to the differences observed between TRPV5 and TRPV6 in the CaM-binding assays. The *in vivo* relevance of the weak interaction of CaM and TRPV5-ARD remains to be determined.

CaM is known to regulate TRPV6, and strong evidence exists for an interaction between CaM and the C terminus of TRPV6 (Niemeyer *et al.*, 2001*;* Lambers *et al.,* 2004).

Others have also reported that CaM interacts with the TRPV6-ARD, in the same region important for CaM binding to TRPV1 (Niemeyer *et al.*, 2001; Lishko *et al.*, 2007; Lambers *et al.*, 2004) (Figure 2.2.4). However, in our assays we do not observe any binding of CaM to TRPV6-ARD under conditions in which CaM does bind the TRPV1 and TRPV5 ARDs (Figure 2.2.1E−I). Like residues previously thought to be involved in tetramerization, the TRPV6-ARD residues implicated in CaM binding (residues 93−102 (Lambers *et al.*, 2004)) form part of the conserved ankyrin repeat core (Figure 2.2.4) and are unlikely to be available for interactions with other proteins. Recent cell-based studies, using a combination of electrophysiology and confocal microscopy, have also shown that the TRPV6 N terminus does not play a significant role in its inhibition by CaM (Derler *et al.*, 2006). Although both TRPV1 and TRPV6 are inhibited by CaM, the failure of the TRPV6-ARD to bind CaM (or ATP, a competitor for the CaM binding site and sensitizer in TRPV1) indicates that CaM regulates different members of the TRPV family through distinct mechanisms.

Regulation of TRPV6 by phosphorylation of ARD residues

It is possible that post-translational modifications are necessary for tetramerization of the TRPV6-ARD or for its association with CaM, and a number of confirmed and putative phosphorylation sites have been identified within the TRPV6-ARD. Phosphorylation of rat TRPV6 by a Src-like kinase leads to an increased Ca^{2+} influx through the channel, and both Y161 and Y162 were identified as putative phosphorylation sites through mutagenesis (Sternfeld *et al.*, 2007). The PTP1B phosphatase also interacts with the first 191 residues of TRPV6 and inhibits Ca^{2+} influx, likely through dephosphorylation of these tyrosine residues (Sternfeld *et al.*, 2007). Mutation of both tyrosines was necessary to prevent up-regulation of the channel by phosphorylation. In the TRPV6-ARD structure, the hydroxyl group of Y162 is exposed at the protein surface, while Y161 is buried and interacts with side chains from Finger 4 (Figures 2.2.3 and 2.2.4). If the observed tyrosine phosphorylation plays a significant role in regulating TRPV6 function *in vivo*, it is unlikely that Y162 is an important physiologic substrate of Src-like kinases and PTP1B, as it is not conserved in all TRPV6 proteins: it is a tyrosine in crayfish, zebrafish, mouse and rat, but a phenylalanine in human, chimpanzee, chicken and frog (Figure 2.2.5). In the TRPV6-ARD structure, Y161 is buried and inaccessible to kinases, and therefore also seems an unlikely target for phosphorylation. As mentioned above, mutations away from ankyrin repeat consensus destabilize individual ankyrin repeats (Zweifel *et al.*, 2003*;* Ferreiro *et al.*, 2007). It is therefore possible that the presence of residues that deviate from the ankyrin repeat consensus in TRPVs makes this region dynamic enough to allow kinase access to Y161. The Y161 side chain has slightly above average *B*-factors for the structure (41.6 $Å^2$ vs 35.4 $Å^2$) and is positioned where Finger 3 transitions from the stable region of conserved TRPV residues to the variable and flexible fingertip (Figure 2.2.7, Appendix). However, making Y161 accessible to a kinase still requires a rather large conformational change.

Two other putative phosphorylation sites have previously been identified in TRPV6 by sequence analysis (Den Dekker *et al.*, 2003). S142 and S144 are putative targets for CaM dependent kinase II (CaMKII) and protein kinase C (PKC), respectively. Both are surface-exposed at the base of Finger 3, and phosphorylation at one of the serines would disrupt recognition of the target site by the other kinase (Figure 2.2.4). It is noteworthy that neither serine is on the concave surface of the ARD that is typically involved in protein−ligand interactions in ankyrin repeat-containing proteins (Mosavi *et al.,* 2004). Still, the possibility that these two putative phosphorylation sites play a role in the balance between PKC activity and increased intracellular Ca^{2+}/CaM-binding in regulation of TRPV6 requires further investigation.

Conclusion

Comparison of the ARD structures from three TRPV proteins reveals conserved residues that produce the twist and stabilize the long fingers of these ARDs. Protein sequence analysis shows that the TRPV ARD "scaffold" is a conserved feature from worms to humans. These conserved residues form a continuous network that stabilizes both the twists and turns of the fingers and helices. The reason for the conservation of this twisted ankyrin repeat structure and stabilization of the base of the elongated Finger 3 will require the elucidation of the structures of TRPV ARDs with some of their interacting partners.

TRPV6 and the closely related TRPV5 channels are critical regulators of calcium homeostasis. Their activity, localization and assembly are all regulated in part through their N-terminal ARD. The crystal structure of the TRPV6 ARD reveals the location and environment of amino acid residues previously identified as mediators of these activities. Many of these residues actually form the conserved core of the TRPV6-ARD and provide the scaffolding that stabilizes the domain, shaping the surface accessible to interacting regulatory partners. The interactions that regulate TRPV6 are therefore likely to come from other residues, not yet identified, that are positioned by this scaffolding. The TRPV6-ARD structure provides a basis for further experiments such as mutagenesis of surface residues to disrupt potential ligand interfaces.

Chapter 2.3: Differential regulation of TRPV1, TRPV3, and TRPV4 sensitivity through a conserved binding site on the ankyrin repeat domain

Introduction

Transient receptor potential channels (TRP), including the six vanilloid (TRPV) channels in warm-blooded vertebrates, have many physiological functions in neuronal and non-neuronal cells (Venkatachalam *et al.*, 2007). TRPV5 and TRPV6 are calcium channels in the gut and kidney important for Ca^{2+} homeostasis (Nijenhuis *et al.*, 2005), whereas TRPV1–V4 are non-selective cation channels that contribute to temperature sensation (Caterina *et al.*, 2007). TRPV1 and TRPV2 activate at noxious temperatures above 42°C and 52°C, respectively, whereas TRPV3 and TRPV4 activate at warm temperatures ~33–39 and 25–34°C, respectively.

Thermosensitive TRPVs are polymodal channels activated by physical stimuli (e.g., temperature) and chemical agonists. For instance, capsaicin and low extracellular pH activate TRPV1 (Tominaga and Caterina, 2004), thymol, carvacrol and eugenol activate TRPV3 (Xu *et al.*, 2003), and extracellular hypotonicity, phorbol esters, and arachidonic acid metabolites activate TRPV4 (Liedtke *et al.*, 2005; Strotmann *et al.*, 2000; Watanabe *et al.*, 2002; Watanabe *et al.*, 2003). 2-Aminoethyl diphenylborinate (2-APB) is promiscuous and activates TRPV1, TRPV2, and TRPV3 (Hu *et al.*, 2004).

Remaining questions include whether TRPV channels have maintained common regulatory mechanisms. Thermosensitive TRPV channels are modulated intracellularly by Ca^{2+}, calmodulin (CaM), and phosphoinositides (Lee and Caterina 2005; Nilius *et al.*, 2008; Zhu 2005). TRPV1 desensitization depends on intracellular Ca^{2+} and CaM (Koplas *et al.*, 1997; Lishko *et al.*, 2007). Similarly, TRPV4 is first potentiated and then inactivated by intracellular Ca^{2+}, again likely through CaM (Strotmann *et al.*, 2003). Like TRPV1, TRPV4 desensitizes after repeated or prolonged stimulations (Güler *et al.*, 2002). In contrast, TRPV3 currents increase with repeated stimulation (Peier *et al.*, 2002; Xu *et al.*, 2002; Smith *et al.*, 2002), and while TRPV3 sensitivity also depends on Ca^{2+} and CaM, the effects differ from TRPV1 and TRPV4 (Xiao *et al.*, 2002). The nature of these differences in homologous temperature-sensitive TRPVs has yet to be determined.

TRPVs have a channel domain homologous to Shaker K^+ channels and cytosolic N- and C-terminal domains, including a conserved N-terminal ankyrin repeat domain (ARD) (Gaudet, 2008). TRPV1-, TRPV2-, and TRPV6-ARD structures have been reported (Lishko *et al.*, 2007; Jin *et al.*, 2006; McCleverty *et al.*, 2006; Phelps *et al.*, 2008). The crystal structure of TRPV1-ARD revealed a bound ATP molecule, and it was shown that ATP and Ca^{2+}-CaM compete for a common binding site on TRPV1-ARD (Lishko *et al.*, 2007). Intracellular ATP sensitizes TRPV1, while both Ca^{2+}-CaM and its binding site on the ARD are necessary to inactivate TRPV1 (Lishko *et al.*, 2007).

We investigated whether the modulatory binding site found on TRPV1-ARD exists in other TRPV channels. We demonstrate that TRPV3- and TRPV4-ARD also bind ATP and Ca^{2+}-CaM. Similar to TRPV1, TRPV4 is sensitized by intracellular ATP and a binding site mutation eliminates this sensitization. In contrast, intracellular ATP prevents TRPV3 sensitization to 2-APB, and binding site mutations confirm a role for the ARD in regulating TRPV3 sensitivity. Moreover, the ARD is key to the previously reported sensitivity of TRPV3 to intracellular Ca^{2+} and CaM (Xiao *et al.*, 2008). Potential physiological roles of this multiligand binding site conserved on several thermosensitive TRPV channels include setting channel responsiveness to stimuli and adaptation to the metabolic state.

Material and methods

Cloning of expression vectors

cDNA fragments encoding ARDs (human TRPV3-ARD residues 115–367 and chicken TRPV4-ARD residues 132–383) and full-length protein (human TRPV3 and chicken TRPV4) were cloned into the *Nde*I and *Not*I sites of pET21-C6H (Jin et al., 2006) and pFastBac-CFLAG (Lishko et al., 2007) vectors, respectively. Baculovirus stocks were generated and used to infect Sf21 cells as described in the Bac-to-Bac manual (Invitrogen). Full-length TRPVs in pcDNA3 were provided by Michael Caterina (Johns Hopkins School of Medicine; rat TRPV2), David Clapham (Harvard Medical School; human TRPV3) and Stefan Heller (Stanford University; chicken TRPV4). All mutants were generated by mutagenesis, and all clones were verified by DNA sequencing.

Expression and purification of TRPV ARDs

The ARDs were expressed in *Escherichia coli* BL21(DE3) by induction with 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG) overnight at room temperature after the cells reached A_{600} = 0.6. Cells were resuspended in lysis buffer (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 20 mM imidazole, and 1 mM phenylmethylsulfonyl fluoride) with 0.1% Triton X-100, 0.2 mg/ml lysozyme, 50 μg/ml RNase A, and 25 μg/ml DNase I and lysed by sonication. The cleared lysate was loaded onto nickel-nitrilotriacetic acid (Qiagen) and eluted by a step gradient containing 50, 100, 150, and 200 mM imidazole, pH 8, in lysis buffer. Ten mM EDTA, pH 8.0, and 1 mM dithiothreitol (DTT) were added after elution. The fractions containing TRPV3-ARD or TRPV4-ARD were pooled and further purified on Q or SP Sepharose FF (GE Healthcare), respectively, in 20 mM Tris, pH 8.0, 5 mM DTT using a linear gradient of 0–0.4 M NaCl. Size exclusion chromatography on a Superdex 75 column (GE Healthcare) in 10 mM Tris-HCl, pH 8.0, 200 mM NaCl, and 1 mM DTT was used for further purification of TRPV3-ARDs, whereas the TRPV4-ARDs were dialyzed in 20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 10% glycerol, and 1 mM DTT. All proteins were concentrated to >7 mg/ml in a Vivaspin centrifugal filter (10,000 molecular weight cut off; Sartorius AG, Goettingen, Germany), flash frozen, and stored at −80°C. TRPV1-ARD, TRPV2-ARD, TRPV5-ARD, and TRPV6-ARD were purified as described previously (Lishko *et al.*, 2007; Jin *et al.*, 2006; Phelps *et al.*, 2008).

ATP- and CaM-agarose pull-down assays

All assays were carried out at 4°C as described previously (Phelps *et al.*, 2008). The ATP-agarose assays were performed in the absence of divalent ions except otherwise noted, in binding buffer (10 mM Tris-HCl, pH 7.5, 50 mM NaCl, 1 mM DTT, and 0.15% n-decyl-β-D-maltopyranoside; except 150 mM NaCl was used for TRPV4-ARD mutant analyses to preserve protein solubility). For ATP competition assays, competing compounds were added to reaction mixtures prior to the agarose slurry. All nucleotides used were sodium salts diluted from 0.5 M stocks adjusted to pH 7 with NaOH. The CaM-agarose assays were performed in binding buffer supplemented with 2 mM $CaCl_2$ or 5 mM EGTA (pH 7.5). In each load lane, the volumes loaded corresponded to 2 µg of protein. Gels were quantified using ImageJ (Abramoff *et al.*, 2004), and shown are the average ± S.D. for at least three independent experiments.

Insect and mammalian cell culture and full-length TRPV protein expression

Sf21 insect cells were maintained in Hink's TNM-FH (Mediatech, Manassas, VA), supplemented with 10% fetal bovine serum, 0.1% pluronic F-68, and 10 µg/ml gentamycin. Cells at 5×10^5 cells/ml were adhered to glass coverslips in medium without pluronic F-68 and infected with baculovirus. HEK293 cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, GlutaMAX (Invitrogen), 100 units/ml penicillin and 100 µg/ml streptomycin. Cells were co-transfected with pNEGFP and pcDNA3 containing the appropriate full-length TRPV using Lipofectamine 2000 (Invitrogen, Carlsbad, CA) according to the manufacturer directions.

Electrophysiology

Insect cells were tested 44–48 h post-infection, and HEK293 cells were tested 20–25 h post-transfection under continuous perfusion using a multichamber perfusion apparatus for agonist application. 2-APB and thymol were dissolved in dimethyl sulfoxide and 4αphorbol 12,13-didecanoate (4αPDD) in ethanol prior to dilution in bath solution. Currents were recorded and analyzed as described (Lishko *et al.*, 2007). Data are presented as mean ± S.E. The intracellular/pipette solution contained 140 mM NaMethanesulfonate, 10 mM HEPES, and either (4 mM NaCl and 10 mM EGTA) for EGTA conditions or (0.6 mM $MgCl_2$ and 10 mM BAPTA, resulting in 0.4 mM free Mg^{2+} according to MaxChelator (Patton *et al.*, 2004)) for BAPTA conditions. The BAPTA conditions were very similar to those used in Lishko *et al.*, 2007. The pH was adjusted to 7.2 with NaOH, and the final osmolarity was ~315 mOsm. As indicated, the intracellular solution was supplemented with 4 mM ATP (sodium salt) or ATPγS (lithium salt) from 0.5 M stocks (pH adjusted to ~7 with NaOH). In EGTA conditions, all ATP should be free ATP, whereas in BAPTA conditions, the presence of 0.6 mM $MgCl_2$ results in 0.001 mM free Mg^{2+}, 0.58 mM Mg-ATP, and 3.42 mM free ATP (Patton *et al.*, 2004). For CaM depletion experiments, the intracellular solution was supplemented with 2 μg/ml CaM85 or an isotype-matched control antibody (Invitrogen). The extracellular/perfusion solution was 150 mM NaCl, 5 mM KCl, 1 mM $MgCl_2$, 2 mM $CaCl_2$, 10 mM HEPES, and 10 mM D-glucose (pH adjusted to 7.4 with NaOH; ~315 mOsm), except for TRPV3 dose-response experiments and TRPV4 voltage step experiments in insect cells, in which the extracellular solution was 150 mM NaGluconate, 10 mM NaCl, 2 mM $CaCl_2$, 10 mM HEPES, and 10 mM D-glucose (pH adjusted to 7.2 with NaOH; ~ 315 mOsm), which produced more stable seals with less leak current at high agonist concentrations.

Data analysis

EC_{50} values were calculated by fitting the average normalized current at −100 mV for a range of agonist concentrations to the Hill Equation, $I(S) = 1 - (K^n/(K^n + S^n))$, where I is the current, K is the EC_{50}, S is the agonist concentration, and n is the Hill coefficient. Tail currents from voltage step experiments in HEK293 cells used for the determination of TRPV4 $V½$ were measured during the first millisecond of a step to a voltage of −160 mV and normalized to the maximum current. Average tail currents were fit to a modified Boltzmann function: $G(V) = G_{\max} - (G_{\max} - G_{\min})/(1 + \exp(zF/RT^{*}(V - V½)))$, where z is the valence of the gating charge and F/RT is 25 mV^{-1}. Statistical analyses were performed using a two-tailed t test, with $p < 0.05$ being considered statistically significant. Data are presented as mean ± S.E.

Results

TRPV3-ARD and TRPV4-ARD bind ATP and Ca^{2+}-CaM

To determine whether the ATP/CaM-binding site on the TRPV1-ARD is conserved in other TRPV channels, the ARDs from all six TRPV channels common to warm-blooded vertebrates were tested for ATP binding in pull-down assays (Figure 2.3.1). As previously observed (Lishko *et al.*, 2007; Phelps *et al.*, 2008), the TRPV1-ARD bound ATP-agarose, while the TRPV2-, TRPV5- and TRPV6-ARDs did not. Both TRPV3-ARD and TRPV4-ARD were precipitated by ATP-agarose (Figure 2.3.1A), suggesting that the TRPV1-ARD ATP-binding site is conserved in TRPV3 and TRPV4. Furthermore, the three ARDs that interact with ATP, the TRPV1-, TRPV3- and TRPV4-ARDs, were also precipitated with CaM-agarose in the presence of Ca^{2+}, and this interaction was eliminated in the presence of EGTA, a Ca^{2+}-chelator (Figure 2.3.1A). As previously determined, the TRPV2-, TRPV5-, and TRPV6-ARDs interacted either very weakly or not at all with CaM-agarose (Figure 2.3.1B) (Lishko *et al.*, 2007; Phelps *et al.*, 2008).

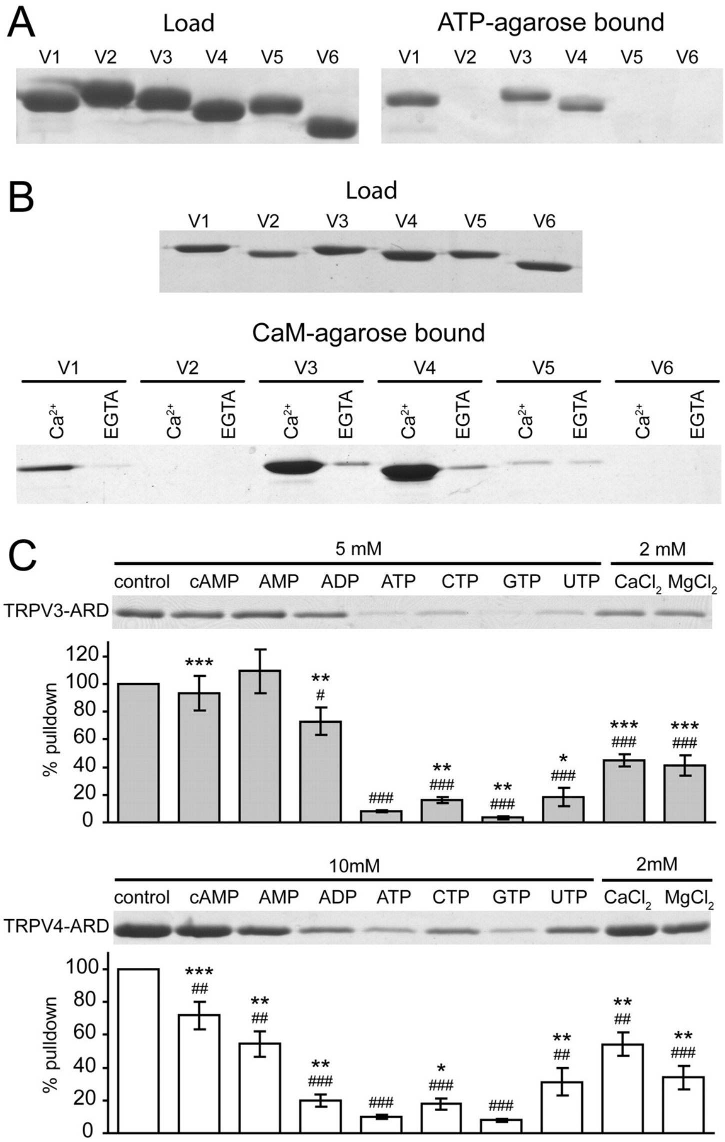


Figure 2.3.1: **Interactions of TRPV ARDs with ATP and CaM.** A) Coomassie-stained gel of an ATP-agarose pull-down assay with the six TRPV ARDs, showing loaded (*left*) and ATP-agarose-bound (*right*) proteins. B) Coomassie-stained gels of a CaM-agarose pull-down assay of the six TRPV ARDs showing loaded protein (*top*) and protein bound in the presence of Ca^{2+} or EGTA (*bottom*). C) Nucleotide specificity of the TRPV3- and

Figure 2.3.1 (Continued) TRPV4-ARD. Coomassie-stained gels of wild-type TRPV3-ARD (*top*) or TRPV4-ARD (*bottom*) bound to ATP-agarose in the presence of the indicated concentration of competing compounds. The *histogram* below each representative gel shows the average amount of protein recovered (±S.D.) in the absence or presence of nucleotide and divalent cations over four experiments. The statistical significance with respect to control (#, $p < 0.05$; ##, $p < 0.01$; ###, $p < 0.001$) and ATP (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$) was determined using two-tailed t tests.

The ATP and Ca^{2+}-CaM binding site is conserved in TRPV3-ARD and TRPV4-ARD

To further characterize the properties of the ATP-binding site on TRPV3 and TRPV4, we tested its specificity in competition assays with other nucleotides. As previously reported with TRPV1 (Lishko *et al.*, 2007), free GTP and ATP most efficiently competed for binding to ATP-agarose for both TRPV3-ARD and TRPV4-ARD (Figure 2.3.1C). Furthermore, Ca^{2+} and Mg^{2+} also reduced binding to ATP-agarose (Figure 2.3.1C). Therefore, the ATP-binding sites on the ankyrin repeats of TRPV3 and TRPV4 have the highest affinity for divalent-free triphosphate nucleotides, with a small preference for purines over pyrimidines, a specificity profile comparable to TRPV1-ARD (Lishko *et al.*, 2007).

The similar nucleotide specificities of TRPV3-ARD, TRPV4-ARD, and TRPV1-ARD strongly suggest that ATP interacts with these domains at a conserved site. The overall sequence conservation of TRPV1, TRPV3, and TRPV4 (Figure 2.3.9) was mapped onto the structure of TRPV1-ARD bound to ATP (Figure 2.3.2A). The most conserved surface encompasses the ATP-binding site. We used mutagenesis to confirm that the conserved phosphate-binding residues are important for the interaction of TRPV3-ARD and TRPV4-ARD with ATP. Lys^{155} and Lys^{160} interact with the triphosphate moiety of ATP in the TRPV1-ARD structure and are important for ATP and CaM binding (Lishko *et al.*, 2007). The corresponding lysines, Lys^{169} and Lys^{174} in TRPV3-ARD and Lys^{178} and Lys^{183} in TRPV4-ARD, were mutated to alanine. Arg^{188} in TRPV3 and Lys^{205} in TRPV4, predicted to lie on the opposite face of the ARDs, were also mutated to alanine and used as negative controls. In pull-down assays, the TRPV3-ARD and TRPV4-ARD lysine mutants showed reduced binding to both ATP and CaM compared with the wild-type proteins and negative control mutants (Figure 2.3.2B and C). In summary, the lysines homologous to Lys^{155} and Lys^{160} in TRPV1 are also necessary for TRPV3 and TRPV4 interactions with ATP and CaM, indicating that the multiligand binding site previously identified in TRPV1 (Lishko *et al.*, 2007) is conserved in TRPV3 and TRPV4.

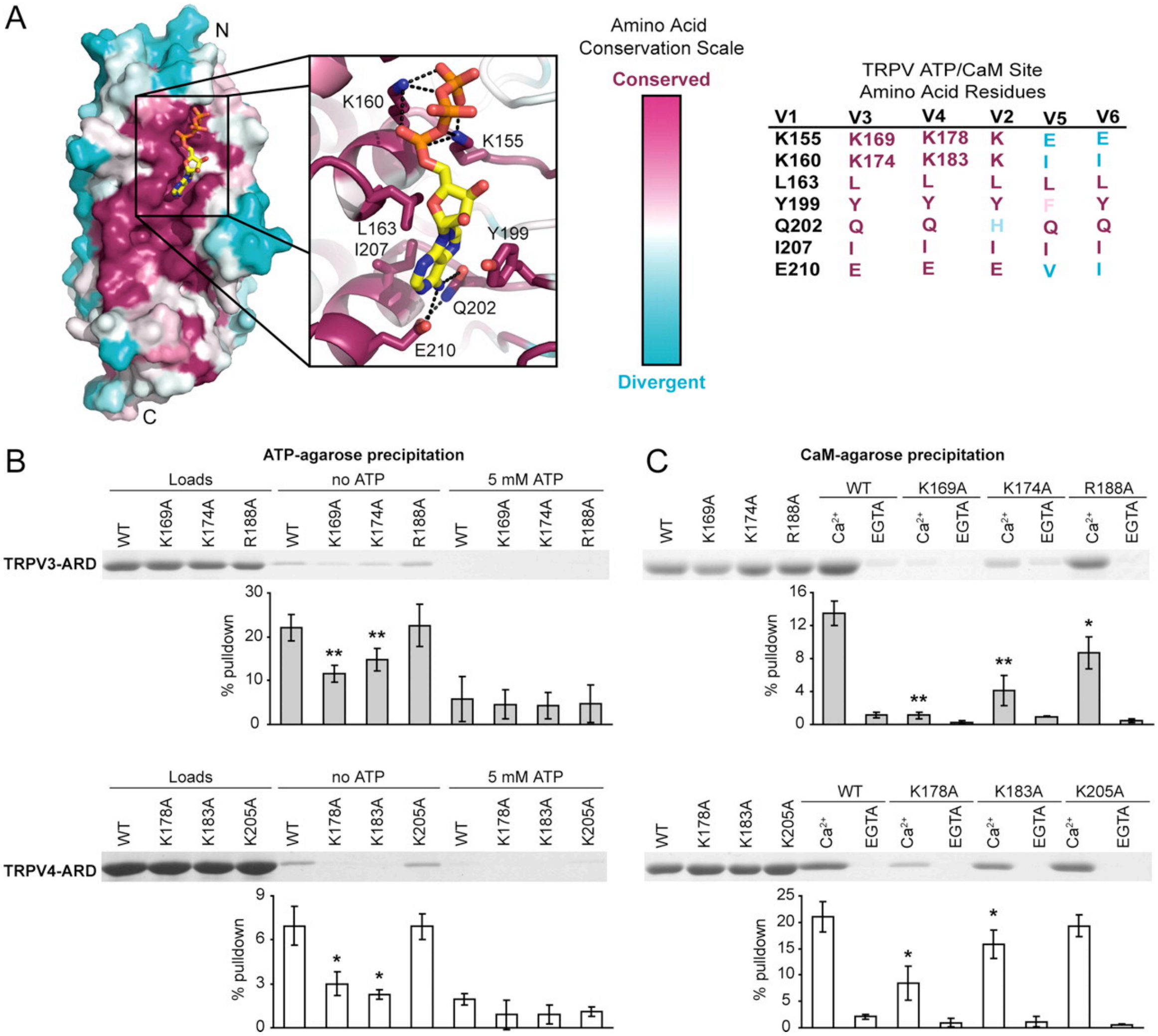


Figure 2.3.2: A conserved ATP/CaM binding site in the ARDs of TRPV1, TRPV3, and TRPV4. A) The amino acid conservation between these three ARDs was calculated and mapped onto the surface of the TRPV1-ARD structure (Protein Data Bank code 2PNN) using Consurf (Landau *et al.*, 2005) based on the alignment in Figure 2.3.9. The most conserved and divergent residues are *purple* and *cyan*, respectively. The ATP binding site is magnified to show the amino acid side chains that contact ATP. The identity of the TRPV1 site and corresponding residues in the other five TRPVs is shown on the *right.* B) Coomassie-stained gels of wild-type and mutant TRPV3-ARD (*top*) or TRPV4-ARD (*bottom*) loaded (*left*) and bound to ATP-agarose in the absence (*middle*) or

Figure 2.3.2 (Continued) presence (*right*) of competing free ATP. *C*, Coomassie-stained gels show wild-type and mutant TRPV3-ARD (*top*) or TRPV4-ARD (*bottom*) loaded (*left*) and bound to CaM-agarose in the presence of Ca^{2+} or EGTA. In B) and C), the average percentage of protein recovered (±S.D.) is plotted below. The statistical significance of the reduction in binding to ATP-agarose or Ca^{2+}-CaM-agarose with respect to wild type (*WT*) was determined by one-tailed *t* tests, with $p < 0.05$ and $p < 0.01$ indicated by * and **, respectively.

TRPV2 is insensitive to intracellular ATP

Electrophysiology experiments demonstrated that intracellular ATP can sensitize TRPV1 and prevent its desensitization to repeated applications of capsaicin (Lishko *et al.*, 2007). Experiments with the K155A and K160A mutants of TRPV1 also indicated that these effects of ATP were through its direct interaction with the TRPV1-ARD. Rat TRPV2 was hypothesized to be a natural negative control; ATP and Ca^{2+}-CaM were not expected to affect its sensitivity, because its ARD did not bind either. Rat TRPV2 expressed in HEK293 cells responded to 2-APB in whole cell patch clamp recordings as reported previously (Hu *et al.*, 2004; Neeper *et al.*, 2007). TRPV2 exhibited similar currents when stimulated with 2-APB in the absence or presence of intracellular ATP (Figure 2.3.3). Furthermore, no significant desensitization or tachyphylaxis was observed in response to repeated 2-APB applications. Therefore, TRPV2 activity was not affected by the presence of intracellular ATP, correlating with the lack of interaction between ATP and the TRPV2-ARD.

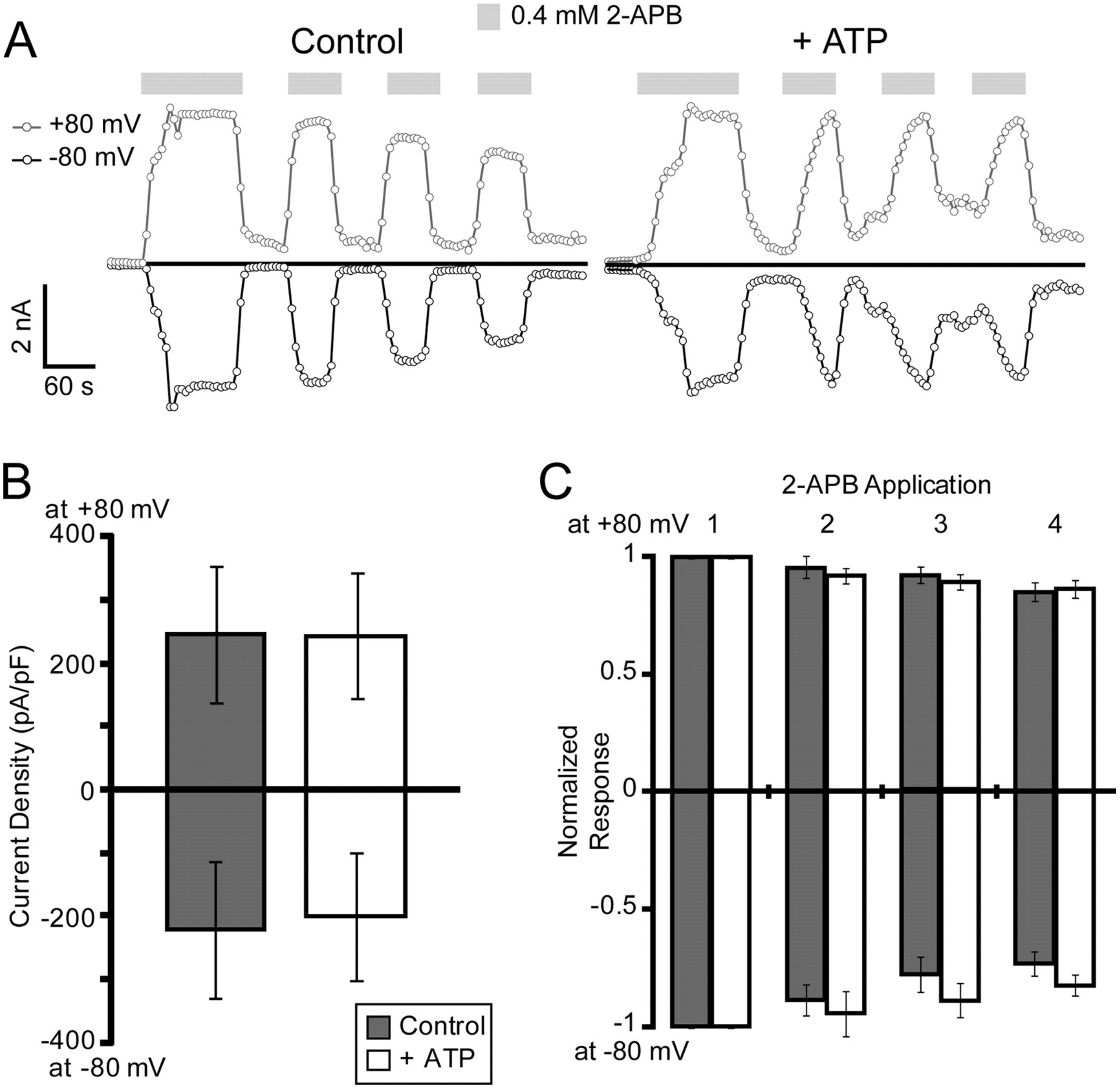


Figure 2.3.3: TRPV2 is insensitive to intracellular ATP. A) Sample whole cell patch clamp recordings from TRPV2 expressing HEK293 cells with (*right*) and without (*left*) 4 mM ATP in the intracellular solution. Currents at +80 mV (*gray*) and −80 mV (*black*) were extracted from linear voltage ramps. *Gray bars* indicate perfusion with 0.4 mM 2-APB, and *black lines* indicate zero current. B) Average maximum current density evoked during the first 2-APB application. C) TRPV2 does not undergo tachyphylaxis. Currents evoked by multiple 2-APB applications at ±80 mV were normalized to the maximum current from the first 2-APB application. For both B) and C), control cells (n = 6) are

Figure 2.3.3 (Continued) colored *gray* and cells with intracellular solution supplemented with 4 mM ATP are colored *white* ($n = 6$).

We attempted to generate a TRPV2-ARD mutant that could bind ATP and/or CaM. We looked at two mutations: D78N, which neutralizes a negatively charged side chain which maps in close proximity of the phosphate-interaction site, and H165Q, to attempt to restore the adenine-binding pocket (Figure 2.3.10). Neither of the single mutants bound to ATP- or CaM-agarose in our assays. The D78N/H165Q mutant bound weakly but significantly to ATP, but not CaM. Because the TRPV2-ARD is only 50% identical to the TRPV1-ARD, it is difficult to determine which other sequence differences may be responsible for the differences in biochemical properties.

TRPV4 is sensitized by intracellular ATP

We used whole cell patch clamp electrophysiology to determine the effect of intracellular ATP on the sensitivity of TRPV4 expressed in insect and HEK293 cells. TRPV4 showed constitutive basal activity in both cell types (Figure 2.3.4 and Figure 2.3.11), similar to previous observations (e.g., Liedtke *et al.*, 2000; Strotmann *et al.*, 2000). In voltage step experiments in insect cells, TRPV4 currents were significantly increased in the presence of intracellular ATP or the non-hydrolyzable ATP analog ATPγS (Figure 2.3.4A). Furthermore, the K178A mutation, which reduces ATP binding, abolished sensitization by ATP (Figure 2.3.4A).

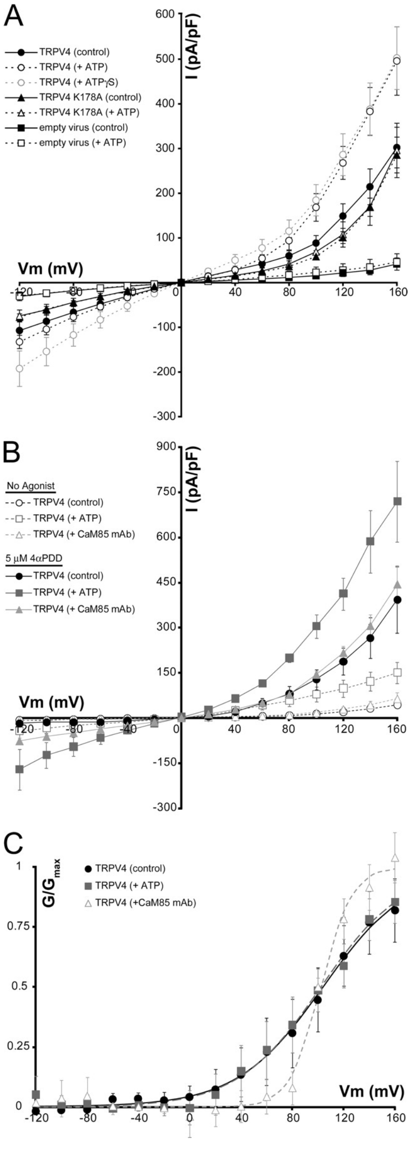


Figure 2.3.4: TRPV4 is sensitized by intracellular ATP. A) Average current density from voltage step experiments in insect cells plotted against holding potential for cells recorded in the absence (*solid symbols*) or presence of intracellular ATP (*black, open symbols*) or ATPγS (*gray, open symbols*). Data from control cells infected with empty virus (n = 4 each; *squares*), wild-type (*WT*) TRPV4 data (n = 7 each; *circles*), and TRPV4 K178A data (n = 7 each; *triangles*) are shown. ATP and ATPγS cause a significant current increase ($p < 0.05$ at $V_m > 100$ mV). B) Average current density plotted against holding potential from voltage step experiments in TRPV4-expressing

Figure 2.3.4 (Continued) HEK293 cells. Data were collected on unstimulated (*open symbols*) and 4αPDD-perfused cells (5 μM; *filled symbols*) with control intracellular solution (*black circles*, $n = 6$), 4 mM ATP (*dark gray squares*, $n = 7$) or an anti-CaM monoclonal antibody (CaM85, *light gray triangles*, $n = 6$). C) Activation curves from TRPV4-expressing HEK293 cells calculated from the average, normalized tail currents measured in the first milliseconds after a step to −160 mV from the cells in B) (control, *black circles*; ATP, *dark gray squares*; CaM85 mAb, *light gray triangles*). *Lines* represent the fit of a modified Boltzmann function to the data.

Similar results were obtained from basal TRPV4 currents in HEK293 cells (Figure 2.3.4B), although the lower constitutive activity in HEK293 cells enabled us to also look at 4αPDD-stimulated activity. Currents observed after perfusion with 4αPDD were also significantly increased by the addition of ATP to the recording solution (Figure 2.3.4B). The effect of ATP was similar in both 4αPDD-stimulated and constitutive conditions (at +100 mV, constitutive currents increased 1.9-fold and 3.0-fold in insect and HEK293 cells, respectively, while 4αPDD-stimulated currents increased 2.4-fold). Furthermore, depleting HEK293 cells of CaM by including a monoclonal anti-CaM antibody in the intracellular solution, as was previously done in TRPV1-expressing cells (Lishko *et al.*, 2007), did not affect the voltage response of unstimulated TRPV4, but did significantly increase inward currents in TRPV4 expressing HEK293 cells treated with 4αPDD ($p < 0.05$ at $V_m \leq -20$ mV).

Similar to TRPV1 (Lishko *et al.*, 2007), this increased current density for TRPV4 in the presence of intracellular ATP appears to be a result of increased whole cell conductance, rather than a shift in the current-voltage relationship. Tail-current analyses from voltage step experiments in TRPV4-expressing HEK293 cells with control, 4 mM ATP, or anti-CaM antibody intracellular solutions (Figures 2.3.4C and 2.3.11) show that the V½ of TRPV4 is not altered by ATP or CaM with *V*½ values of 104 ± 36.4, 102 ± 8.3, and 102 ± 13.6 mV for the control, ATP and anti-CaM experiments, respectively. Of note, the anti-CaM antibody increases the steepness of the *G*/*V* curve, suggesting that CaM may affect the intersubunit cooperativity of TRPV4. Overall, these results strongly suggest that the previously observed potentiation or desensitization by binding of intracellular ATP or CaM, respectively, to the N-terminal ankyrin repeats of TRPV1 is conserved in TRPV4.

ATP lowers the agonist sensitivity of TRPV3

Similar to previously published reports using mammalian cells (Xiao *et al.*, 2008; Chung *et al.*, 2004), TRPV3 expressed in insect cells is sensitized by repeated applications of 2-APB (Figure 2.3.5A). Once sensitized, TRPV3 also showed biphasic currents (Figure 2.3.5A) where the initial outward rectified current (I_1) is followed by an off-response with the appearance of a less rectified, higher amplitude current that is slower to inactivate (I_2), similar to the currents reported in HEK293 cells and primary keratinocytes overexpressing TRPV3 (Chung *et al.*, 2005). The sensitization of TRPV3 to repeated agonist applications is in contrast to what is observed with TRPV1, which is desensitized by repeated agonist applications (Koplas *et al.*, 1997; Lishko *et al.*, 2007). Also unlike TRPV1 and TRPV4, intracellular ATP blocked the sensitization of TRPV3 to repeated 2-APB applications (Figure 2.3.5B). The same effect was observed when ATPγS was used, supporting the idea that it is ATP binding, not an ATP hydrolysis-dependent process, that prevents TRPV3 sensitization. There is no significant difference between the currents observed during the first and twelfth 2-APB applications in presence of intracellular ATP or ATPγS. Furthermore, the currents observed on the twelfth 2-APB application with the control cells are significantly larger than in cells with intracellular ATP or ATPγS (Figure 2.3.5B). Additionally, while biphasic currents and off-responses were observed for seven of the nine control cells tested, none of the ATP (0/6) or ATPγS (0/7) cells showed biphasic currents or off-responses.

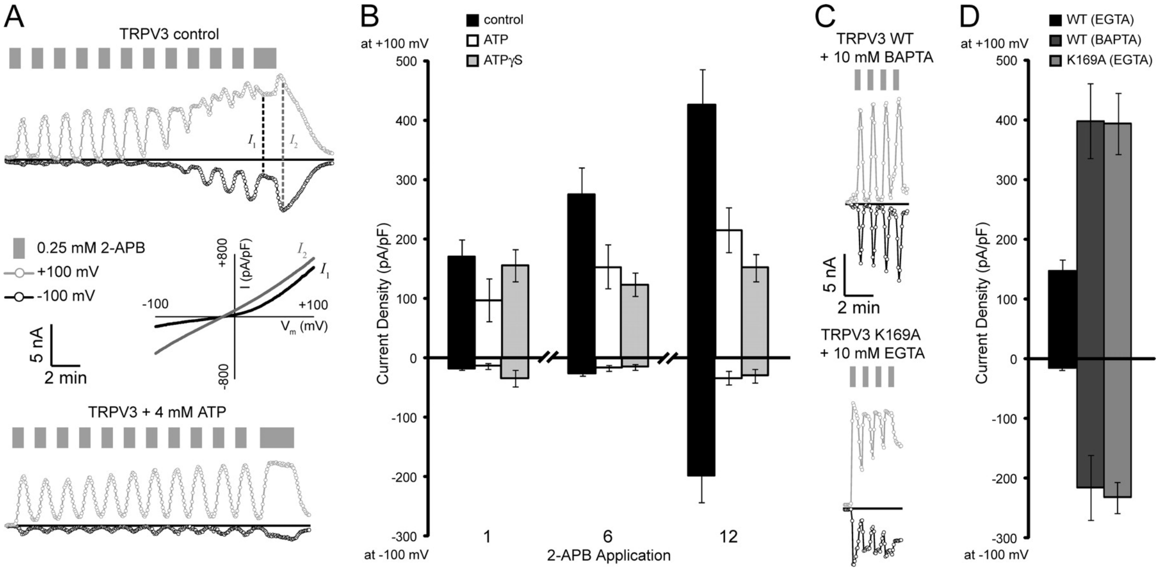


Figure 2.3.5: Sensitization of TRPV3 in insect cells. A) Sample whole cell patch clamp recordings from baculovirus-infected insect cells expressing wild-type TRPV3. Shown are currents at +100 (*gray circles*) or −100 mV (*black circles*) extracted from linear voltage ramps from a control cell (*top*) and cells with intracellular ATP (*bottom*). Applications of 0.25 mM 2-APB are indicated by *gray bars* and zero current by *black lines*. For the control cell the *dashed lines* indicate the time points for the *I–V* traces

Figure 2.3.5 (Continued) (plotted as current density *versus* membrane voltage) from type 1 (I_1) and off-response type 2 (I_2) currents, which are shown below the control. B) Control cells are sensitized by repeat applications of 2-APB, and this is blocked by ATP and ATPγS. Average current densities (pA/pF) at +100 and −100 mV are shown for the first, sixth, and twelfth applications of 2-APB to TRPV3-expressing cells with intracellular solutions containing no nucleotide (control; *black bars*, $n = 9$), ATP (*white*, $n = 6$) or ATPγS (*gray*, $n = 7$). C) Sample whole cell recordings from insect cells expressing wild-type TRPV3 with BAPTA as the intracellular calcium buffer (*top*) and TRPV3 K169A with EGTA as the intracellular calcium buffer (*bottom*) collected and displayed as in A). D) Average maximum current density at +100 and −100 mV from a 30 s application of 0.25 mM 2-APB for wild-type (*WT*) TRPV3 with EGTA (*black bars*) or BAPTA (*dark gray bars*), and K169A TRPV3 with EGTA (*gray bars*). Note that C) and D) are on the same scales as A) and B), respectively.

The sensitization of TRPV3 is dependent on the strength of the intracellular Ca^{2+} buffer. When BAPTA, a more rapid and specific Ca^{2+} buffer, was used in place of EGTA, TRPV3 was pre-sensitized, showing large responses to the first application of 2-APB and little increased sensitivity to subsequent 2-APB applications (Xiao *et al.*, 2008). This behavior could also be reproduced in our insect cell system (Figure 2.3.5C and D). Also, TRPV3 K169A (one of the ATP/CaM site mutants that no longer bound ATP or CaM) (Figure 2.3.2) showed initial current densities similar to those of wild-type TRPV3 in the presence of BAPTA, even when EGTA was used as the Ca^{2+} buffer (Figure 2.3.5). The TRPV3 K169A currents were similar to the I_2 currents observed with sensitized wild-type

TRPV3, with large amplitudes, little rectification, and slower deactivation after removal of 2-APB. Consistent with a sensitized state, the average current density from the first 2-APB application for TRPV3 K169A was as large as that for wild-type TRPV3 either from the twelfth 2-APB application in experiments with EGTA as the Ca^{2+} buffer, or the first 2-APB application when pre-sensitized with BAPTA as the Ca^{2+} buffer (Figure 2.3.5). These results show that disruption of ATP/CaM binding eliminates the sensitivity of TRPV3 to intracellular Ca^{2+} levels and indicate that interaction of ATP and/or Ca^{2+}-CaM on the N-terminal ankyrin repeats regulates TRPV3 sensitization.

To further characterize the mechanism by which intracellular ATP regulates TRPV3 sensitivity, dose-response relationships were measured for two different TRPV3 agonists, 2-APB and thymol, in the absence or presence of intracellular ATP (Figure 2.3.6). Dose-response experiments were carried out with BAPTA to pre-sensitize TRPV3 and remove any confounding effects on the dose response from repeated agonist applications. Additionally, the chloride ions in the extracellular solution were replaced with gluconate. Replacing chloride with gluconate lowers the agonist-induced TRPV3 currents, allowing us to also determine the dose response of TRPV3 K169A, which was otherwise difficult to inactivate after the first agonist application (Figure 2.3.12). Thymol concentrations above 1 mM were toxic and as a result saturated currents could not be recorded in the presence of ATP. The responses of TRPV3 to thymol in the absence or presence of ATP were both normalized to the maximum current density from the experiments without ATP. ATP increases the EC_{50} of both agonists by ~3-fold (Figure 2.3.6), indicating that intracellular ATP reduces the sensitivity of TRPV3 to its agonists.

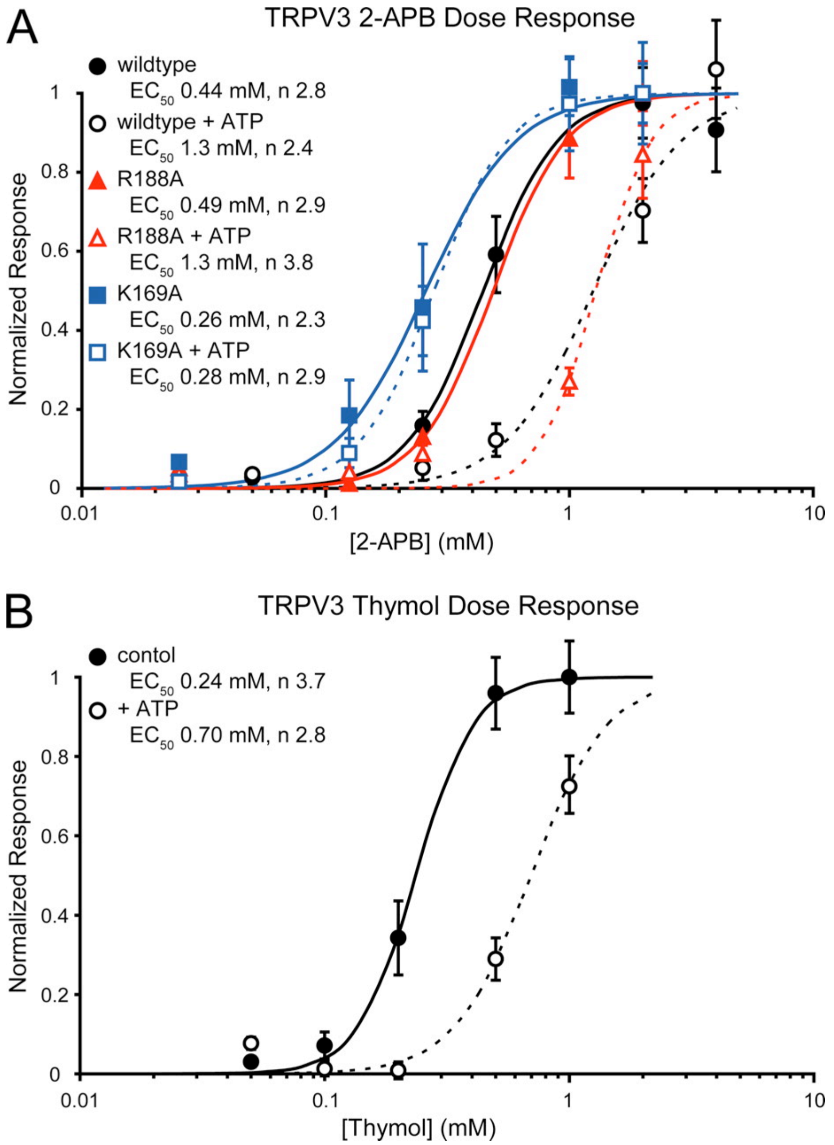


Figure 2.3.6: ATP lowers the sensitivity of TRPV3 to chemical agonists. A) Dose response of TRPV3 to 2-APB. The dose response of wild-type (*black circles*), R188A (*red triangles*), and K169A (*blue squares*) TRPV3 to 2-APB were determined from control cells (*filled symbols*) and cells with intracellular ATP (*open symbols*). Normalized responses (based on the average maximum current density at −100mV) are plotted against the concentration of 2-APB. Fits of the data to the Hill equation are shown as *solid* (control cells) or *dashed lines* (+ ATP), and the resulting EC$_{50}$ and Hill

Figure 2.3.6 (Continued) coefficients (*n*) values are listed for each sample. B) Dose response of wild-type TRPV3 currents to thymol, measured as in A), showing control cells (*filled circles*; *solid line*) and cells with intracellular ATP (*open circles*; *dashed line*).

TRPV3 K169A showed an increased sensitivity to 2-APB with an EC_{50} ~2-fold lower than wild-type TRPV3, and it is unchanged in the presence of ATP (Figure 2.3.6). The changes in sensitivity between K169A and wild-type TRPV3 are not due to differences in expression (Figure 2.3.13) and instead could result from the loss of ATP and CaM binding in K169A. In contrast, the behavior of the TRPV3 R188A control mutant is indistinguishable from wild type either in the presence or absence of intracellular ATP (Figure 2.3.6). These results strongly support the role of ATP binding to the TRPV3-ARD in altering agonist sensitivity.

TRPV3 behaved similarly in HEK293 cells. Intracellular ATP or ATPγS reduced sensitization compared with control (EGTA), whereas intracellular BAPTA caused significantly increased current densities for the first three 2-APB applications (Figure 2.3.7). In HEK293 cells, we could also test the effects of CaM depletion using an anti-CaM monoclonal antibody. Addition of anti-CaM antibody into the intracellular solution led to immediate sensitization and significantly larger current densities for all 2-APB applications, while currents in the presence of an isotype-matched control antibody were indistinguishable from control (Figure 2.3.7B). Furthermore, the 2-APB-induced currents from CaM-depleted cells were similar to those observed with the K169A mutant in insect cells in that they did not inactivate upon agonist removal and showed little rectification (compare Figure 2.3.5C with Figure 2.3.7A). Taken together, the results from HEK293 and insect cells indicate that there is direct role for CaM binding to the conserved ARD site in TRPV3 inactivation and that ATP binding to the same site can maintain TRPV3 in a low sensitivity state. This is in agreement with a previous report that sensitization of TRPV3 results from a loss of CaM binding (Xiao *et al.*, 2008) and further demonstrates a role for the ARD in this CaM-mediated regulatory mechanism.

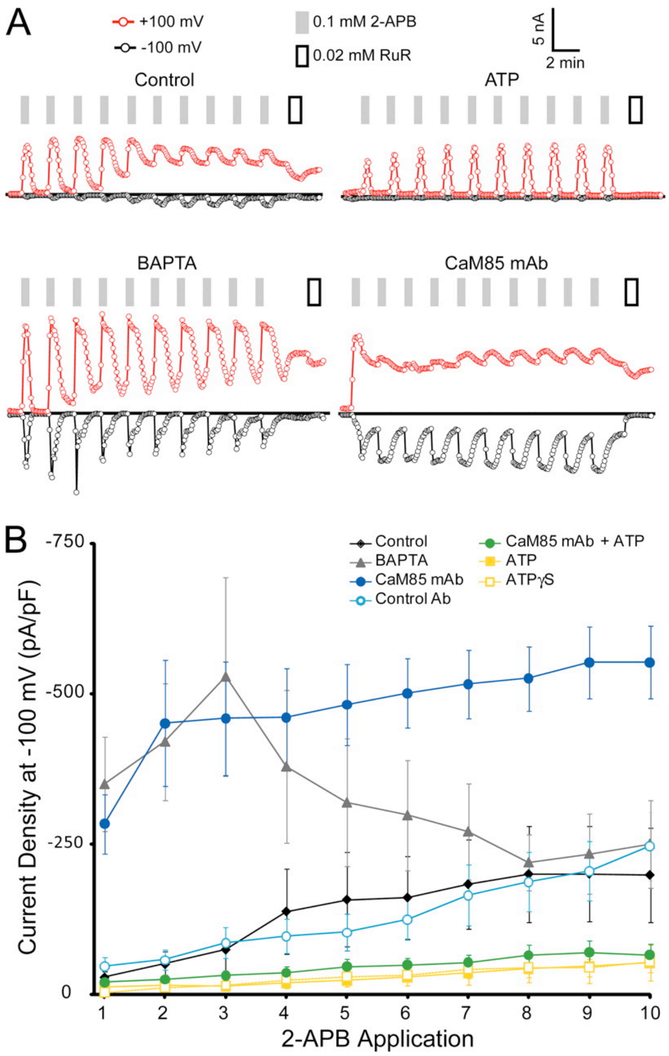


Figure 2.3.7: Ca^{2+}-CaM and ATP decrease the sensitivity of TRPV3 in HEK293

Figure 2.3.7 (Continued) cells. A) Sample whole cell patch clamp recordings from transiently transfected HEK293 cells expressing wild-type TRPV3. Shown are currents at +100 (*red circles*) or −100 mV (*black circles*) extracted from linear voltage ramps from cells with different intracellular solutions; control (*top left*), 4 mM ATP (*top right*), 10 mM BAPTA (*lower left*), and 2 μg/ml anti-CaM antibody (*Ab*) (CaM85, *lower right*). Application of 0.1 mM 2-APB to the cells is shown by *gray bars. White bars* indicate application of 20 μM ruthenium red (*RuR*), a channel blocker. B) Average current density at −100 mV (in pA/pF or picoampere per picofarad) from 10 consecutive applications of 2-APB for cells with control intracellular solution (*black diamonds*), 10 mM BAPTA (*gray triangles*), CaM85 monoclonal antibody (*mAb*; *dark blue circles*), isotype matched control antibody (*light blue circles*), ATP (*yellow squares*), ATPγS (*open yellow squares*), and both ATP and CaM85 monoclonal antibody (*green circles*).

Discussion

We find that the binding of ATP and Ca^{2+}-CaM to the N-terminal ARD observed in TRPV1 (Lishko *et al.*, 2007) is conserved in two of the three other thermo TRPVs, TRPV3, and TRPV4. Intracellular ATP increased TRPV4 currents in response to voltage steps, indicating a sensitizing effect, similar to the effect of intracellular ATP on TRPV1, increasing currents in response to capsaicin. In contrast, the response of TRPV3 to agonists is reduced in the presence of intracellular ATP. More precisely, intracellular ATP prevented the sensitization of TRPV3 to repeated applications of 2-APB and increased the EC_{50} for agonists. Furthermore, mutagenesis and electrophysiology data support a role for the ATP/CaM binding site on the ARD in regulating both TRPV3 and

TRPV4 sensitivity. Therefore, although the biochemical properties of the TRPV1-, TRPV3-, and TRPV4-ARDs are similar, there are marked differences in functional consequences of modulatory interactions with the ARD.

Of the four thermosensitive TRPVs, only TRPV2 did not bind to ATP or Ca^{2+}-CaM through its ARD. This is likely a result of several amino acid substitutions within and around the conserved ATP/CaM binding site (Figures 2.3.2 and 2.3.10) (Jin *et al.*, 2006; McCleverty *et al.*, 2006 Lishko *et al.*, 2007). Accordingly, we saw no significant effects of intracellular ATP on TRPV2 currents (Figure 2.3.3). Although TRPV2-ARD does not interact with ATP or CaM, it is still important for TRPV2 function, since deletions of parts of the ARD impair activation by 2-APB or heat and surface localization (Neeper *et al.*, 2007).

Our data do not directly demonstrate a physical interaction between ATP and the TRPV ARDs under the patch clamp conditions where the effects of intracellular ATP on channel sensitivity were observed. It is therefore difficult to rule out an indirect effect of intracellular ATP. However, several observations support a direct binding of ATP to the ARDs. First, similar results are obtained in two different cell types, HEK293 and insect cells, ruling out factors that are not conserved in both cell types. Second, the effects of ATP can be observed in the absence of divalent cations and/or presence of chelator in the intracellular solution and are reproduced by ATPγS, a poorly hydrolyzable ATP analog. This argues against an ATP-hydrolysis-dependent process (e.g., phosphoinositide synthesis). Third, the disruption of the ligand-binding site on the ARD by mutagenesis, confirmed biochemically, eliminated the effect of ATP on channel function in TRPV1 (Lishko *et al.*, 2007), TRPV3, and TRPV4. This supports a direct role for ATP binding to the ARD in regulating TRPV channel sensitivity.

What might be the physiological purpose of intracellular ATP-meditated regulation of TRPV ion channels? As suggested above, the overall role of the ATP/CaM-binding site on the ARD may be to tune the sensitivity of TRPV channels. Regulation by intracellular ATP has also been observed in other ion channels, including TRP channels TRPC5 (Dattilo *et al.*, 2008), TRPM4 (Nilius *et al.*, 2005), and TRPM6 (Thébault *et al*., 2008). K_{ATP} channels use several nucleotide-binding sites to sense nucleotide levels and have been implicated in sensing metabolic levels in tissues ranging from muscles to the pancreas to neurons, tying membrane potential to the metabolic level of the cell (Bryan *et al.*, 2007). Furthermore, the C-terminal domain of ClC-type chloride channels binds adenine nucleotides (Meyer *et al.*, 2007), and, at least under some circumstances, intracellular adenine nucleotides inhibit ClC channels, although the ATP-mediated regulation of ClCs remains controversial (Accardi, 2008). Hence, intracellular ATP may play an important role in modulating physiological functions of multiple channel families including TRPV channels. The data on fluctuations of nucleotide concentration in cellular physiology are still sparse, but some studies suggest that such variations may be important (Ataullakhanov and Vitvitsky, 2002). Thus, changes in cellular nucleotide concentrations reflecting the metabolic state, either local or global, could directly affect TRPV channel sensitivity.

Alternatively, ATP may act as a cofactor in sensing Ca^{2+} levels. ATP binding to TRPV ARDs is sensitive to the divalent cation concentration: only free ATP has high affinity for the binding site. High concentrations of Ca^{2+} disrupt the interaction with ATP, presumably through Ca^{2+} chelation by the triphosphate moiety and favor the interaction with Ca^{2+}-CaM. Of note, although most ATP is chelated by Mg^{2+} *in vivo*, the cellular concentration of free ATP is still significant, ranging from 0.3–0.7 mM (Taylor *et al.*, 1991 and references therein). It was suggested that the competition of ATP with CaM for the same binding site on the ARD could provide sensitivity to global Ca^{2+} levels while making the channel less sensitive to transient local Ca^{2+} concentration changes that rapidly dissipate (Tadross *et al.*, 2008). That is, the competition between ATP and CaM affects the kinetic and thermodynamic parameters of the channel modulation by Ca^{2+}. In such a scenario, ATP could be considered a cofactor tuning the sensitivity of TRPV channels to intracellular Ca^{2+}.

The different modulatory effects of the ATP/CaM binding site on TRPV3 *versus* TRPV1 and TRPV4 may have arisen to provide different basal sensitivity and/or feedback mechanisms. That is, the physiological roles of these channels, which are still being uncovered (see Vennekens *et al.*, 2008 for a recent review), likely require different adaptation and potentiation mechanisms. TRPV3, unlike TRPV1 and TRPV4, is sensitized by repeated agonist applications. The data presented here (Figure 2.3.7) and by others (Xiao *et al.*, 2008) clearly show that TRPV3 is sensitized by the removal of CaM. Here, we further show that these effects are mediated through the conserved ATP/CaM site in the TRPV3-ARD (Figure 2.3.5). Moreover, ATP binding maintains the

TRPV3 channel in a low sensitivity state, even though it also prevents CaM binding. We hypothesize that TRPV3 undergoes a conformational change in the open state that decreases the ability of TRPV3 to bind CaM, making TRPV3 easier to open and slower to close. The channel is slow to revert back to the CaM-binding state, and therefore further stimulations result in an increased population of the sensitized TRPV3 state. On the other hand, according to our model, ATP binding to the ARD holds TRPV3 in a lower sensitivity state, requiring higher agonist concentrations to activate the channel (Figure 2.3.6) and preventing the transition to the sensitized state.

The structural similarity of the ligand-free TRPV2-ARD (Jin *et al.*, 2006; McCleverty *et al.*, 2006) and ATP-bound TRPV1-ARD (Lishko *et al.*, 2007) suggests that ligand binding causes little conformational change in the ankyrin repeats. This is supported by a recent survey of ankyrin repeat structures; ligand binding typically imposes little conformational change on ankyrin repeats (Gaudet, 2008). The molecular basis for the differences between TRPV3 and its close homologs, TRPV1 and TRPV4, may instead originate from distinct pathways within the protein to decode the bound regulatory ligand on the ARD and communicate to the channel gate. Notably, the TRPV3-ARD sequence diverges from both the TRPV1-ARD and TRPV4-ARD sequences, with several deviations on the side opposite the conserved binding site (Figure 2.3.9). Hence, these sequence variations may contribute to the different responses to intracellular ATP in TRPV1 and TRPV3 by engaging distinct interactions with other regions of the channel. Differences outside of the ARD likely contribute as well. In fact, mutagenesis screens on TRPV1 (Myers *et al.*, 2008) and TRPV3 (Grandi *et al.*, 2008) identified other regions involved in channel sensitivity and activation, which might act in concert with the ARDs to regulate channel activity (Figure 2.3.8). In TRPV3, the pore region has been implicated in both heat activation (Grandi *et al.*, 2008) and regulation by extracellular calcium (Xiao *et al.*, 2008). The intracellular membrane proximal regions at both the N and C termini were implicated in 2-APB activation (Hu *et al.*, 2009). Further experiments will be required to determine how all these different local interactions converge to effect channel gating. The conserved biochemical properties but distinct modulatory outputs of ATP binding to the ARD of TRPV1 and TRPV4 *versus* TRPV3 provide a starting point to design such experiments.

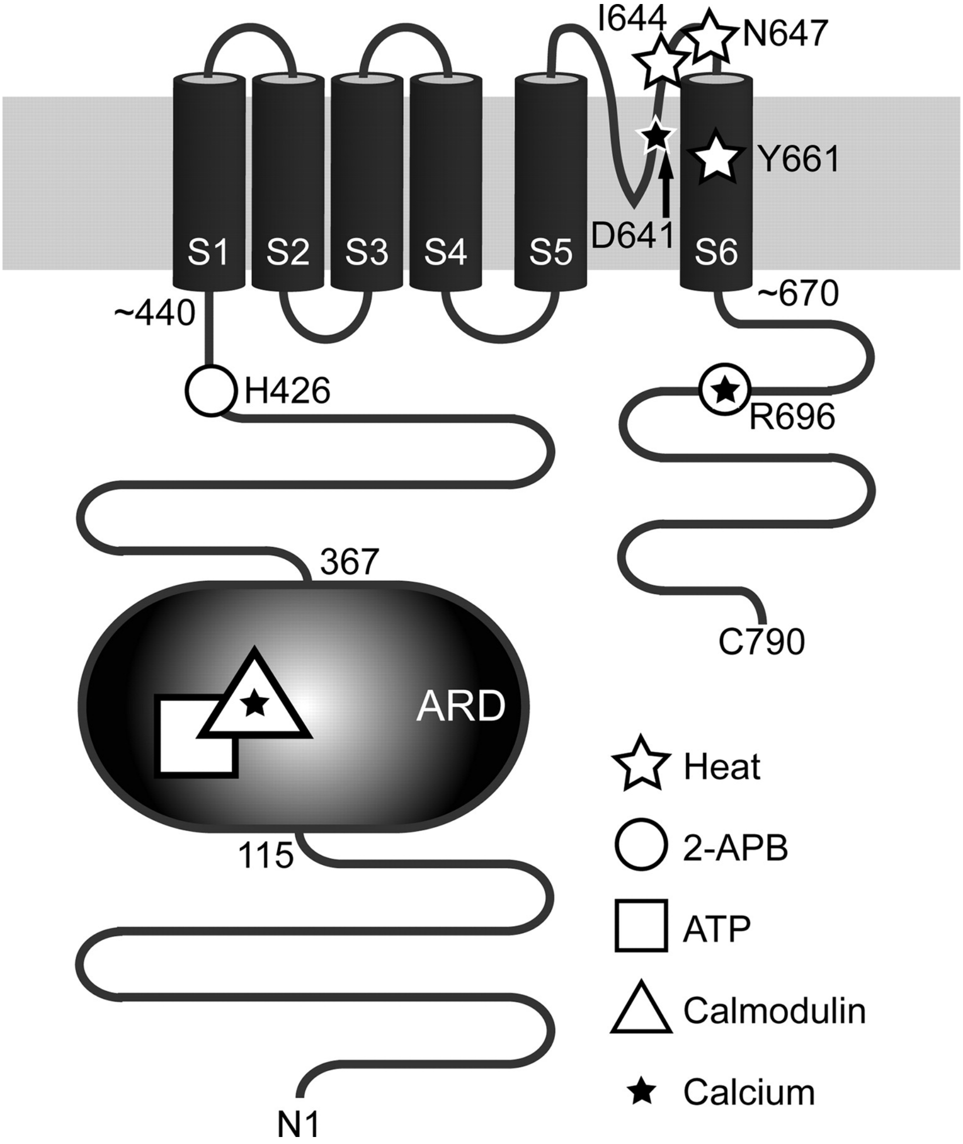


Figure 2.3.8: Topology of TRPV3 and location of functionally important sites. The ARD is important in the tuning of TRPV3 sensitivity and interacts with ATP (*square*) and Ca^{2+}-CaM (*starred triangle*; see also Xiao *et al.*, 2008). The previously identified sites of heat activation (Ile^{644}, Asn^{647}, and Tyr^{661} marked as *white stars* (Grandi *et al.*, 2008)), activation by 2-APB (cytoplasmic residues His^{426} and Arg^{696}, *circles* (Hu *et al.*, 2009)), and calcium sensitivity (intracellular site Arg^{696} (Grandi *et al.*, 2008)) and extracellular site Asp^{641} (Xiao *et al.*, 2008)) are also indicated.

Chapter 3: Structure and function of MP20

Introduction

MP20 is a member of the PMP20/EMP/MP20/claudin superfamily (Van Itallie *et al*., 2006). It is the second most abundant membrane protein in fiber cells of the lens, where it is expressed predominantly in the cortical regions (Alcalá *et al.*, 1975). The function of MP20 in the lens is currently unclear, but its importance in proper lens function is demonstrated by the fact that mutations in MP20 have been associated with cataracts and blindness (Pras *et al*., 2002; Steele *et al*., 1997; Steele *et al*., 2000; Shi *et al.,* 2011). In humans, the F105V mutation in MP20 was found to be associated with presenile cataracts in an inbred Iraqi Jewish family (Pras *et al*., 2002). Another more recent study found that the G154E mutation in MP20 causes autosomal recessive congenital cataracts (Ponnam *et al.,* 2008). In mice, two mutations have been characterized; G15V and C51R (Steele *et al*., 1997; Puk *et al.,* 2011), both resulting in reduced lens size and formation of severe cataracts. These findings indicate that MP20 is conserved and critical for the normal formation of mammalian eye lenses.

Proposed functions of MP20

It has been suggested that MP20 is either involved in the formation of membrane junctions or in membrane fusion. Being a member of the PMP20/EMP/MP20/claudin superfamily, it would not be surprising if MP20 were involved in junction formation, because most members of this superfamily have been implicated in tight junction formation. For instance, PMP22 is a component of tight junctions in the epithelium of the liver and intestine (Notterpek *et al.,* 2001) and is expressed at the blood-brain barrier (Roux *et al.,* 2004). Other family members found in tight junctions in epithelia include Cldn1, Cldn6, Cldn11, Cldn14 and Cldn19 (Van Itallie *et al,* 2006). In support of MP20 forming membrane junctions, MP20 is known to interact with galectin-3, a known cell adhesion modulator (Ervin *et al*., 2005; Gonen *et al.,* 2001), presumably through two glycosylated tryptophans in MP20, Trp 43 and Trp 61 (Ervin *et al*., 2005). In addition, MP20 is stored in intracellular vesicles in young lens fiber cells, and insertion of MP20 into lens fiber plasma membranes coincides with the formation of an extracellular diffusion barrier that restricts passage of material into the lens core (Grey *et al.,* 2003). Taken together, these results support a role for MP20 in junction formation, but do not rule out the possibility of its involvement in membrane fusion.

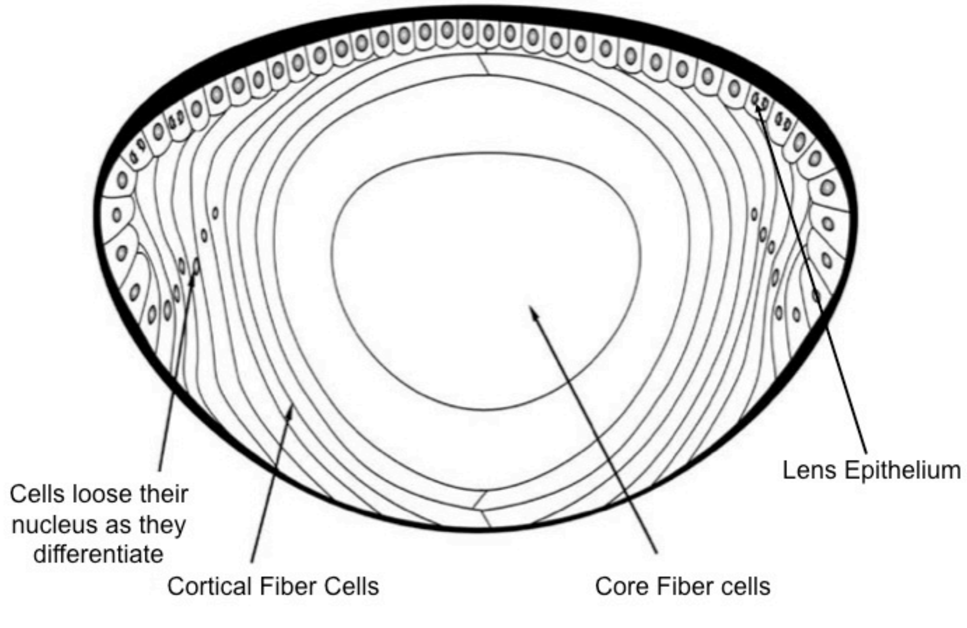


Figure 3.1 Schematic showing the distribution of cell types in the eye lens. Figure adapted from Gupta *et al.,* 2004.

Recent studies demonstrated that MP20 is also required for cell fusion in the lens (Shi *et al.,* 2009). The lens consists of three distinct components, the epithelium, the cortex, and the core (Figure 3.1). The lens cortex contains multiple layers of fiber cells of different ages in which macromolecules are free to diffuse across cells of the same layer but not across different layers. Within the lens core, the fiber cells are completely fused to enable free diffusion of material throughout the entire core (Shi *et al.,* 2009). This organization was described as the stratified lens synctium (Shi *et al.,* 2009). When the MP20 gene was knocked out in mice, cell fusion and free diffusion of material across fiber cells no longer occurred (Figure 3.2), leading to the conclusion that MP20 is necessary for fusion of fiber cells in the lens (Shi *et al.,* 2009). However, the mechanism by which MP20 is involved in cell fusion remains unclear.

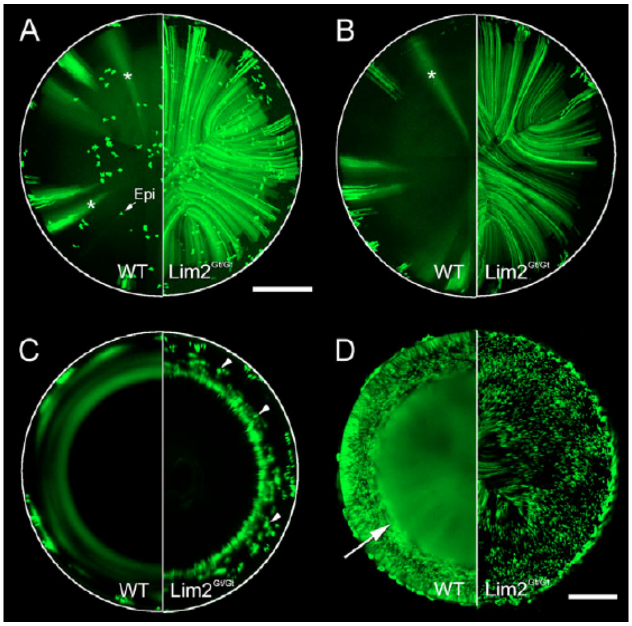


Figure 3.2: Knocking out MP20 prevents the formation of a lens synctium (Shi *et al.,* 2009). Fluorescence imaging of the diffusion of heterologously expressed GFP in the lens of wild-type (left panels) and MP20 knockout mice (right panels). A) Anterior

Figure 3.2 (Continued) views, B) posterior views, C) Equatorial views and D) Center of the lens. Epi: epithelial cells. Scale bars are 500 µm (A-C) and 250 µm (D).

Structure of MP20

No structure for any member of the PMP20/EMP/MP20/claudin superfamily is currently available. Members of the superfamily share a common topology, which features four transmembrane domains and two extracellular loops (Figure 3.3). The first extracellular loop contains a conserved "GLWxxC (8-10 aa) C" motif. In most claudins, the proteins that form tight junctions and paracellular channels (Van Itallie *et al*., 2006), the first extracellular loop was shown to influence the ion selectivity of the paracellular channels, and the second loop was shown to be a receptor for bacterial toxins (Van Itallie *et al*., 2006). While the cytoplasmic C terminus of family members usually contains a PDZ domain that interacts with and signals to other proteins (Van Itallie *et al*., 2006), the C terminus of MP20 lacks a PDZ domain, suggesting that MP20 is regulated or mediates signaling in a way that differs from other family members.

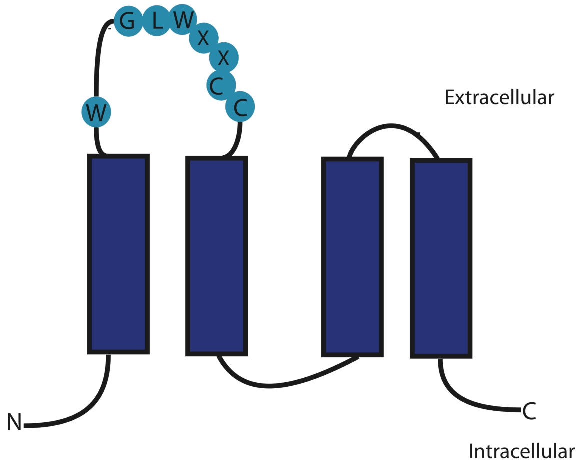


Figure 3.3: Schematic representation of the topology of members of the PMP20/EMP/ MP20/claudin superfamily.

The first 25 N-terminal amino acids are needed for integration of MP20 into the cell membrane (Chen *et al.,* 2003). Furthermore, it was demonstrated that a GFP fusion to the C terminus of MP20 does not interfere with normal plasma membrane localization, whereas a GFP fusion to the N terminus causes MP20 to be distributed throughout the cytoplasm in 293 T-Rex cells (Chen *et al.,* 2003). Thus, proper trafficking of MP20 to the plasma membrane depends on the accessibility of the N terminus. This finding may explain why the G15V and C51R mutations cause severe cataracts in the mouse (Steele *et al.* 2007).

Like other members of the PMP20/EMP/MP20/claudin superfamily, MP20 forms homo-oligomers *in vitro* (Ervin *et al*., 2005; Gonen *et al*., 2008). The largest oligomer detected by cross-linking studies was a hexamer (Jarvis *et al*., 1995). Single-particle EM studies also showed that purified MP20 exists as oligomers of different size, but when the protein was purified in the presence of the metal chelators EDTA and EGTA, MP20 eluted predominantly as a tetramer from a size exclusion chromatography column (Gonen *et al.*, 2008). Using negative-stain EM, MP20 purified in the presence of metal chelators showed predominantly four-fold symmetry (Gonen *et al.*, 2008).

My goal was to work towards determining a high-resolution structure of MP20, which would provide not only clues to the function of this proteins, but could also serve as a template for homology modeling of other members of the PMP20/EMP/MP20/claudin superfamily. In parallel, I was working on setting up an *in vitro* functional assay for MP20 to establish whether MP20 mediates junction formation or membrane fusion.

Results

Purification of MP20 from sheep lenses and reconstitution into proteoliposomes

MP20 was purified from sheep lenses following a previously published protocol (Gonen *et al.,* 2001), with slight modifications. Briefly, lenses were extracted from sheep eyeballs by dissection, and the lens cores were homogenized in 10 mM Tris, pH 8.0. The membranes were stripped by subsequent washes with 4 M urea and 20 mM sodium hydroxide to remove peripheral membrane proteins. The stripped membranes were then finally homogenized in 10 mM Tris, pH 8.0. MP20 was extracted by incubating the membranes with 1% decyl maltoside (DM). The protein was run over a Mono Q anion exchange column and collected in the flow-through. Unlike the previously published protocol, a carboxymethoxyl (CMFF) column was used instead of a hydroxyapatite column to remove the lipids. A CMFF column has been used before to purify MP20 (Jarvis *et al.*, 1995), and it avoids readjustment of the pH, which is required for the use of a hydroxyapatite column. The flow-through from the Mono Q column was added to the CMFF column, and MP20 was eluted with 100 mM NaCl. Finally, MP20 was run over a Superose 12 10/300 GL size exclusion column. The yield of purified MP20 from 50 lens cores ranged from 50 to 60 μg, which was insufficient for structural studies but allowed performing functional assays.

For functional assays, MP20 was reconstituted into lipid vesicles. 30 μl of 0.1 mg/ml MP20 was mixed with 6 μl of 5 mg/ml of dimyristoyl phosphatidylcholine (DMPC) in 1.2% octyl glucoside (OG), giving a lipid-to-protein ratio of 10:1. The volume was increased to 60 μl with dialysis buffer, and the mixture was dialyzed against 100 ml of buffer containing 10 mM Tris, pH 8, 100 mM NaCl, 2 mM EDTA, 2 mM EGTA and 0.01% NaN_3. The buffer was exchanged daily for one week until all detergent was removed and vesicles formed.

Expression and purification of galectin-3

To be able to assess the effects of galectin-3 binding on the function of MP20, His-tagged sheep galectin-3 was expressed in *E. coli* and purified using α-lactose affinity chromatography following an established protocol (Pelletier *et al.*, 2002). The protein was highly pure, and its identity was confirmed by Western blotting with an HRP-conjugated anti-His antibody (Figure 3.4).

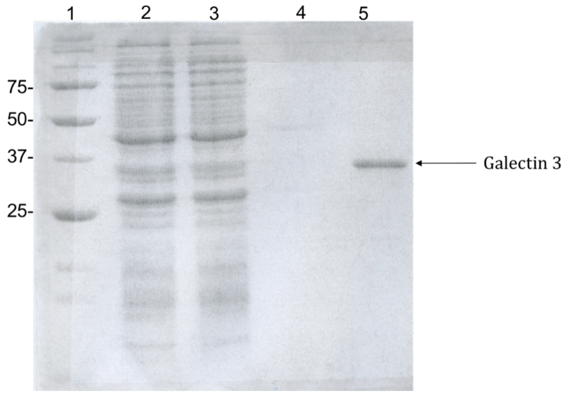


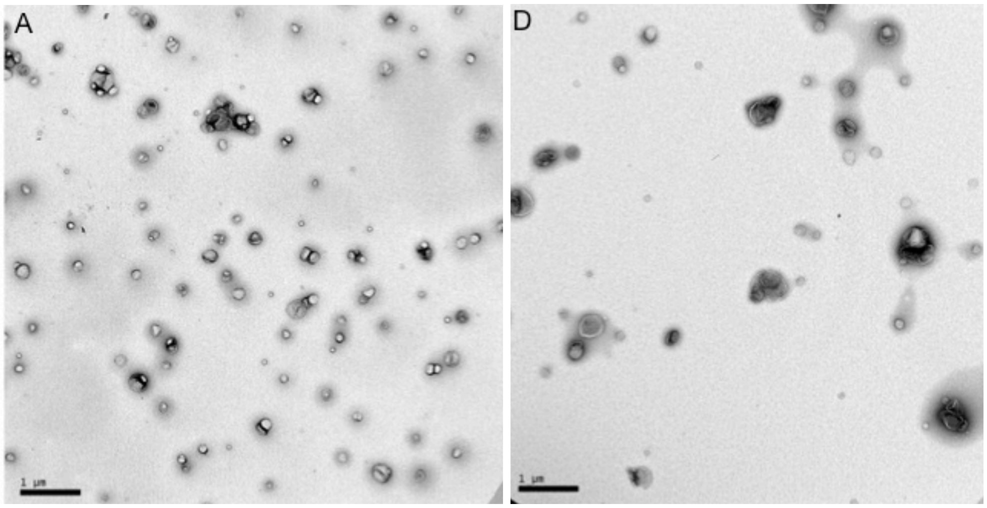
Figure 3.4: Purification of recombinant galectin-3 from *E. coli*. Coomassie blue-stained 15% SDS PAGE gel. Lane 1: molecular weight markers; lane 2: cleared lysate; lane 3: flow-through from α-lactose affinity chromatography; lane 4: wash of α-lactose affinity column with no lactose; lane 5: elution from α-lactose affinity column with 100 mM lactose.

EM-based *in vitro* functional assay for MP20

The MP20 proteoliposomes (Figure 3.5A) were used to test the hypothesis that the interaction of MP20 with galectin-3 induces the formation of membrane junctions. When His-tagged galectin-3 was added to MP20 proteoliposomes, no change could be detected (Figure 3.5B). However, when Ca^{2+} was added (final concentration of 10 mM) to the MP20 proteoliposomes, the size of the vesicles appeared to increase (Figure 3.5C). When MP20 proteoliposomes were pre-incubated with galectin-3 before addition of Ca^{2+}, the vesicles appeared to have a similar size as the control MP20 proteoliposomes (Figure 3.5D). These finding suggested that binding of galectin-3 to MP20 does not induce the formation of membrane junctions, which would have caused the formation of vesicle clusters. Instead, it appeared as if MP20 has a weak, Ca^{2+}-dependent membrane fusion activity, leading to the formation of larger vesicles, and that binding of galectin-3 to

MP20 prevents MP20-mediated vesicle fusion. Lactose is known to bind to galectin-3 and has been shown to compete with galectin-3 binding to MP20 (Gonen *et al.,* 2005). When lactose was added to the MP20 vesicles before the addition of galectin-3 and Ca^{2+}, the size of the MP20 proteoliposomes appeared to increase (Figure 3.5E). This result suggested that inhibition of the MP20 membrane fusion activity by galectin-3 involves the lactose-binding site of galectin-3. This finding is in agreement with a previous result that showed that galectin-3 binds to MP20 through its glycosylations (Gonen *et al.,* 2005). Because the effects were subtle, the diameters of the MP20 vesicle under different conditions were measured and compared (Figure 3.5F). The results supported the impressions obtained by visual inspection of the vesicles in the electron microscope.

Figure 3.5



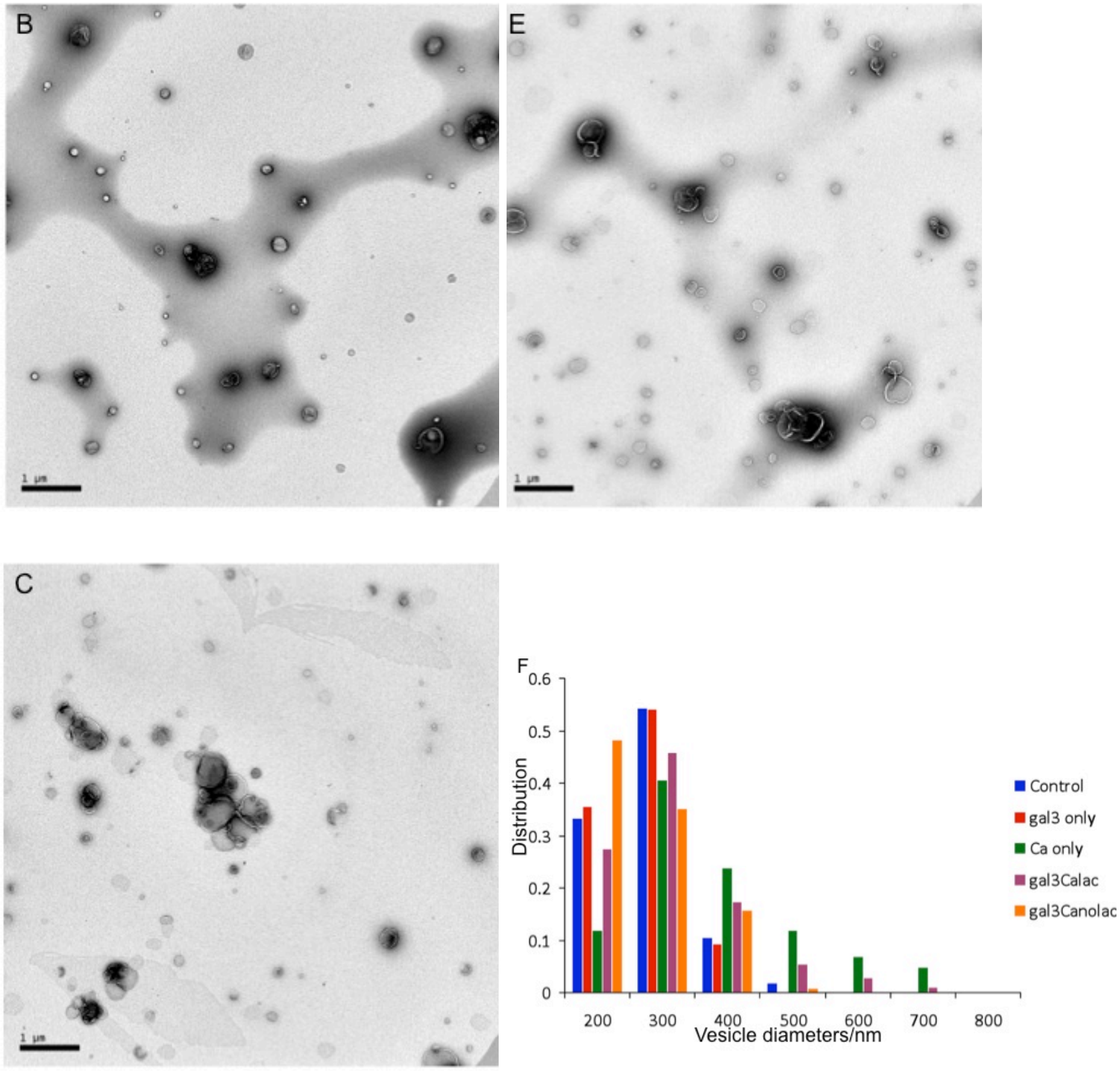


Figure 3.5 (continued): *In vitro* assay suggests that MP20 mediates vesicle fusion in the presence of Ca^{2+} and that galectin-3 inhibits MP20-mediated membrane fusion. A) MP20 proteoliposomes alone, B) MP20 proteoliposomes in the presence of galectin-3, C) MP20 proteoliposomes in the presence of Ca^{2+}, D) MP20 proteoliposomes in the presence of Ca^{2+} and galectin-3, E) MP20 proteoliposomes in the presence of Ca^{2+}, galectin-3 and lactose, F) graph showing the distribution of MP20 proteoliposome diameters under different conditions.

Fluorescence-based vesicle fusion assay

The EM-based *in vitro* assay suggested that MP20 may have a weak membrane fusion activity, but the observed effects were subtle. It was thus necessary to confirm this result using a different method, for which I chose a fluorescence-based fusion assay. Fluorescently labeled lipids (Rhodamine-PE) were incorporated into MP20 proteoliposomes. The proteoliposomes containing purified MP20 were incubated with 40 μM Rhodamine-PE and the labeled vesicles were isolated using a PD-10 desalting column. The fluorescently labeled MP20 proteoliposomes were mixed with an equal volume of unlabeled lipid vesicles of defined size. Vesicle fusion would lead to a dilution of the fluorescently labeled lipids, and the resulting increase in fluorescence due to dequenching would allow the extent of vesicle fusion to be quantified by fluorimetry. The sample was split into two aliquots, and 10 mM Ca^{2+} (final concentration) was added to one of the aliquots, and the fluorescence was measured for 20 minutes. No changes in fluorescence were detected for either sample (Figure 3.6A). The experiment was repeated with fluorescence-labeled MP20 proteoliposomes and non-labeled MP20 proteoliposomes, but again, no change in fluorescence was detected upon addition of Ca^{2+} (Figure 3.6A).

This result contradicted the result obtained with the EM-based fusion assay, which showed an increase in average size of MP20 proteoliposomes upon exposure to 10 mM Ca^{2+}. To follow up on this contradiction in results, the fluorescence-labeled MP20 proteoliposomes were imaged under the electron microscope. Their appearance differed significantly from the unlabeled MP20 proteoliposomes, in that the labeled MP20 proteoliposomes were much larger, and a fraction of the vesicles were broken (Figure 3.6B). This observation suggests that incorporation of fluorescent lipids into MP20 proteoliposomes affected the integrity of the proteoliposomes.

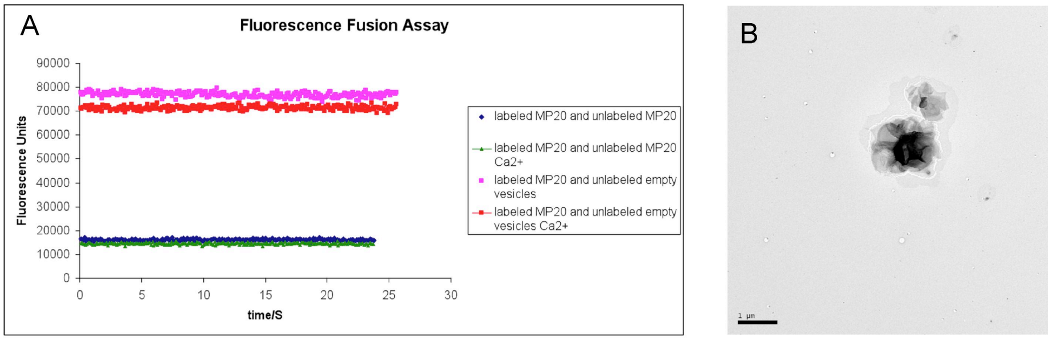


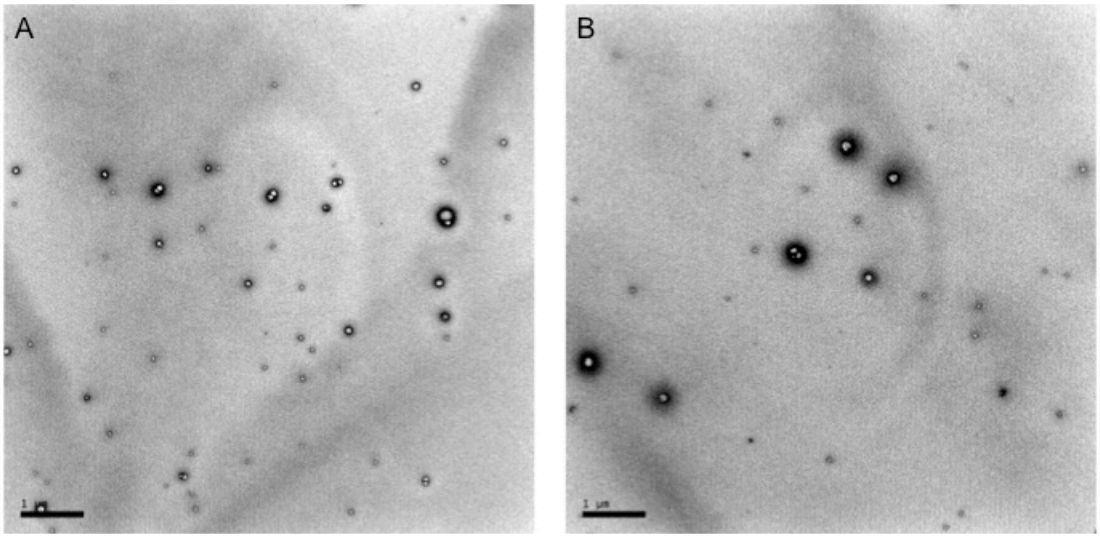
Figure 3.6: Fluorescence-based assay for MP20-mediated vesicle fusion. A) Graph showing fluorescence units over time. Fluorescence of labeled vesicles containing MP20 mixed with unlabeled MP20 proteoliposomes was measured over time in the presence (Green) and absence (Blue) of Ca^{2+}. Fluorescence of labeled vesicles containing MP20 mixed with unlabeled empty lipid vesicles was measured over time in the presence (Red) and absence (Pink) of Ca^{2+}. B) EM image of MP20 proteoliposomes after incorporation of fluorescently labeled lipids, showing that the vesicles are large and clumped together. Scale bar: 1 µm.

EM-based assay with pure lipid vesicles

To better understand the conflicting results of the two vesicle assays, I decided to investigate the effects of Ca^{2+} on pure lipid vesicles. Vesicles formed either with DMPC or a 50:50 mixture of DMPC and lens lipids were prepared by reconstitution. Lens lipids were obtained by collecting the flow-through from the CMFF column used in the purification of MP20 from sheep lenses. The lens lipids were quantified by measuring

UV absorbance at 230 nm and calibrating the value with known quantities of *E. coli* lipids in the same buffer. The lipid vesicles were observed before and after addition of 10 mM Ca^{2+} using negative-stain EM. The size of the DMPC vesicles did not appear to change upon addition of 10 mM Ca^{2+} (Figure 3.7A and B), but the size of the vesicles containing lens lipids appeared to increase upon addition of Ca^{2+} (Figure 3.7C and D). This result suggested that the increase in vesicle size in the EM-based vesicle assay might not have been MP20-mediated membrane fusion but instead was a result of the presence of lens lipids that remained associated with MP20 when the protein was purified from the lens.

Figure 3.7



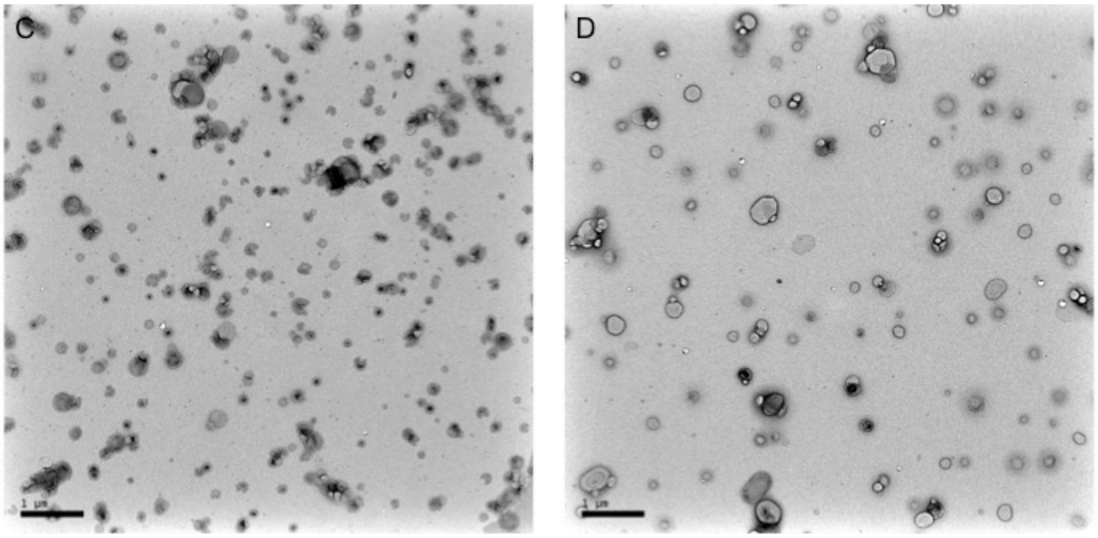


Figure 3.7 (continued): Effect of Ca^{2+} on lipid vesicles. A, B) EM images of DMPC lipid vesicles before (A) and after (B) addition of 10 mM of Ca^{2+}. C, D) EM images of lipid vesicles containing lens lipids before (C) and after (D) addition of 10 mM Ca^{2+}. Scale bars: 1 μm.

Localization of MP20 mutants

Mutations in MP20 cause cataract formation, but the molecular basis for how these mutations result in cataract is not understood. Mutations may directly interfere with the function of MP20, but alternatively mutations could also interfere with proper folding and/or trafficking of MP20 to the plasma membrane.

To determine the localization of the MP20 mutants and to assess the effect of the mutations on membrane targeting, I expressed EGFP-tagged MP20 in HEK 293T cells and imaged the cells by fluorescence light microscopy. HEK 293T cells were grown on 35-mm glass bottom dishes and transfected with pMIG-MP20 using lipofectamine 2000. Mutations in MP20 were introduced using the Quikchange site-directed mutagenesis kit.

Cells were stained with Hoechst 33342 prior to viewing in the Nikon Imaging Center at Harvard Medical School. Images were recorded on a spinning disk confocal microscope. MP20-EGFP fluorescence was excited at 488 nm and collected at 525/50 nm. Hoechst 33342 was excited at 355/50 nm and collected at 420 nm.

As expected, wild-type MP20-EGFP localized to the plasma membrane (Figure 3.8A). The F105V mutant, a mutation that causes presenile cataract (Pras *et al.*, 2002), also localized to the plasma membrane and showed a similar distribution as wild-type MP20 (Figure 3.8B). This result suggests that the F105V mutation does not affect folding or localization of the protein and therefore likely promotes cataractogenesis in human adults by interfering with the activity of MP20. The G154E mutant, a mutation that causes congenital cataract (Ponnam *et al.,* 2008), displayed a different localization pattern in HEK 293T cells (Figure 3.8C). None of the $MP20^{G154E}$ localized to the cell surface membrane but instead appeared trapped in the endoplasmic reticulum. This finding suggests that the mutation disrupts trafficking of MP20 to the plasma membrane. These localization results may explain why the F105V mutation causes late-onset cataracts, whereas the G154E mutation causes early-onset cataracts in humans.

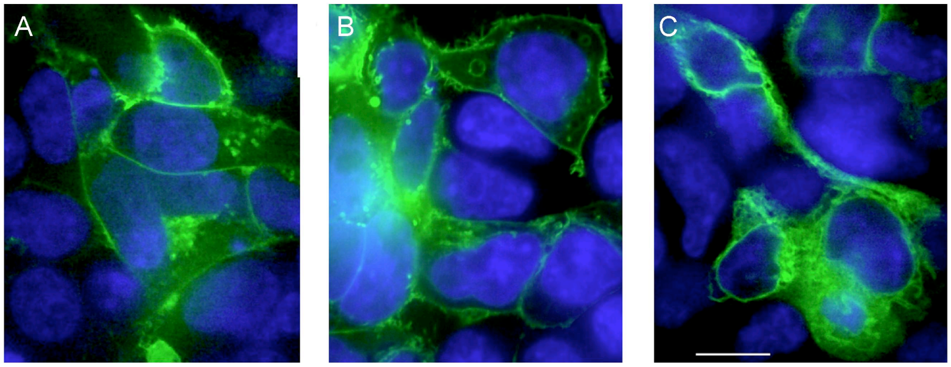


Figure 3.8: Localization of wild-type MP20 and cataract-causing mutants in HEK293T cells. Wild-type MP20 (A) and the F105V mutant (B) localize to the plasma membrane, but not the G154E mutant (C). Scale bar: 10 μm.

Expression and purification of MP20 from insect cells

Recombinant N-terminally His-tagged MP20 was expressed and purified from Sf9 insect cells. Briefly, 2 l of insect cells at a density of 2.0 x 10^6/ml were infected at an MOI of 0.1. The cells were harvested after 96 h, lyzed by sonication, and membranes were isolated by centrifugation. Screening various detergents, including Fos-choline 11, decyl maltoside (DM) and octyl glucoside (OG), showed that only 1% Fos-choline 11 could extract MP20 from the insect cell membranes. Since MP20 purified from sheep lenses behaved best in DM, Fos-choline 11 was exchanged with 0.3% DM on the Ni-NTA affinity column, which was followed by size-exclusion chromatography (Figure 3.9A). Using this procedure, I reproducibly obtained a yield of 0.5 mg from 2 l of insect cell culture. I was also able to reconstitute the recombinant protein into lipid vesicles using the same protocol I used for native protein (Figure 3.11B), and the resulting proteoliposomes were indistinguishable from those obtained with MP20 isolated from sheep lenses (Figure 3.11A).

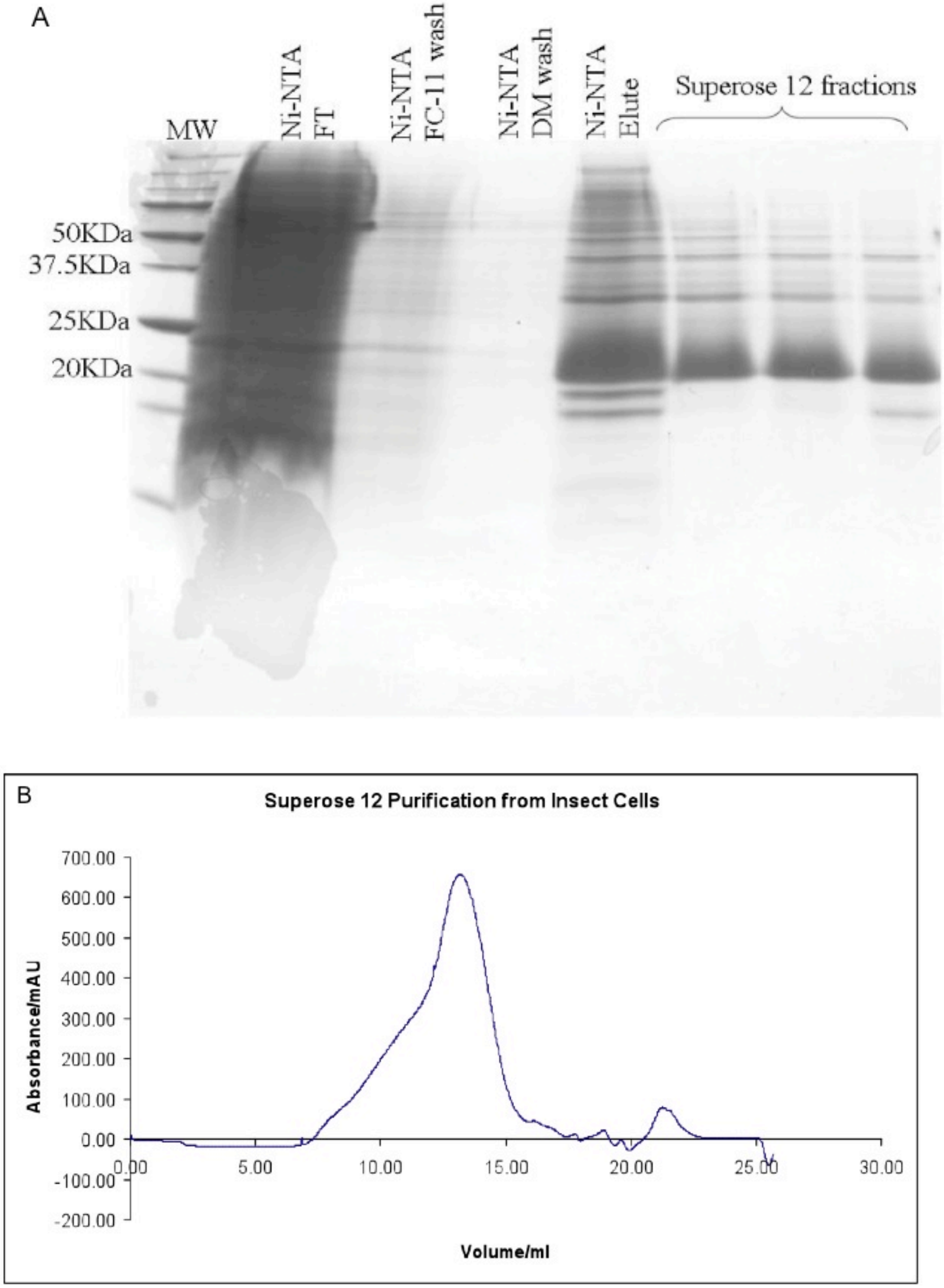


Figure 3.9: Purification of MP20 from insect cells. A) Coomassie blue-stained gel of the MP20 purification from insect cells. B) Gel filtration profile of MP20 after Ni-NTA affinity chromatography.

Expression and purification of MP20 from *Pichia pastoris*

The quantity and quality of recombinant MP20 expressed in Sf9 cells was very encouraging but still somewhat limiting for comprehensive structural studies. Therefore, I attempted to express MP20 in *P. pastoris*. This expression system, if successful, would provide an economical and convenient approach for large-scale production of MP20. For the expression of MP20, a protease-deficient *P. pastoris* strain, SMD1163, was used and transformed with pPICZA-10xHis-MP20. 500 ml of cells was first grown at 30^{o}C for 24 hours in Buffered Glycerol Complex Medium (BMGY), spun down and the cells were added to 2 l of Buffered Methanol Complex Medium (BMMY) at a density of $OD_{600} = 1$. The cells were grown at 30^{o}C, harvested 24 hours later, and lyzed with an EmulsiFlex-C3 microfluidizer (Avestin). The cells were first centrifuged at 15,000 x g for 15 minutes to remove cell debris and then centrifuged at 100,000 x g for 30 minutes to collect the membranes. Subsequently, His-tagged MP20 was purified based on the same protocol used for purification from insect cells. Based on a preliminary trial, the yield obtained from *P. pastoris* was approximately the same as that obtained from insect cells, given the same biomass of cells harvested. In addition, the protein purified from *P. pastoris* appeared to show higher purity based on a Coomassie blue-stained gel (Figure 3.10B) and could also be reconstituted into DMPC vesicles (Figure 3.11C). With these sources for recombinant MP20, I began to set up 2D and 3D crystallization screens.

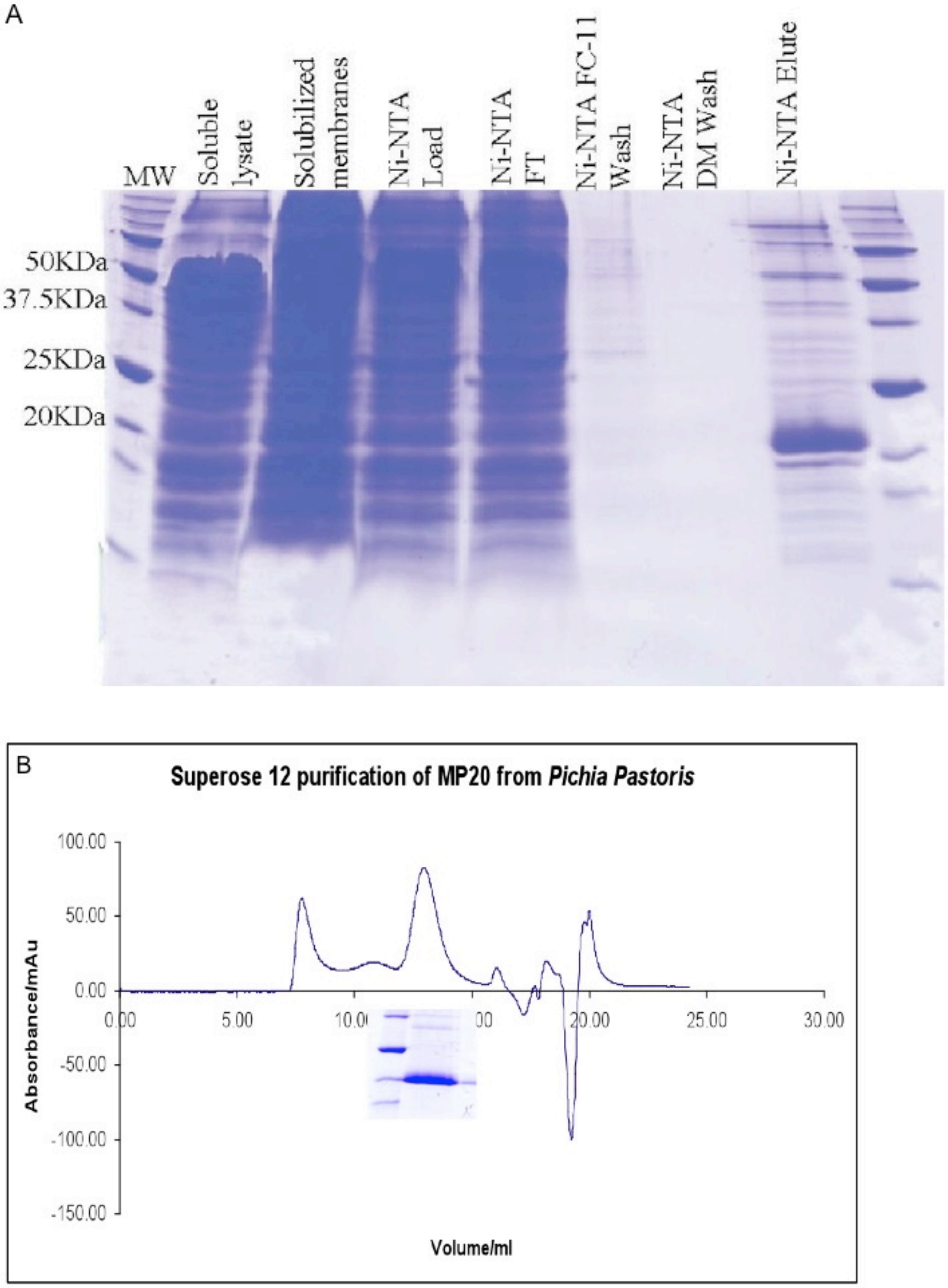


Figure 3.10: Purification of MP20 from *P. pastoris.* A) Coomassie blue-stained gel of MP20 purification from *P. pastoris*. B) Gel filtration of MP20 purified from *P. pastoris.*

Figure 3.10 (Continued) The Coomassie blue-stained gel under the peak at 12.5 ml shows the purity of the isolated MP20.

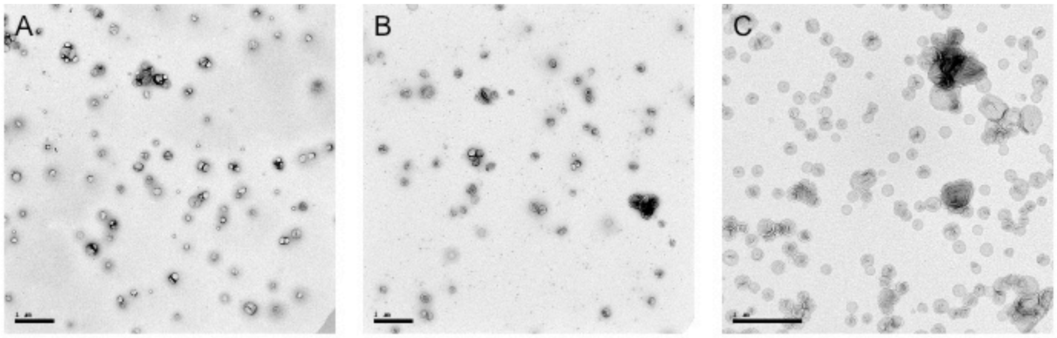


Figure 3.11: Comparison of DMPC vesicles with reconstituted MP20 obtained from different sources. A) Native MP20 isolated from sheep lenses. B) MP20 expressed in and purified from Sf9 insect cells. C) MP20 expressed in and purified from *P. pastoris*. Scale bars: 1 μm.

<u>2D crystallization of MP20</u>

Various reconstitution parameters were tested to produce 2D crystals of MP20, including different lipids (EPL, DMPC, POPC and Sph (refer to captions for Table 3.1 for full name of lipids)), different lipid-to-protein ratios (LPRs) (0.25 - 1), and different pH of the dialysis buffer (Table 3.1). Using 10 mM Tris, pH 8, 100 mM NaCl, 20 mM $MgCl_2$ and 0.01% NaN_3 as the dialysis buffer, MP20 reconstituted into lipid vesicles (Figure 2.12) with almost every lipid except EPL, but no 2D crystalline arrays formed.

Table 3.1

Screen 1					
	LPR	lipid	buffer	protein conc	outcome

Table 3.1 (continued)

1	0.25	DMPC	pH 6	1 mg/ml	aggregates
2	0.5	DMPC	pH 6	1 mg/ml	vesicles
3	0.75	DMPC	pH 6	1 mg/ml	aggregates and vesicles
4	0.25	EPL	pH 6	1 mg/ml	aggregates
5	0.5	EPL	pH 6	1 mg/ml	aggregates
6	0.75	EPL	pH 6	1 mg/ml	aggregates
Screen 2					
1	0.25	DMPC	pH 6	1 mg/ml	aggregates
2	0.5	DMPC	pH 6	1 mg/ml	some badly shaped vesicles
Screen 3					
1	0.5	DMPC	pH 6	1 mg/ml	aggregates & very small vesicles
2	0.75	DMPC	pH 6	1 mg/ml	aggregates & very small vesicles
3	0.5	EPL	pH 6	1 mg/ml	aggregates & very small vesicles
4	0.75	EPL	pH 6	1 mg/ml	aggregates & very small vesicles
5	0.5	DMPC	pH 8, $MgCl_2$	1 mg/ml	many large vesicles
6	0.75	DMPC	pH 8, $MgCl_2$	1 mg/ml	many large vesicles (smaller than above)
Screen 4					
1	0.5	DMPC	pH 8, $MgCl_2$	1 mg/ml	some large vesicles
2	0.75	DMPC	pH 8,$MgCl_2$	1 mg/ml	large vesicles with striations

Table 3.1 (Continued)

3	0.5	POPC	pH 8, MgCl2	1 mg/ml	aggregates
4	0.75	POPC	pH 8, MgCl2	1 mg/ml	aggregates
5	0.5	POPC	pH 6	1 mg/ml	aggregates
6	0.75	POPC	pH 6	1 mg/ml	aggregates
7	0.75	DMPC	pH 8, $MgCl_2$	2 mg/ml	aggregates
8	0.5	DMPC	pH 6	1 mg/ml	tbd
Screen 5					
1	0.3	DMPC	pH 8, $MgCl_2$	1 mg/ml	Large, clustered vesicles
2	0.5	DMPC	pH 8, $MgCl_2$	1 mg/ml	large, clustered vesicles
3	0.75	DMPC	pH 8, $MgCl_2$	1 mg/ml	large, clustered vesicles
4	0.3	POPC	pH 8, $MgCl_2$	1 mg/ml	large, clustered vesicles
5	0.5	POPC	pH 8, $MgCl_2$	1 mg/ml	large, clustered vesicles
6	0.75	POPC	pH 8, $MgCl_2$	1 mg/ml	large, clustered vesicles
7	0.3	Sph	pH 8, $MgCl_2$	1 mg/ml	large, clustered vesicles

8	0.5	Sph	pH 8, $MgCl_2$	1 mg/ml	large, clustered vesicles with striations

Table 3.1 (continued): Outcome of 2D crystallization screens. EPL: *E. coli* polar lipids, POPC: 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine, DMPC: 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine, Sph: sphingomyelin

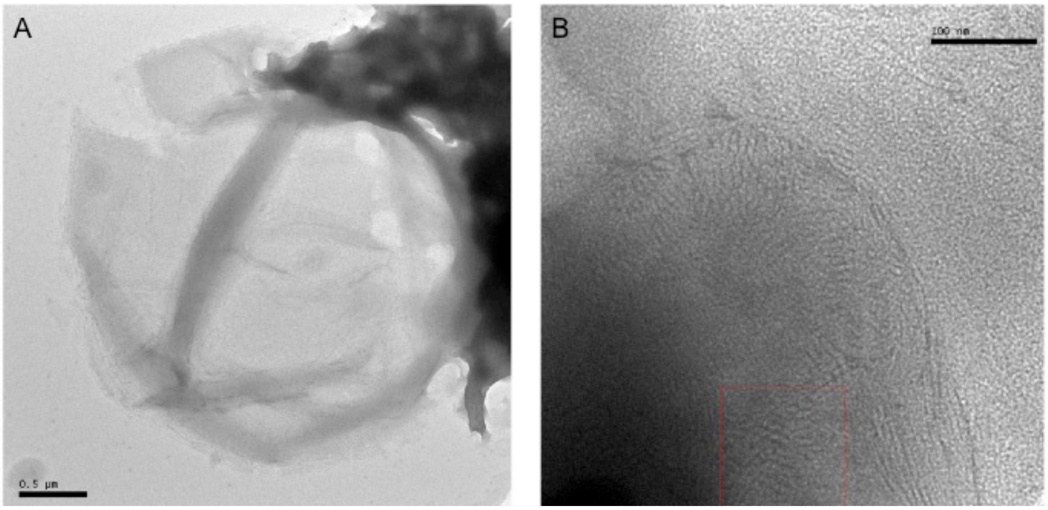


Figure 3.12: Images of vesicles from 2D crystallization trials. A) Low-magnification image of a negatively stained vesicle at 5200x. B) Higher-magnification image of the same vesicle showing striations. Scale bars: 0.5 µm in A) and 100 nm in B).

3D crystallization of MP20

With recombinant MP20 produced in insect cells, I used a Mosquito crystallization robot to set up hanging drop crystallization trials using commercially available screens (MB Class I and II from Qiagen). Crystals were first obtained with 10 mM Tris, pH 8, 100 mM NaCl and 0.3 % DM and 35% PEG 600 as the precipitant. Although these crystals could be reproduced manually (Figure 3.13A), they were very small and not suitable for analysis. More than a month after the conditions were set up, crystals also formed in the same buffer with 30% PEG 600. The crystals were on average 100 μm x 100 μm in size and thus of suitable dimensions for analysis by x-ray diffraction (Figure 3.13B). The same conditions also produced crystals with MP20 purified from *P. pastoris*.

To test whether the crystals were composed of protein or detergent, crystallization screens were set up with DM ranging in concentration from 0.3 mg/ml to 3 mg/ml with 30% PEG 600 as the precipitant. Spherulites similar to the ones obtained with MP20 appeared when the concentration of DM was higher than 1 mg/ml (Figure 3.13D), but the morphology of the crystals obtained with and without MP20 differed slightly. To further study their composition, the crystals were washed three times in mother liquor before being dissolved in Laemmeli buffer. While all three washes contained MP20, the dissolved crystals did not (Figure 3.14). This result suggests that the crystals are formed by the detergent and not MP20.

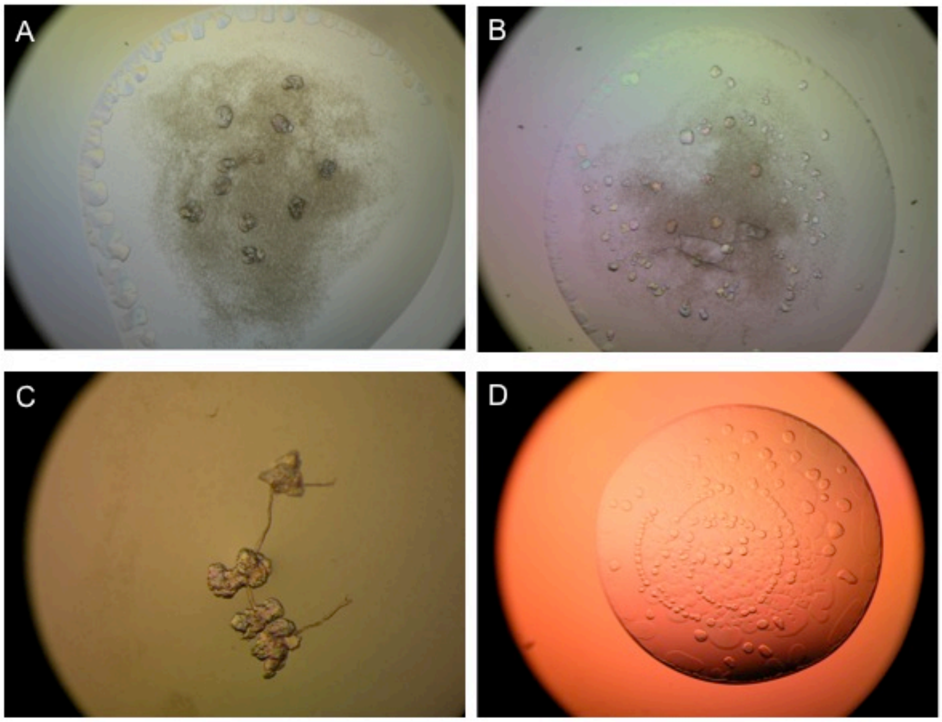


Figure 3.13: 3D crystals. A, B) With MP20 purified from insect cells, small compact disc-like crystals appeared either overnight in the presence of 35% PEG 600 (A) or after a month in the presence of 30% PEG 600 (B). C) With MP20 purified from *P. pastoris*, crystals formed after a month with 30% PEG 600. D) Decyl maltoside at a concentration of 1 mg/ml formed crystals with 30% PEG 600.

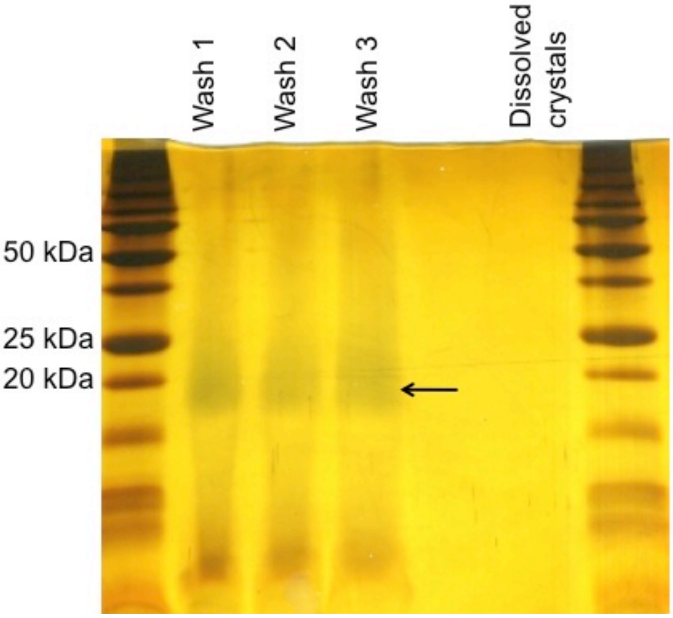
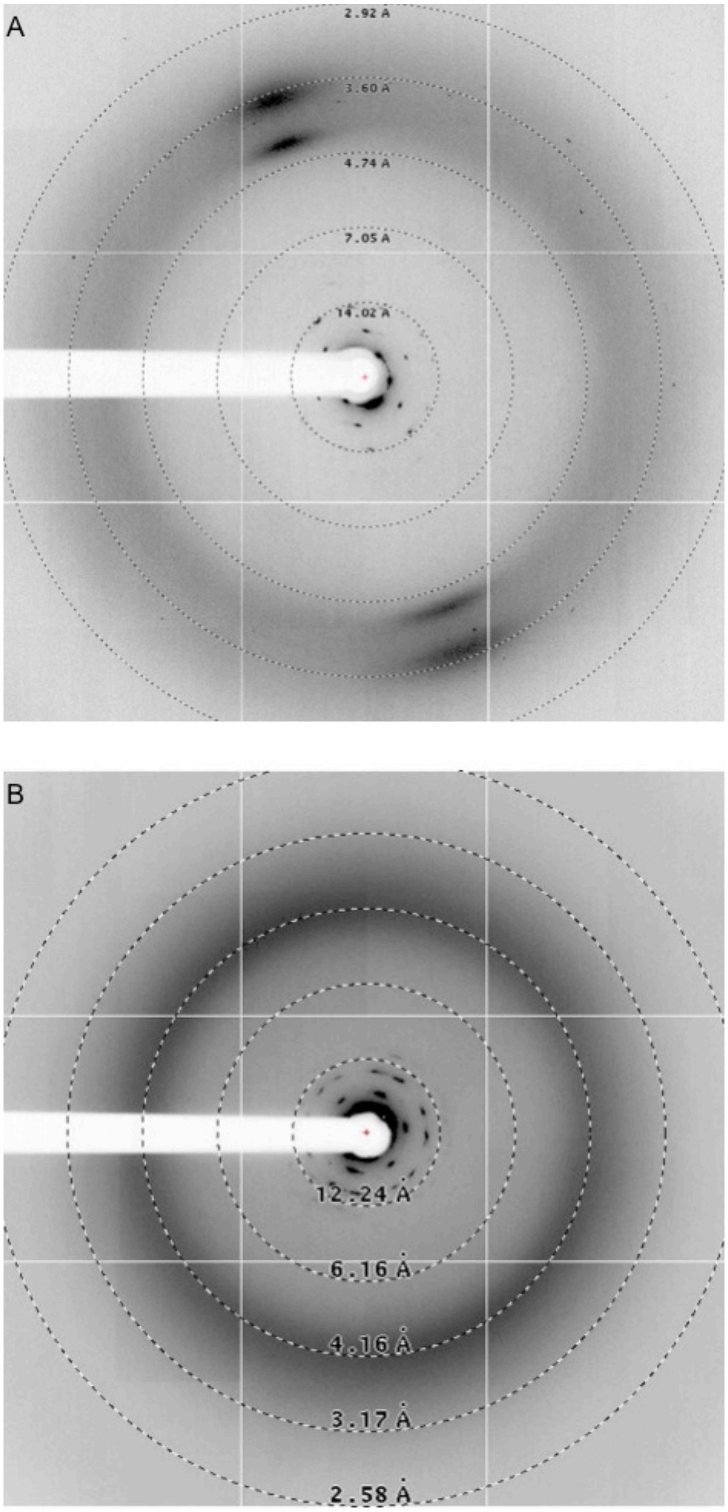


Figure 3.14: Silver-stained SDS-PAGE gel of the 3D crystals. MP20 was only detected in the three washes (arrow), but not in the dissolved crystals.

As further confirmation that the crystals were formed by DM, the crystals were sent to a synchrotron for x-ray diffraction. The diffraction patterns from the crystals obtained with MP20 purified from insect cells (Figure 3.15A) were very similar to those from crystals obtained with MP20 purified *P. pastoris* (Figure 3.15B, C). It was difficult to determine the unit cell size of the crystals but based on the spacings in between the diffraction spots, it was very unlikely that the crystals contained ordered MP20 in detergent micelles.

Figure 3.15



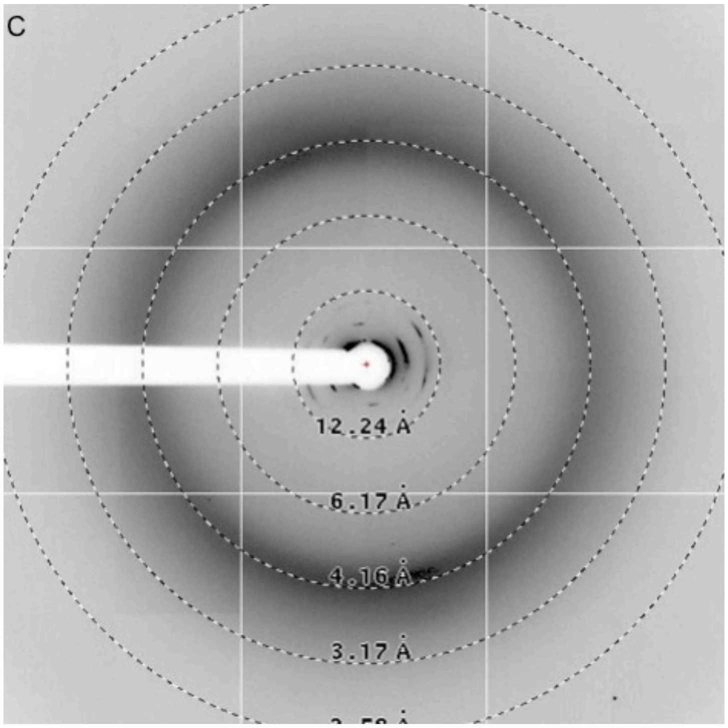


Figure 3.15 (Continued): Diffraction patterns of crystals obtained with recombinant MP20. A) Diffraction patterns of a crystal obtained with MP20 purified from insect cells. B, C) Diffraction patterns of a crystal obtained with MP20 purified from *P. pastoris*. Patterns in B) and C) were taken at 0° and 90°.

Discussion

MP20 remains an enigmatic protein that is recalcitrant to structural and functional characterization. While my efforts to study the function of MP20 *in vitro* did not allow me to assign a function to MP20, they did rule out the possibility that MP20 mediates the formation of membrane junctions through interactions with galectin-3. This hypothesis was based on studies performed by Grey *et al.,* 2003 and Gonen *at al.,* 2007. To test whether interactions of galectin-3 with MP20 induces membrane junction formation, I adapted an *in vitro* vesicle clustering assay that was previously used to show that cleavage of the lens-specific water channel protein aqupaorin-0 induces the formation of membrane junctions (Gonen *et al.,* 2004). However, vesicles with reconstituted MP20 showed no indication of clustering when incubated with galectin-3 (Figure 3.5B), ruling out the possibility that direct interactions of galectin-3 with MP20 are the basis for membrane junction formation.

Instead, MP20 proteoliposomes appeared to grow larger when they were exposed to Ca^{2+} (Figure 3.5C), suggesting that MP20 may mediate membrane fusion in a Ca^{2+}-dependent manner. This result was consistent with recent cell biological studies that implicated MP20 in the hemi-fusion of lens fiber cells (Shi *et al.,* 2009). Furthermore, galectin-3 appeared to inhibit the Ca^{2+}-dependent fusion activity of MP20 (Figure 3.5D). While the increase in size of MP20 proteoliposomes upon incubation with Ca^{2+} was reproducible, the effect was subtle (Figure 3.5F). Attempts to detect vesicle fusion by using a fluorescence-based assay did not show any indication that MP20 fuses membranes in a Ca^{2+}-dependent way (Figure 3.6). Furthermore, exposure of vesicles formed by lens lipids to Ca^{2+} appeared to grow larger (Figure 3.7C and D), indicating that lens lipids have an intrinsic tendency to promote membrane fusion. The most likely interpretation of the results obtained with the MP20 proteoliposomes is thus that vesicle fusion was facilitated by lens lipids that co-purified with MP20 isolated from the lens. The inhibitory effect of galectin-3 on the Ca^{2+}-dependent fusion of MP20 proteoliposomes may then simply be the result of the bound galectin-3 preventing vesicles from coming close enough together to allow fusion. In conclusion, my experiments provided no evidence supporting the notion that MP20 is involved in the fusion of lens fiber cells as suggested by Shi *et al.,* 2009.

The localization of MP20 in HEK 293T cells provided first clues as to how mutations in MP20 may cause cataracts. As previously shown (Chen *et al.,* 2003), wild-type MP20 expressed in HEK 293T cells localized to the plasma membrane (Figure 3.8A). MP20 carrying the F105V mutation, which causes late-onset cataracts in humans (Pras *et al.,* 2002), also localized to the plasma membrane (Figure 3.8B). In contrast, MP20 carrying the G154E mutation, which causes congenital cataracts in humans (Ponnam *et al.,* 2008), showed a predominantly an ER localization, suggesting that the mutation interferes with the folding of MP20 and causes the protein to be retained in the ER (Figure 3.8C). Hence, the G154E mutation is likely to cause early-onset cataract by interfering with correct trafficking of MP20, whereas the F105V mutation likely causes late-onset cataract by interfering with the as yet unclear biological activity of MP20.

2D crystals of MP20 (Gonen *et al.,* 2008) as well as 3D crystals (Tamir Gonen, unpublished results) have previously been obtained but have proved difficult to reproduce. Efforts to reproduce these crystals have also been significantly hampered by the small amount of MP20 that can be purified from sheep lenses. The successful expression of MP20 both in insect cells (Figure 3.9) and *P. pastoris* (Figure 3.10) will greatly facilitate future efforts to grow high-quality crystals of MP20. Initial trials to reproduce the 2D crystals were unsuccessful (Figure 3.12), but 3D crystallization screens did yield crystals (Figure 3.13). Analysis of these crystals by x-ray diffraction (Figure

3.15) and SDS-PAGE (Figure 3.14) suggests, however, that they were most likely formed by DM, the detergent used to purify MP20. In conclusion, more effort will be needed to produce 2D or 3D crystals that will allow structure determination of MP20. Such efforts will include developing better purification strategies, screening of MP20 homologs from different species as well as different expression constructs, adding lipids during crystallization conditions, and using antibodies to generate larger surfaces for the formation of crystal contacts.

Materials and Methods

Purification of MP20 from sheep lenses

Sheep eyeballs were purchased from Wolverine Packing and dissected to retrieve the lenses. The cores were isolated from the lenses and homogenized in 10 mM Tris, pH 8.0. The membranes were first stripped by subsequent washes with 4 M urea and 20 mM sodium hydroxide to remove peripheral membrane proteins and then homogenized in 10 mM Tris, pH 8.0 (Gonen *et al.,* 2001).

The membranes were incubated with 1% decyl maltoside (DM) (Anatrace) for 1 hour at room temperature and then centrifuged at 100,000 x g to remove insoluble components. The supernatant was applied to a Mono Q anion exchange column (GE Healthcare) and the flow-through containing MP20 was incubated with carboxymethoxyl (CMFF) resin (GE Healthcare) in 10 mM Tris pH 8 and 0.3% DM. MP20 was eluted from the CMFF resin with 100 mM NaCl in the same buffer, and run over a Superose 12 10/300 GL size exclusion column. The buffer used for size exclusion chromatography was 10 mM Tris, pH 8, 100 mM NaCl and 0.3% DM. The yield obtained from 50 lenses ranged from 50 to 60 μg, as quantified by using the BCA assay (Piercenet).

Expression and purification of galectin-3

His-tagged ovine galectin-3 was expressed in *E. coli* and purified using α-lactose affinity chromatography as described (Pelletier *et al.*, 2002). Briefly, ovine galectin-3 was cloned into a pET15 vector, which adds an N-terminal His tag. 2 l of BL21DE3 (pLys) (Invitrogen) were grown at 37°C, and at an $OD_{600} = 0.5$, the cells were induced with 0.1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) (Sigma Aldrich). The cells were harvested 3 hours post induction, lysed by sonication, and centrifuged at 100,000 x g for 15 minutes at 4°C. The supernatant was applied to an α-lactose affinity column, and galectin-3 was eluted with 100 mM lactose. The purity of galectin-3 was ascertained by SDS-PAGE, and the identity of the protein was confirmed by Western blotting with an HRP-conjugated anti-His antibody (Qiagen). For use in functional assays, 100 μl of galectin-3 at 0.3 mg/ml was dialyzed against 100 ml of buffer containing 20 mM Tris, pH 7.5, and 150 mM NaCl.

Preparation of lipid vesicles and MP20 proteoliposomes

Lens lipids were obtained from the flow-through of the CMFF column used in the purification of MP20 from sheep lenses. The concentration of the lens lipids was quantified by measuring UV absorbance at 230 nm and comparing the value with those of known quantities of *E. coli* lipids in the same buffer. Vesicles were formed by dialyzing

1 mg/ml of lipid against buffer containing 10 mM Tris, pH 8, 100 mM NaCl, 2 mM EDTA, 2 mM EGTA and 0.01% NaN_3 with daily buffer exchanges for one week.

MP20 proteoliposomes were obtained by mixing 1 mg/ml of dimyristoyl phosphatidylcholine (DMPC) with 0.1 mg/ml of MP20 to a total volume of 60 μl. DMPC was diluted from a stock solution prepared with 1.2% octyl glucoside (OG). The mixture was dialyzed as described above.

EM-based vesicle assay

The vesicle assay was set up with a total volume of 20 μl. 5 μl of MP20 proteoliposomes were added to five different Eppendorf tubes at room temperature. Galectin-3 in 10 mM Tris, pH 8, and 100 mM NaCl was added to two tubes, and galectin-3 in the same buffer but containing 100 mM lactose was added to another tube. After 30 min, 10 mM of Ca^{2+} was added to the tubes containing galectin-3 and to an additional tube of vesicles. After another 30 min, samples were prepared by negative staining and imaged on a CM10 electron microscope. The diameters of the vesicles were measured using Image J (Collins, 2007).

Electron microscopy

Samples were applied to glow-discharged, continuous-carbon-coated grids. After washing twice with water, grids were stained with two drops of 0.75% uranyl formate. Grids were inspected with a Philips CM10 electron microscope equipped with a tungsten filament and operated at an acceleration voltage of 100 kV. Images were recorded on a Gatan 1K x 1K CCD camera at a magnification of 2950 x.

Fluorescence-based vesicle assay

500 μl of MP20 proteoliposomes were incubated with 40 μM of Rhodamine-PE (Avanti Polar Lipids) for 30 minutes at 37°C and then for 2 h at room temperature. Labeled proteoliposomes were isolated using a PD-10 desalting column (GE Healthcare) and imaged using negative-stain EM. In parallel, empty lipid vesicles of defined size were prepared according to Schmidt *et al.* (2010). Briefly, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE) (Avanti Polar Lipids), and 1-cholesterol (Sigma) were dissolved in chloroform, mixed at a 1:1:1 molar ratio, and dried under high vacuum for 4 h. The lipid film was resuspended in 20 mM triethanolamine, pH 8.0, 100 mM NaCl and subjected to five cycles of freeze-thawing, followed by 10 cycles of extrusion through a 0.4 μm polycarbonate filter membranes (Whatman).

Vesicle fusion was quantified using a fluorimeter (Photon Technology International). A mixture of equal volumes of labeled MP20 proteoliposomes and unlabeled lipid vesicles were split into two aliquots, and 10 mM Ca^{2+} was added to one of the two aliquots. The fluorescence of both samples was measured for 20 minutes following addition of Ca^{2+}. The same experiment was repeated by misixing fluorescence-labeled with non-labeled MP20 proteoliposomes.

Localization of wild-type and mutant MP20 in HEK 293T cells

HEK 293T cells were cultured at 37°C with 5% CO_2 in Dulbecco's modified Eagle's medium containing 10% heat-inactivated fetal bovine serum. Cells grown on 35-mm glass bottom dishes (MatTek Corporation) were transfected with pMIG-MP20 using lipofectamine 2000 (Invitrogen). Mutations in MP20 were introduced using site-directed mutagenesis (Strategene). Cells were stained with Hoechst 33342 (Invitrogen) prior to imaging in the Nikon Imaging Center at Harvard Medical School. Images were recorded on a Yokogawa spinning disk confocal on a Nikon Ti inverted microscope equipped with a 100x Plan Apo NA 1.4 objective lens and the Perfect Focus System for continuous maintenance of focus. MP20-EGFP fluorescence was excited at 488 nm and collected at 525/50 nm. Hoechst 33342 was excited at 355/50 nm and collected at 420 nm.

Expression and purification of MP20 from insect cells

Recombinant N-terminally His-tagged MP20 was expressed and purified from Sf9 insect cells. Briefly, 2 l of insect cells at a density of 2.0 x 10^6/ml were infected at an MOI of 0.1. The cells were harvested after 96 hours, lysed by sonication, and membranes were isolated by centrifugation at 100,000 x g for 30 minutes at 4^oC. The membranes were solubilized with 1% Fos-choline 11 in 20 mM HEPES, pH 7.4, 300 mM NaCl, 20 mM imidazole at 4^oC, and centrifuged at 100,000 x g for 1 h at 4^oC. The supernatant was incubated for 1 h with Ni-NTA agarose (Qiagen). The resin was first washed with 0.3% Fos-choline 11 and then with 0.3% DM, and the protein was eluted with 500 mM imidazole in 20 mM HEPES, pH 7.4, 300 mM NaCl, 0.3% DM. The eluted protein was further purified on a Superose 12 13/300 column (GE Healthcare) in buffer containing 10 mM Tris, pH 8, 100 mM NaCl and 0.3 % DM.

Expression and purification of MP20 from *Pichia pastoris*

Recombinant N-terminally His-tagged MP20 was expressed in and purified from *P. pastoris* as according to the Invitrogen handbook for *P. pastoris.* For the expression of MP20, a protease-deficient *P. pastoris* strain, SMD1163, was transformed with pPICZA-10xHis-MP20. 500 ml of cells containing the construct of interest was first grown at 30°C for 24 hours in Buffered Glycerol Complex Medium (BMGY), spun down and the cells were added to 2 l of Buffered Methanol Complex Medium (BMMY) at a density of OD_{600} = 1. The cells were grown at 30°C, harvested 24 hours later, and lyzed with an EmulsiFlex-C3 microfluidizer (Avestin). The cells were first centrifuged at 15,000 x g for 15 minutes to remove cell debris and then centrifuged at 100,000 x g for 30 minutes to collect the membranes. Membranes were solubilized and MP20 purified as described for MP20 expressed in insect cells.

2D crystallization

1 mg/ml of MP20 (final concentration) was mixed with 0.2 to 0.75 mg/ml lipids to obtain lipid-to-protein ratios in the range of 0.2 to 0.75 (Table 3.1). The mixtures were placed in dialysis buttons and dialyzed against different buffers (Table 3.1). The dialysis buffers were exchanged daily for a week. The outcome of the reconstitution trials was assessed by negative-stain EM.

3D crystallization

Hanging drop crystallization trials using commercially available screens (MB Class I and II from Qiagen) were set up with MP20 at a concentration of 2.4 mg/ml. Using the Mosquito crystallization robot, screens were set up with 0.1 μl drops at a 1:1 ratio of protein and precipitant over 80 μl of reservoir solution. The screens were incubated at room temperature. Crystals appeared overnight with buffer containing 10 mM Tris, pH 8, 100 mM NaCl, 0.3% DM and 35% PEG 600 as the precipitant. Crystals were reproduced with hanging drop screens set up manually at room temperature with 1:1 volume ratio of protein and precipitant, and varying the PEG 600 from 25% to 35%. The largest crystals, ~ 100 μm x 100 μm, formed with 30% PEG 600. These crystals were sent to the synchrotron at Argonne National Labs for analysis by x-ray diffraction. To verify the identity of the crystals, crystals were washed three times in mother liquor (30% PEG600 in milliQ water) before being dissolved in Laemmli buffer and analyzed by SDS-PAGE.

Chapter 4: Monolayer purification and Affinity Grid for single-particle electron microscopy

Introduction

Single-particle EM is a powerful approach to obtain low- to medium-resolution structural information of macromolecules, and in favorable cases can even produce density maps at near-atomic resolution (Cheng and Walz, 2009; Grigorieff and Harrison, 2011; Zhou, 2008). The technique requires that the macromolecules are prepared for the vacuum in the electron microscope, which is done by adsorbing the molecules to an EM grid and embedding them either in negative stain or vitrified ice (Cheng and Walz, 2008). The specimens are then imaged using low-dose procedures, and the projection images of individual molecules in different orientations are combined computationally to calculate a 3D reconstruction of the imaged molecule (Cheng and Walz, 2009; Grigorieff and Harrison, 2011; Zhou, 2008). Over the past few years, advances in the design of electron microscopes, automation of data collection, new algorithms in image processing, and the increasing power of computer processors have made collection and processing of single-particle EM data very efficient (Cheng and Walz, 2008). Indeed, the technological improvements in the methodology not only increased the number of low- and intermediate-resolution structures that have been determined but also allowed determination of near-atomic resolution density maps for several viruses and virus-like particles (Grigorieff and Harrison, 2011).

Monolayer crystallization

Two-dimensional (2D) crystals are a different kind of specimen often studied by EM, using the electron crystallography approach (Raunser and Walz, 2009). 2D crystals are typically formed by reconstituting membrane proteins with lipids at a low lipid-to-protein ratio (Lévy *et al.,* 1999). However, 2D crystallization has also been applied to soluble proteins, taking advantage of lipid monolayers. The application of lipid monolayers to form 2D protein arrays was first introduced by Uzgris and Kornberg (1983). When lipids dissolved in an organic solvent are placed over an aqueous buffer, the hydrophobic tails of the lipids will steer away from the buffer and thus crowd together to form a layer while the hydrophilic head groups of the lipids orient themselves towards the aqueous buffer. This process leads to the spontaneous formation of a lipid monolayer at the air/water interface. Proteins in the aqueous buffer can interact with the lipids and are thus concentrated on the monolayer. The fluidity of the lipid monolayer enables the bound proteins to laterally diffuse, and interactions between proteins can lead to the formation of 2D protein crystals.

The formation of 2D arrays of proteins on lipid monolayers can occur either through non-specific electrostatic interaction between the protein and the lipid head groups or through specific protein-ligand interactions. In the latter case, lipids containing unique ligand head groups can be synthesized to recruit proteins that are engineered with a tag with strong affinity for the ligand. For example, lipids with (N-(5-amino-1-carboxypentyl)iminodiacetic acid)succinyl] nickel salt (Ni-NTA) head groups have been successfully used to form 2D arrays of a wide variety of His-tagged proteins, including cholera toxin (Kornberg and Ribi, 1987), RNA polymerase II (Darst *et al.,* 1988), HIV reverse transcriptase (Kubalek *et al.,* 1994), MHC (Celia *et al.*, 1999), the 20S proteasome (Thess *et al.,* 2002), and proteinase (Plisson *et al.*, 2003). In the case of the 20S proteasome, complexes that were His-tagged on the side of the barrel-shaped complex displayed exclusively side-on views, whereas proteasomes His-tagged at their ends displayed end-on views, demonstrating specific recruitment of the His tag by the Ni-NTA lipid.

Monolayer purification

Compared with x-ray crystallography and NMR, structural studies by single-particle EM require a much smaller quantity of sample, making proteins and macromolecular complexes that cannot be obtained in large amounts amenable to structural analysis. Currently, however, single-particle EM still requires a purified sample. Biochemical purification of a complex can be a long and arduous process, and producing a suitable sample for EM imaging has become the rate-limiting step in single-particle EM. To address this issue, monolayer purification has been conceived as a tool to simplify the protein purification process and to combine it with EM specimen preparation into a single, fast and convenient step. Monolayer purification is based on a lipid monolayer that contains lipids whose head groups have been functionalized with a Ni-NTA group. If such a monolayer is cast over a solution that contains a His-tagged version of the target protein, only the target protein should specifically bind to the Ni-NTA lipids, making it possible – at least in theory – to isolate the target protein from an impure protein solution or even cell extracts. The lipid monolayer can then simply be picked up with a carbon-coated EM grid and prepared by negative staining or vitrification for EM imaging. This procedure is outlined in Figure 4.1.

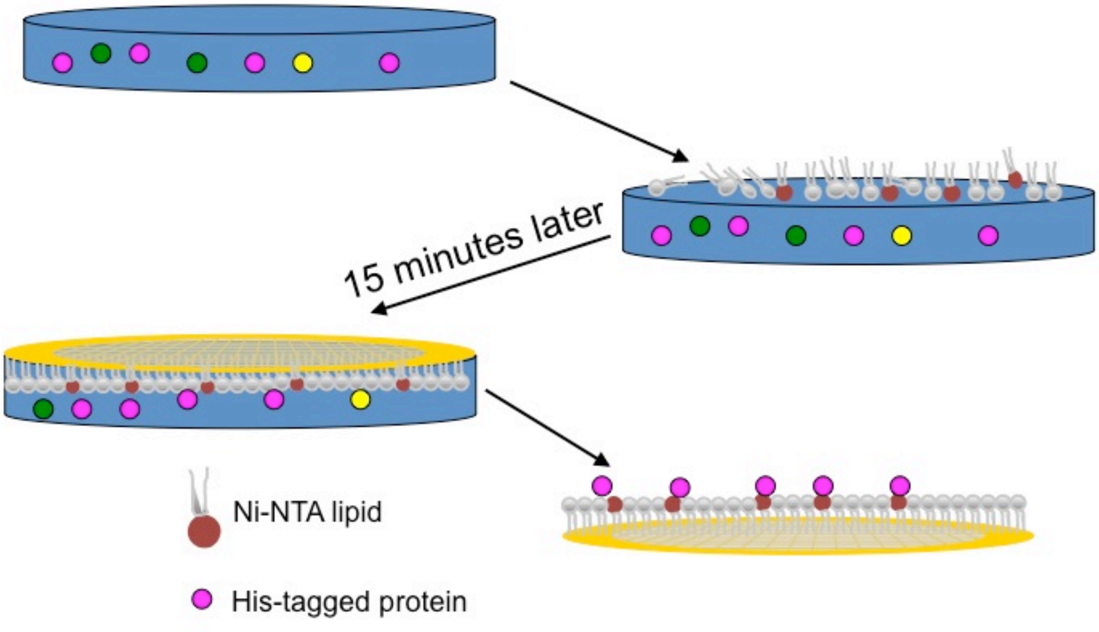


Figure 4.1 Schematic for monolayer purification. Adapted from Kelly *et al.*, 2008a.

The idea of monolayer purification was explored by Dr. Deborah Kelly, a former postdoctoral fellow in the Walz laboratory. The first experiment was aimed at establishing the specific recruitment of a target complex to the Ni-NTA lipids in the monolayer. Indeed, increasing concentrations of Ni-NTA lipid in the monolayer led to increasingly efficient recruitment of His-tagged transferrin (Tf)–transferrin receptor (TfR) complex, whereas a lipid monolayer containing no Ni-NTA lipid did not recruit any Tf-TfR complex. The experiments also established that a Ni-NTA lipid concentration of 2% was ideal for the preparation of negatively stained specimens, whereas preparation of vitrified specimens required a higher Ni-NTA lipid concentration of 20%. This experiment was followed up by adding purified His-tagged Tf-TfR complex to growth medium and various cell extracts and recruiting it to Ni-NTA lipid-containing monolayers. While Tf-TfR complex was recruited, images of the samples also revealed non-specifically bound proteins. This background could be eliminated by addition of imidazole (up to a final concentration of 50 mM) to the cell extracts prior to casting the lipid monolayer. To further evaluate the specificity of monolayer purification, proteins adsorbed to lipid monolayers were eluted with a high concentration of imidazole (300 mM) and analyzed by SDS PAGE, Western blotting, and mass spectrometry. The results established that only Tf and TfR were present in the monolayer-purified samples, demonstrating exceptional specificity of the method. His-tagged Tf-TfR complex prepared by monolayer purification from insect cell extract was also vitrified and imaged by cryo-EM. The images allowed calculation of a 3D map at 20 Å resolution, demonstrating that monolayer purification is compatible with structure determination by cryo-EM. After establishing the basics of the monolayer purification method, it was applied to prepare ribosomes from *E. coli* by expressing a His-tagged version of the human homolog of Rpl3. After preparing a cell extract, monolayer purification was used to isolate the His-tagged complexes. Mass spectrometry analysis of the proteins recruited to the lipid monolayer identified exclusively ribosomal proteins, and showed that samples purified by monolayer purification – unlike complexes purified by conventional Ni-NTA affinity purification – were extremely pure. Finally, cryo-EM of vitrified monolayer purification samples allowed calculation of a 3D map for the 50S ribosomal subunit at 22 Å resolution. These results were published in Kelly *et al.* (2008a).

Affinity Grid

Monolayer purification requires that a lipid monolayer is cast over every sample to be analyzed. To simplify the use of monolayer purification, the Affinity Grid was developed. The Affinity Grid consists of an EM grid (covered by a continuous or holey carbon film) with a pre-deposited Ni-NTA lipid-containing monolayer. Preparing an Affinity Grid sample thus only requires the application of a drop of solution containing the His-tagged target protein. After a short incubation, the Affinity Grid can be washed and prepared by negative staining or vitrification for subsequent EM imaging. The Affinity Grid was tested with His-tagged Tf-TfR complex and ribosomes. The results showed that the Affinity Grid not only worked and was easier to use than monolayer purification, but that it had additional advantages. Sample preparation with the Affinity Grid appeared to be milder than monolayer purification, because EM images of vitrified ribosomes recruited to Affinity Grids revealed that many of them were still associated with RNA, which was never observed with ribosomes prepared by monolayer purification. Moreover, while complexes prepared by monolayer purification tended to cluster, no clustering was observed with complexes prepared on an Affinity Grid. Finally, adsorption of the lipid monolayer to the carbon film of an EM grid also appeared to make it less sensitive to the disruptive effects of detergents and glycerol. It was thus possible to use the Affinity Grid to prepare aquaporin-9 from solubilized membranes of insect cells expressing a His-tagged version of this membrane protein. These results were published in Kelly *et al.* (2008b).

Another advantage of the Affinity Grid is the possibility to assemble complexes directly on the grid. This strategy was demonstrated for the interaction of glycoprotein 1 (GP1) from the Machupo virus envelope with the Tf-TfR complex. The Affinity Grid was incubated with a His-tagged 19-kDa domain of Machupo GP1, which then recruited untagged Tf-TfR complex. Images of vitrified samples showed Tf-TfR complex when the Affinity Grid was pre-incubated with His-tagged GP1, but no Tf-TfR complex was present without pre-incubation of the Affinity Grid with His-tagged GP1, implicating that the Tf-TfR complex was recruited to the Affinity Grid through its interaction with GP1. This result was published in a review (Kelly *et al.*, 2010a).

Adaptation of the Affinity Grid to non-His-tagged proteins

Taking advantage of the possibility to assemble complexes on the Affinity Grid, a strategy was developed to adapt the Affinity Grid to the isolation of proteins that either do not carry a tag or carry a tag different from a His tag. The strategy is based on His-tagged protein A, which is bound to the Affinity Grid covered with a Ni-NTA lipid-containing monolayer. The protein A-decorated Affinity Grid can then be used to immobilize an antibody against the target protein, which can thus be specifically recruited from a protein mixture (Figure 4.2).

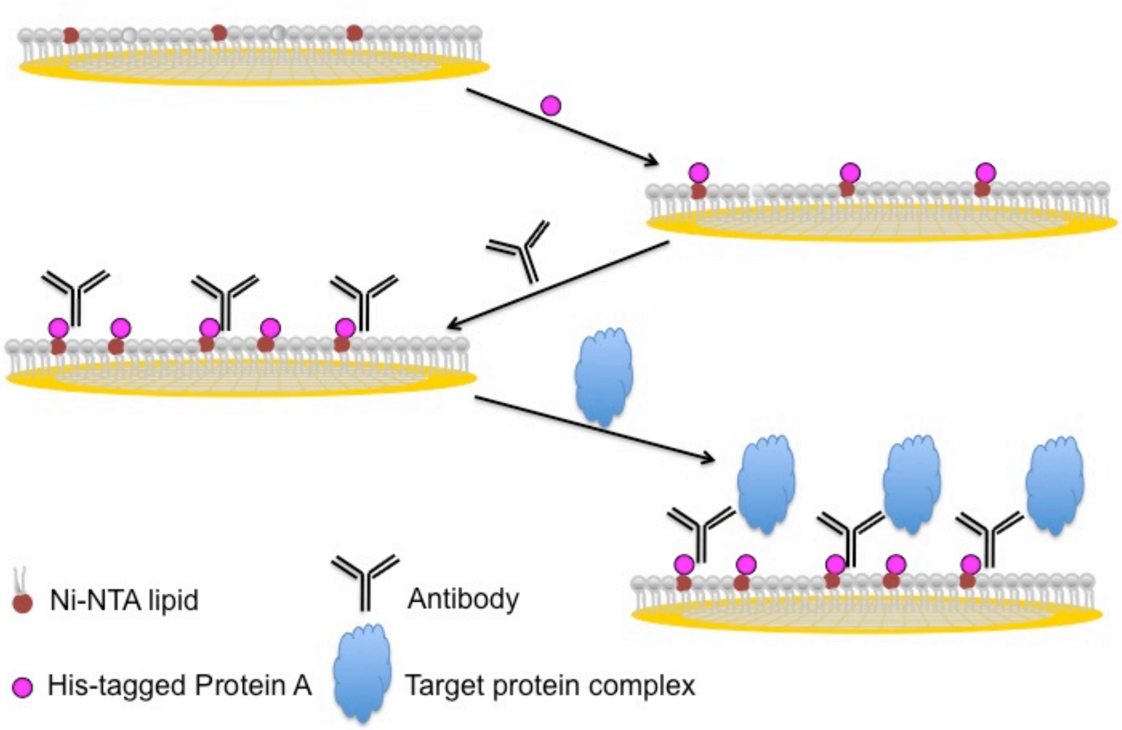


Figure 4.2: Schematic demonstrating the use of His-tagged protein A and antibody to recruit non-His-tagged complexes to an Affinity Grid. Adapted from Kelly *et al.*, 2010b.

The protein A/antibody strategy was tested by using Affinity Grids to prepare untagged as well as myc- and Flag-tagged ribosomes from *E. coli* extracts using antibodies against ribosomal subunit 3, myc and Flag tag, respectively. In all cases, the procedure was successful, revealing ribosomes specifically recruited to the Affinity Grids. The protein A/antibody strategy was then applied to prepare untagged human RNA polymerase II from HEK-293T cell extract using a specific antibody against subunit Rpb1. The resulting Affinity Grid preparations allowed cryo-EM data collection on vitrified specimens and calculation of a 3D map of human RNA polymerase II at a resolution of 25 Å. These results were published in Kelly *et al.* (2010b).

Application of the Affinity Grid to the extracellular domain of the Notch receptor

After establishing and testing monolayer purification and Affinity Grid with test specimens, the technique was finally applied to study the structure of the Notch extracellular domain (NECD). A His-tagged version of the *Drosophila* NECD was expressed in Sf9 cells in a secreted from. The NECD expressed at such low levels that conventional purification by Ni-NTA affinity chromatography from the growth medium was unsuccessful. The medium containing the His-tagged NECD was thus applied to Affinity Grids, which were negatively stained and imaged in the electron microscope. Despite consisting of a linear arrangement of 36 EGF repeats, the NECD particles were globular, and 3D reconstructions using cryo-negatively stained specimens revealed three defined conformations. The same three conformations were also seen in 3D reconstructions of the extracellular domain of human Notch 1, demonstrating that the structure of the NECD and its conformational states are conserved from flies to humans. Furthermore, antibody labeling the positions of the C terminus and of the binding site for the Notch ligand, Delta, successfully identified the positions of the respective domains in the structure. Finally, mass measurement by scanning transmission electron microscopy revealed that the NECD forms dimers in solution. These results were published in Kelly *et al.* (2010c).

Further improvement of the Affinity Grid

Deborah Kelly developed the monolayer purification and Affinity Grid techniques and successfully applied them to the structural study of several complexes. However, although several members of the Walz group tried to use these methods for various other complexes, nobody else was successful. The problem is a very high background of proteins that bind non-specifically to the lipid monolayer, which makes it impossible to identify the protein of interest in EM images. Also, the protein A/antibody strategy to prepare non-His-tagged complexes is cumbersome and requires multiple incubation steps. To find an easier way to use monolayer purification/Affinity Grid for the preparation of non-His-tagged proteins and to minimize non-specific binding, I decided to explore different affinity systems. For several reasons I focused on the interaction of glutathione-S-transferase (GST) with glutathione (γ,l-glutamyl-l-cysteinylglycine, GSH). First, the GSH–GST interaction is more specific than the Ni-NTA–His tag interaction and should therefore lead to a reduction in background binding. Second, the GSH–GST interaction is often used for pull-down experiments and it is therefore well established. Third, a library of yeast strains overexpressing GST-tagged open reading frames is commercially available and would thus minimize the effort to obtain GST-tagged target proteins.

Results

Design of a GSH lipid

Based on the structure of the Ni-NTA lipid, I designed a lipid functionalized with a GSH group (Figure 4.3). The GSH group is linked to the headgroup of a 1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine (DOPE) lipid through a 6-carbon linker. The linkage of the GSH group to the lipid is similar to the linkage of GSH to beads in commercially available GSH resins, such as glutathione-agarose (Piercenet). 200 mg of this lipid was synthesized by Avanti Polar Lipids.

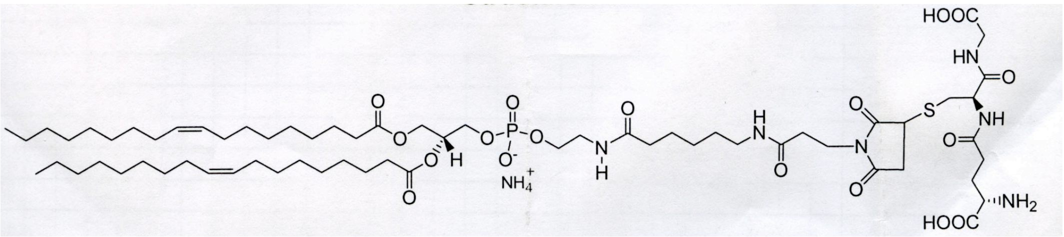


Figure 4.3: Chemical structure of the GSH lipid synthesized by Avanti Polar Lipids.

Testing the use of GSH lipid-based Affinity Grids to assemble complexes

I first tested the possibility of assembling complexes on a GSH-based Affinity Grid. I chose the complex formed by the multifunctional guanine nucleotide exchange factor (GEF) TRAPPII with the small Rab GTPase Ypt1, the structure of which had recently been determined in the Walz laboratory (Yip *et al.,* 2010). I expressed GST-tagged Ypt1 in *E. coli* and purified it using GSH-sepharose and size-exclusion chromatography (Figure 4.4A). I also purified TRAPPII from *S. cerevisiae* by TAP purification. By binding GST-Ypt1 to a freshly prepared Affinity Grid containing 10% GSH-lipid, I was able to recruit untagged TRAPPII complex. The GSH-based Affinity Grid was first incubated with 3.5 μl of 0.1 mg/ml GST-Ypt1. After 2 minutes the grid was blotted and then 3.5 μl of 0.005 mg/ml purified TRAPPII was added. After 5 minutes of incubation, the grid was washed three times with buffer before being stained with 0.75% uranyl formate. The prepared sample was inspected under the electron microscope, and images were recorded at a magnification of 42,000x on a 2K x 2K CCD camera (Figure 4.4B). From 1327 CCD images, I selected 3,309 particles and sorted them into 50 classes using the SPIDER software (Frank *et al.,* 1996). Out of the 50 resulting class averages of TRAPPII, only one revealed an additional density representing the bound GST-Ypt1

(Figure 4.4C, panel 2). This result suggested that most of the TRAPPII complexes were not recruited to the Affinity Grid through GST-tagged Ypt1, but non-specifically associated with the lipid monolayer.

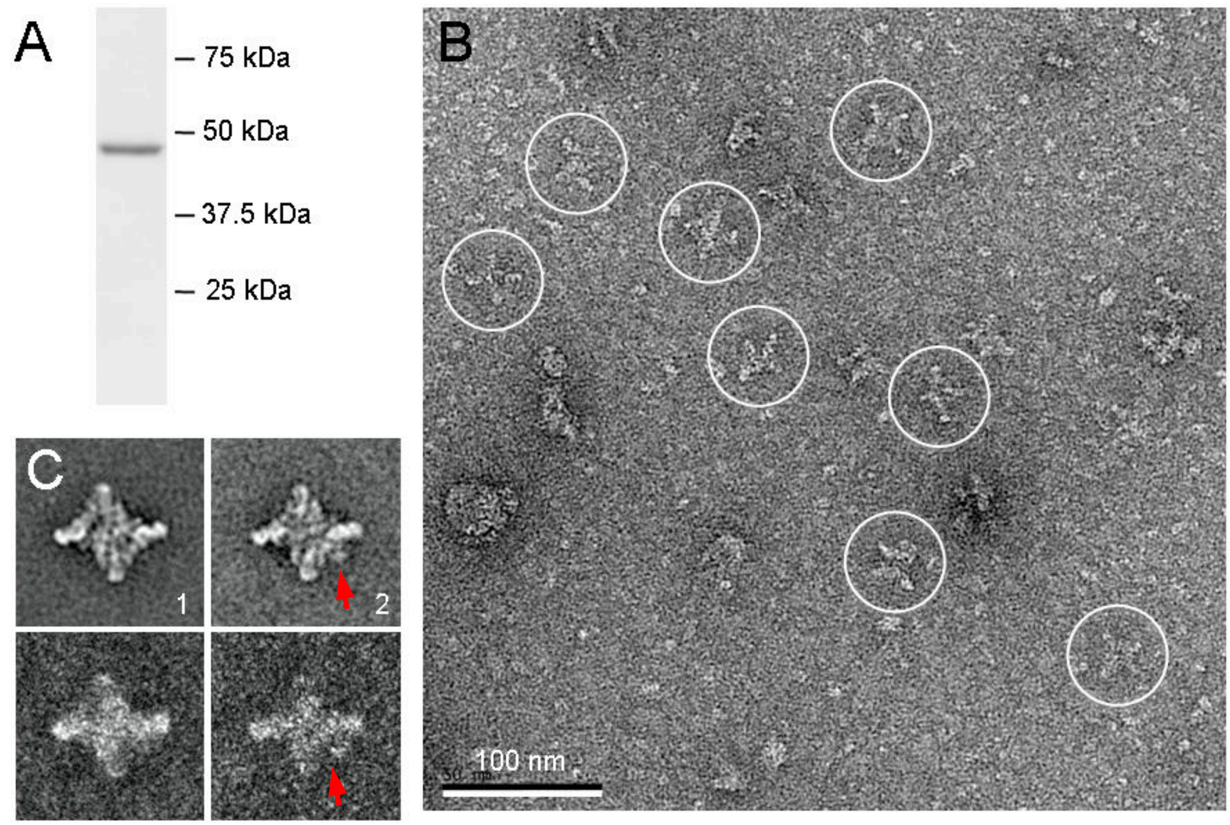


Figure 4.4: Recruitment of untagged TRAPPII complexes to a GSH lipid-based Affinity Grid through GST-tagged Ypt1. A) SDS-PAGE gel of GST-tagged Ypt1 purified from *E. coli*. B) Micrograph of negatively stained TRAPPII complexes recruited to a GSH-based Affinity Grid through GST-tagged Ypt1. C) Top panels: Class averages of TRAPPII complexes. The majority of class averages did not reveal a density for GST-tagged Ypt1 (panel 1) and only one average showed a density for bound GST-tagged Ypt1 (panel 2). Bottom panels: Corresponding variance maps. Red arrows point to densities representing GST-tagged Ypt1 bound to the TRAPPII complex.

Testing the use of GSH lipid-based Affinity Grids for protein purification

To test whether the GSH lipid could be used to purify GST-tagged complexes from cell extracts, I decided to use the ribosome as test specimen, which was previously used to establish the Ni-NTA lipid-based Affinity Grid. I selected the strain from the GST yeast library that overexpresses a GST-tagged version of the ribosomal Rpl5 subunit. The cells were grown, lysed in the presence of buffer containing 10 mM Tris, pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT and protease inhibitors, and the lysate was centrifuged for 30 minutes at 10,000 x g to remove unbroken cells. The protein concentration of the clarified cell lysate was determined by absorbance at 280 nm and adjusted by dilution to 0.5 mg/ml. The cell lysate was incubated for 2 minutes on an Affinity Grid containing a lipid monolayer with 2% GSH lipid. The sample was then negatively stained and inspected in the electron microscope. Images were collected at a magnification of 42,000x on a CCD camera (Figure 4.5). While the images clearly revealed ribosomal particles, there was still a significant background of non-specifically bound proteins, indicating that the GSH lipid did not solve the problem of non-specific binding. Moreover, very similar images were obtained with control specimens that were prepared with Affinity Grids that were covered with a monolayer that did not contain any GSH lipid.

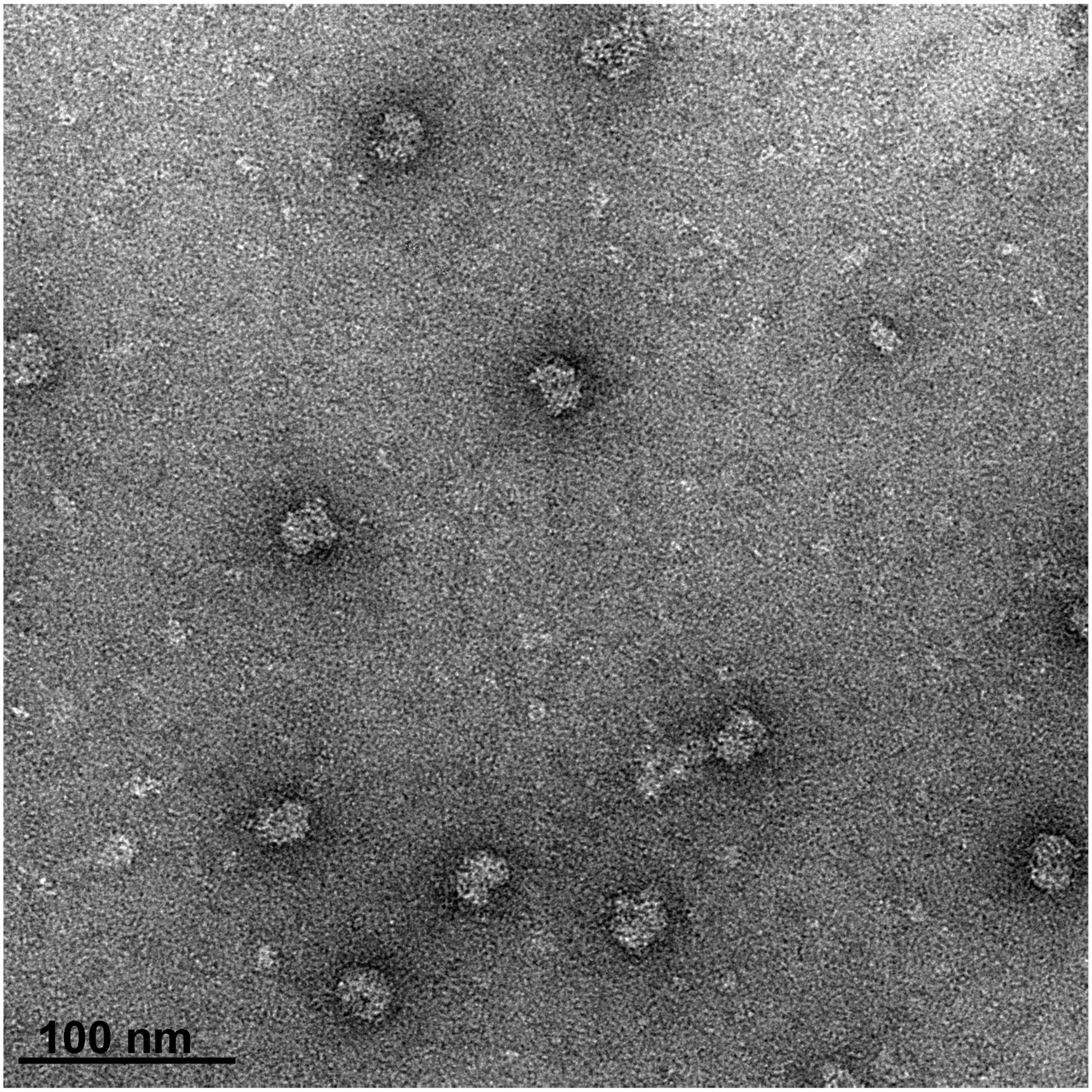


Figure 4.5: Representative EM image of GST-tagged ribosomes recruited from *E. coli* extract to a lipid monolayer containing GSH lipids. While ribosomal particles are clearly visible, the image also shows the presence of many smaller proteins. The scale bar is 100 nm.

The results obtained with the TRAPPII and GST-tagged Ypt1 and with the recruitment of GST-tagged ribosomes suggested that adsorption to a lipid monolayer is largely non-specific in nature. As this behavior was not reported in previous papers describing the monolayer purification and Affinity Grid techniques, it was necessary to reproduce the previously published data.

Reproducing previously published data

I first tried to reproduce the result described in Kelly *et al*. (2008a) that the amount of bound His-tagged Tf-TfR complex correlates with the percentage of Ni-NTA lipid added to the lipid monolayer. Briefly, a 0.01 mg/ml solution of Tf-TfR complex was overlaid with a lipid monolayer formed by the filler lipid dilauroyl phosphatidylcholine (DLPC) and increasing concentrations of functionalized Ni-NTA lipid (0%, 2%, 20%, and 40%). Despite different Ni-NTA lipid concentrations in the monolayers, images taken of all specimens, including the one produced with a lipid monolayer containing no Ni-NTA lipids, showed comparable concentrations of bound Tf-TfR complex (Figure 4.6). Because this result directly contradicted the result published in Kelly *et al*. (2008a) and suggested that most of the Tf-TfR complexes adsorbed to the lipid monolayer through non-specific interactions, Daniel Zachs, EM technician in the Walz group, was asked to independently do the same experiment. Daniel Zachs obtained the same results, seeing no significant difference in the amount of adsorbed Tf-TfR complex with lipid monolayers containing no or different percentages of Ni-NTA lipid. Thus, the only conclusion that can be drawn is that His-tagged Tf-TfR complex predominantly adsorbs to lipid monolayers through non-specific interactions rather than through specific recruitment of the His tag by the Ni-NTA lipids.

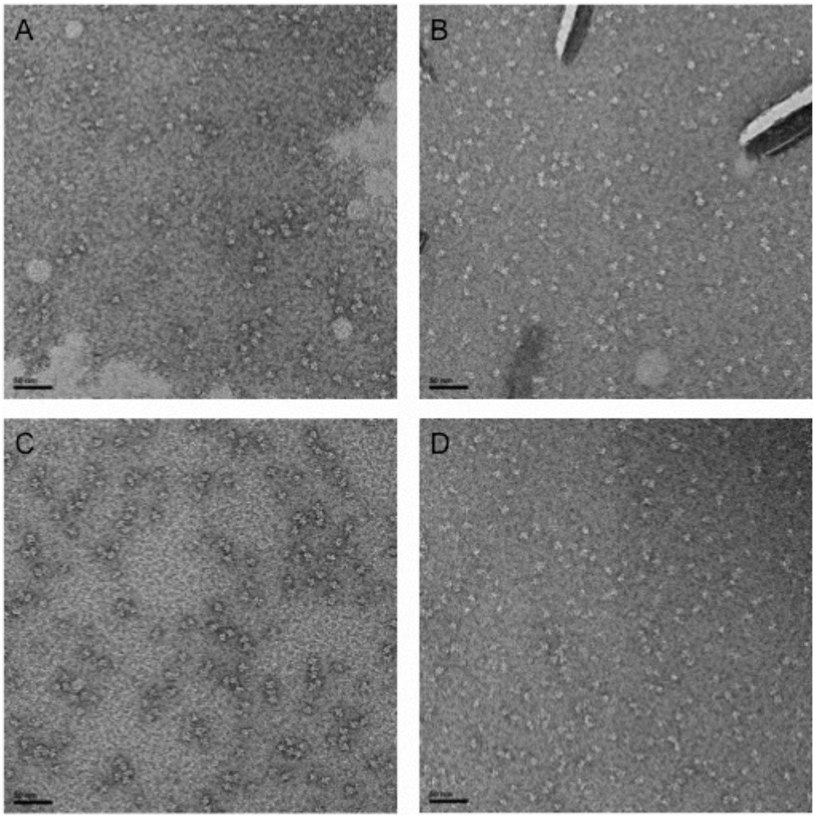


Figure 4.6: Recruitment of His-tagged Tf-TfR complex to lipid monolayers containing various amounts of Ni-NTA lipids. Panels show representative EM images of specimens obtained with 0% Ni-NTA lipid (A), 2% Ni-NTA lipid (B), 20% Ni-NTA lipid (C), and 40% Ni-NTA lipid (D). All images show a similar concentration of bound Tf-TfR complex, suggesting that most of the binding is non-specific. The scale bars are 50 nm.

Because our attempts to reproduce the first experiment described in Kelly *et al*. (2008a) produced a result opposite from the published one, I decided to also repeat the other experiments described in that paper. I thus attempted to use a Ni-NTA lipid-containing monolayer to specifically recruit His-tagged Tf-TfR complex that was added to insect cell extract. As described in the paper, a monolayer with 2% Ni-NTA lipid was used.

However, cell extract at a concentration of 1 mg/ml instead of 7 mg/ml was used, and His-tagged Tf-TfR complex was added to a final concentration of 0.0075 mg/ml rather than only 0.0004 mg/ml. In addition, 50 mM imidazole was added to the cell extract to suppress non-specific binding. As negative control, the experiment was repeated with a lipid monolayer containing no Ni-NTA lipid. Images of specimens prepared with the lipid monolayer lacking Ni-NTA lipid showed no Tf-TfR complex but a very high background of non-specifically bound proteins (Figure 4.7A). In contrast to the previously published results, despite using a lower concentration of cell lysate and an increased amount of His-tagged Tf-TfR complex, images of specimens prepared with the lipid monolayer containing 2% Ni-NTA lipid showed no enrichment in Tf-TfR complex, and the same high background of non-specifically bound protein (Figure 4.7B).

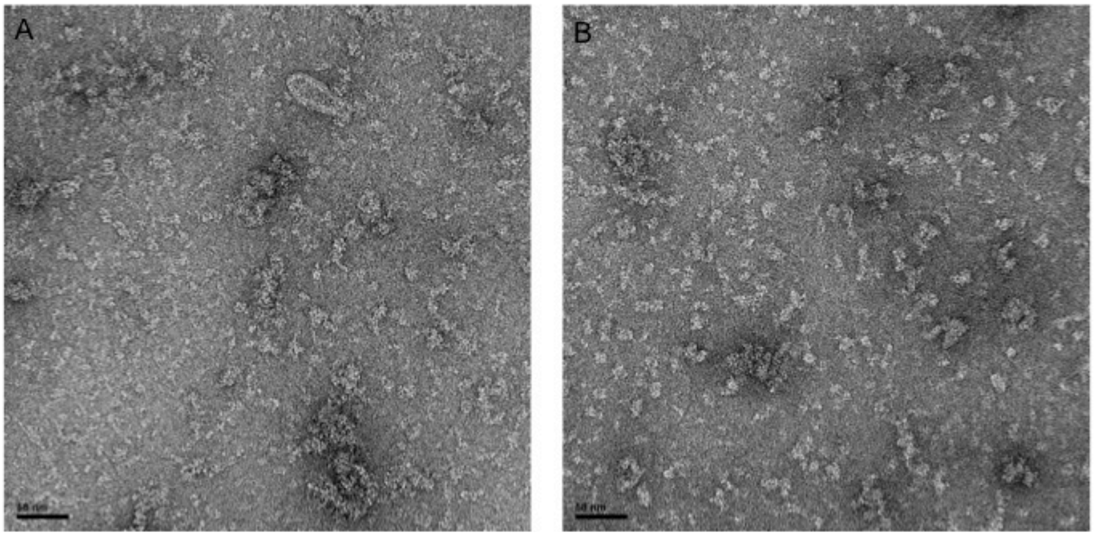


Figure 4.7: Recruitment of His-tagged Tf–TfR complex from insect cell lysate using a Ni-NTA lipid-containing monolayer. A) Representative EM image of negatively stained particles adsorbed to a lipid monolayer containing no Ni-NTA lipid.

Figure 4.7 (Continued) B) Representative EM image of negatively stained particles adsorbed to a lipid monolayer containing 2% Ni-NTA lipid. The scale bars are 50 nm.

To further analyze the adsorbed complexes, specimens with monolayers containing 20% Ni-NTA lipid were prepared, and the adsorbed complexes were eluted with 300 mM imidazole as described in Kelly *et al*. (2008a). The released proteins were sent for identification by mass spectrometry. The mass spectrometry results showed the presence of both Tf and TfR, but – in contrast to the published results – also a large number of contaminating proteins (Table 4.1). This experiment was repeated with a very similar outcome. Contrary to the published results, my own results thus suggest that monolayer purification cannot be used to specifically recruit only His-tagged Tf-TfR complex from insect cell extract.

Protein identified by Mass Spectrometry	Full name of protein	No. of unique tryptic peptides found
TFR	Transferrin receptor 1	20
TF	Serotransferrin	13
YWHALE	14-3-3 protein epsilon	11
TUBB2c	Tubulin beta-2C	9
VCP	Transitional endoplasmic reticulum ATPase	9
HSPA8	Heat shock cognate 71 kDa protein	9
TUBA4A	Tubulin alpha-4A	8
ACTA2	Actin	8
SERBINB4	Serpin B4	7
CALM1	Calmodulin	7
EEF1A2	Elongation factor 1-alpha 2	7
EEF2	Elongation factor 2	7
ACTB	Actin, cytoplasmic 1	6
ACTBL2	Beta-actin-like protein 2	6
HSP90AA1	heat shock 90kDa protein 1	5
PSMC3	26S protease regulatory subunit 6A	4
HSPA5	HSPA5 protein	4
TUBB1	Tubulin beta-1	4
ARF1	ADP-ribosylation factor 1	4

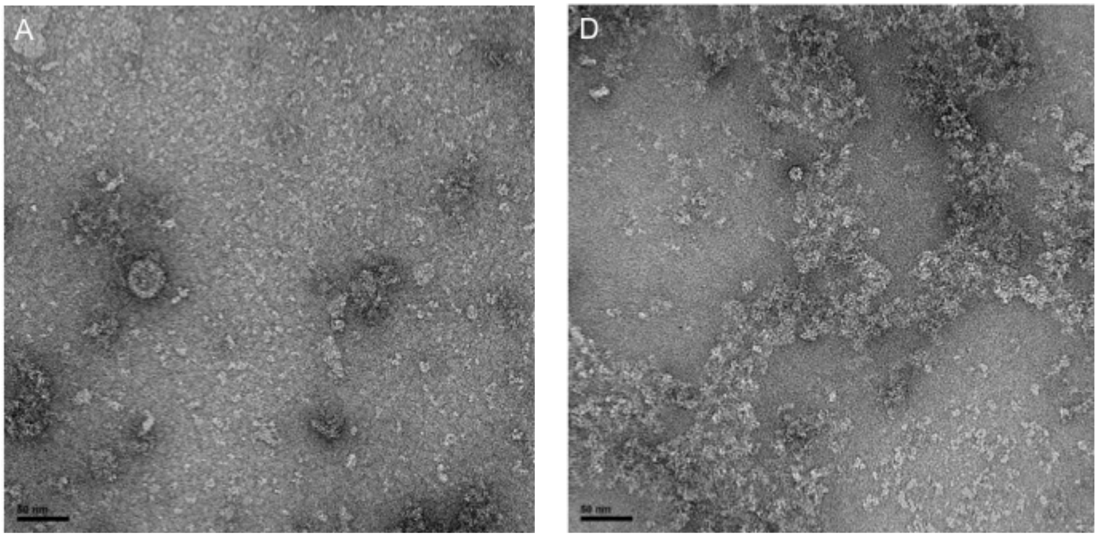
Table 4.1: Mass spectrometry data detailing the proteins eluted from a lipid monolayer containing 20% Ni-NTA lipid. His-tagged Tf-TfR complex was added to insect cell extract, and the extract was overlaid with a lipid monolayer containing 20% Ni-NTA lipids. The proteins adsorbed to the monolayer were eluted with 300 mM imidazole and analyzed by mass spectrometry. Only proteins are shown for which more than four tryptic peptides were identified.

Improvements of the Affinity Grid

The above experiments established that recruitment of His-tagged proteins to a Ni-NTA lipid-containing monolayer is not nearly as specific as previously published, and the high background of non-specifically bound proteins makes this technique unsuitable for single-particle EM specimen preparation. To be able to use lipid monolayers for this purpose, it was essential to find ways to drastically reduce the extent of non-specific protein binding to the lipid monolayer.

A method often used in biochemistry to reduce non-specific protein interactions is to increase the ionic strength of the solution by increasing the salt concentration. This idea was expanded upon by also testing whether it is possible to reduce non-specific protein binding by forming the monolayer with a lipid that lacks charges from the phosphodiester and head groups. Lipid monolayers with DLPC, the commonly used filler lipid, and dilauroyl glycerol (DLG), missing both the phosphodiester and head groups, were formed over 0.1 mg/ml *E. coli* cell extract. Visual inspection of the resulting monolayer specimens suggested that increased salt concentrations did not significantly reduce non-specific binding to the DLPC monolayer (Figure 4.8A-C) but it did appear to reduce non-specific binding to the DLG monolayer (Figure 4.8D-F).

Figure 4.8



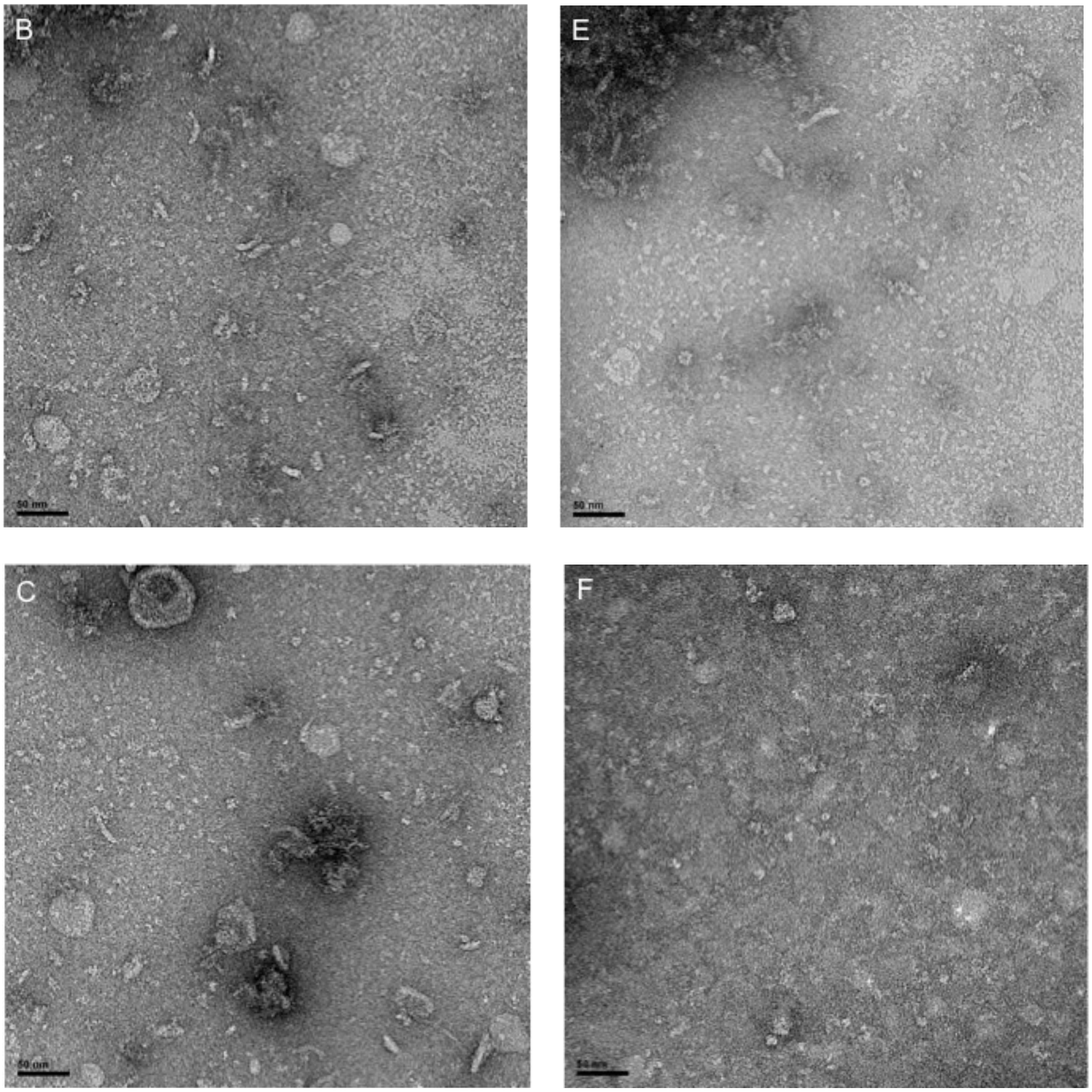
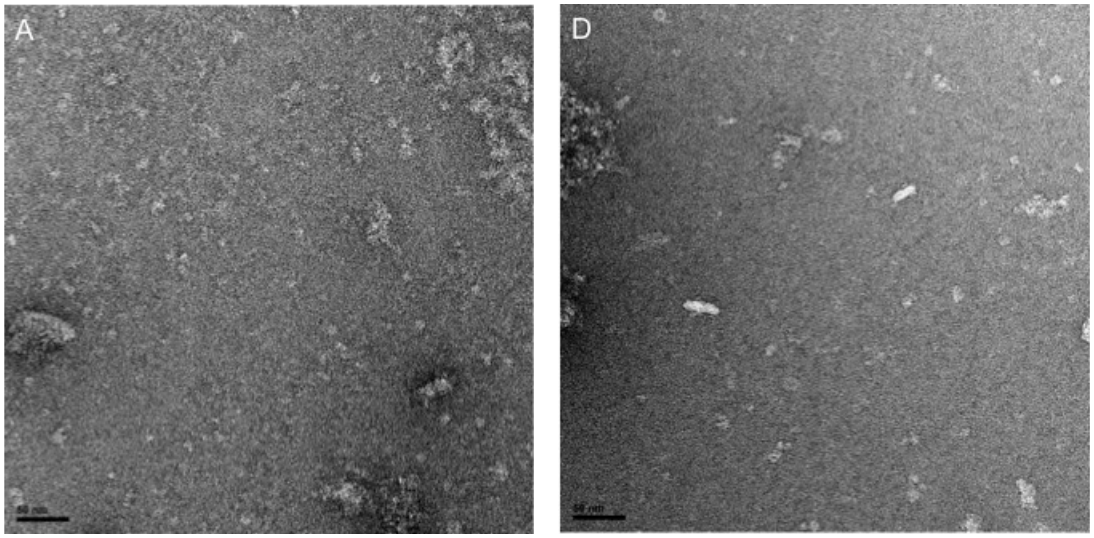


Figure 4.8 (continued): Effect of salt concentration on non-specific protein binding to monolayers formed by DLG and DLPC. A-C) Representative EM images of monolayer specimens obtained with a DLPC monolayer and 150 mM NaCl (A), 300 mM NaCl (B) and 500 mM NaCl (C). D-F) Representative EM images of monolayer specimens obtained with a DLG monolayer and 150 mM NaCl (D), 300 mM NaCl (E) and 500 mM NaCl (F). The scale bars are 50 nm.

To further reduce non-specific binding, PEG lipids were added to the DLG monolayer, because in a study using self-assembly monolayers, PEG resulted in the least protein adsorption out of all the head groups tested (Prime *et al.,* 1991). 0.1 mg/ml *E. coli* cell lysate containing 500 mM NaCl was overlaid with a DLG monolayer containing 0%, 5%, 10%, 20%, 50% and 100% of either PEG350-DOPE or PEG2000-DOPE. PEG350-DOPE was not effective at suppressing non-specific protein adsorption (Figure 4.9), as non-specifically bound proteins were present even in the monolayer sample obtained with 100% of the PEG350-DOPE lipid (Figure 4.9F).

Figure 4.9



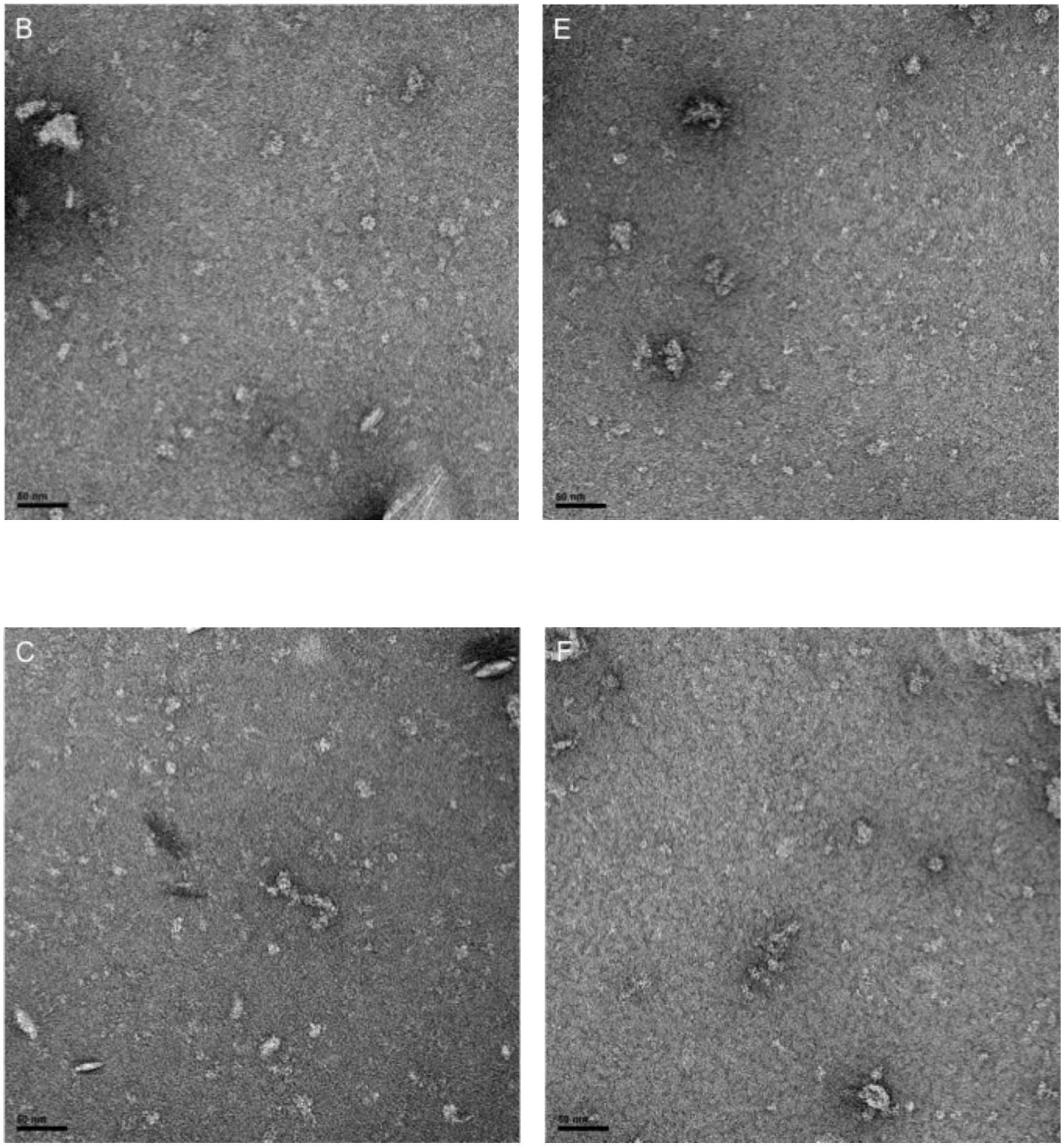


Figure 4.9 (continued): Effect of PEG350 lipid on non-specific protein binding to DLG monolayers. *E. coli* cell lysate containing 500 mM NaCl was overlaid with DLG monolayers containing no PEG350 lipid (A), 5% PEG350 lipid (B), 10% PEG350 lipid, (C), 20% PEG350 lipid (D), 50% PEG350 lipid (E), and with a monolayer containing 100% PEG350 lipid (F). The scale bars are 50 nm.

Although there was no clear correlation between the percentage of PEG2000-DOPE lipid in the monolayer and the level of non-specific protein binding, samples obtained with a 100% PEG2000-DOPE monolayer (Figure 4.10D) showed less non-specifically bound protein than samples obtained with a 100% DLG monolayer (Figure 4.10A). Repetition of this experiment with a 0.1 mg/ml His-tagged Tf-TfR complex solution instead of *E. coli* cell extract showed the same result (Figure 4.11B).

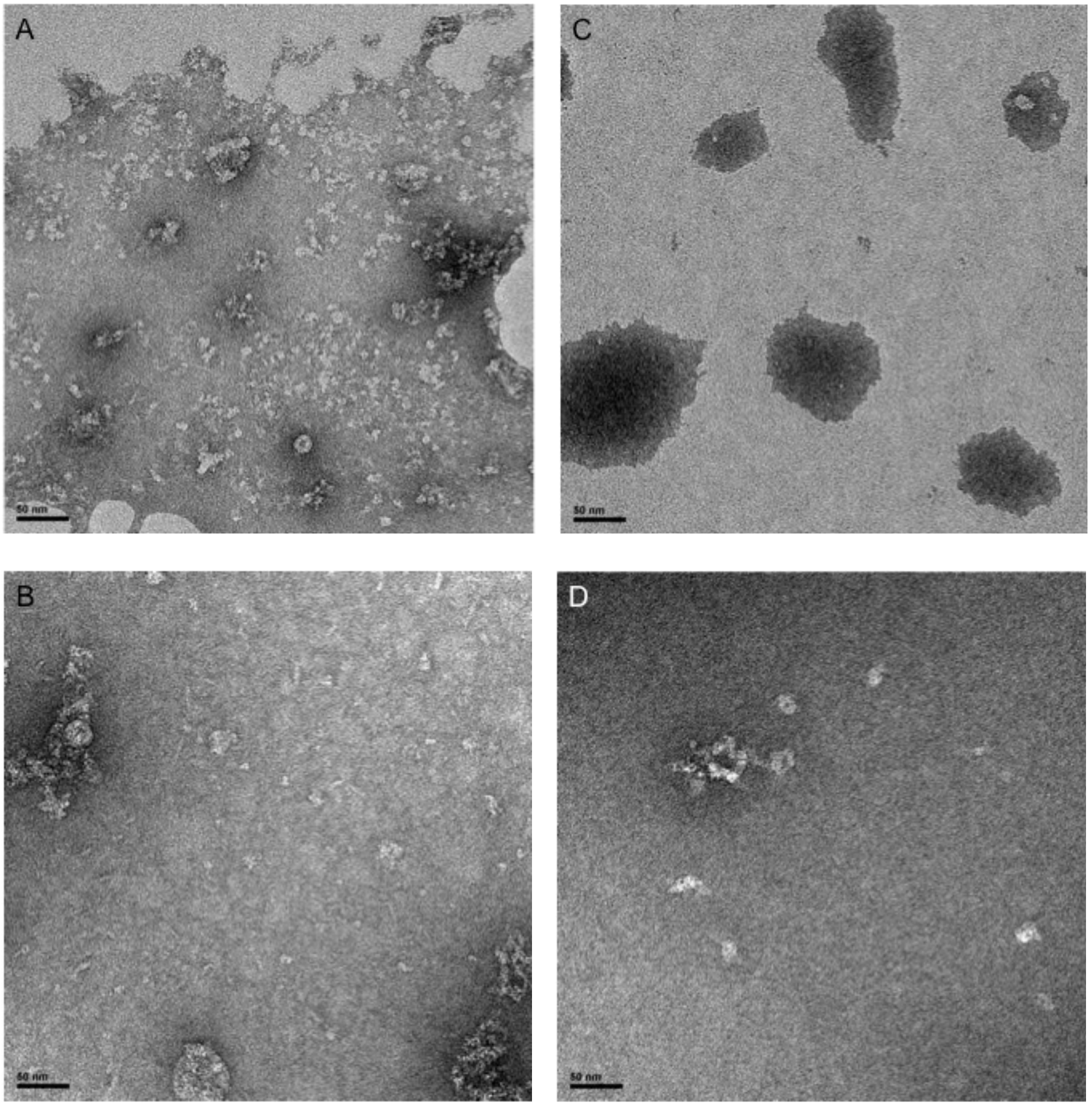


Figure 4.10: Effect of PEG2000 lipid on non-specific protein binding. *E. coli* cell lysate containing 500 mM NaCl was overlaid with DLG monolayers containing no PEG2000-lipid (A), 5% PEG2000 lipid (B), 50% PEG2000-lipid (C), and with a monolayer containing 100% PEG2000-lipid (D). The scale bars are 50 nm.

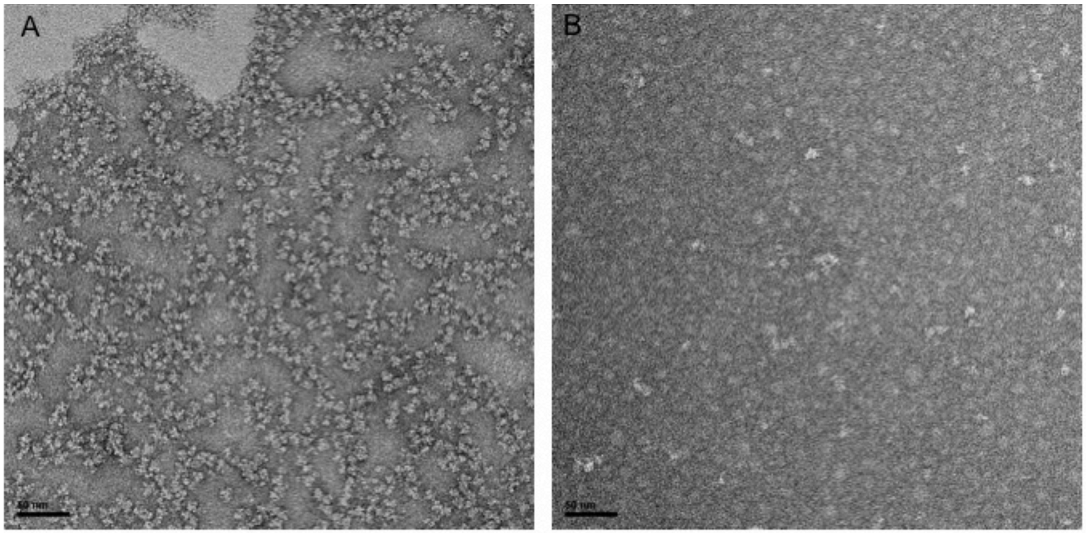


Figure 4.11: Non-specific adsorption of His-tagged Tf-TfR complex on a DLG monolayer (A) in comparison to a PEG2000 lipid monolayer (B). The scale bars are 50 nm.

Lipid synthesis

In parallel, in order to adapt the monolayer purification and Affinity Grid techniques to include the PEG feature, I synthesized GSH-conjugated dioleyl phosphatidyl ethanolamine (DOPE) lipid (**8** in Figure 4.12). Unlike the commercially purchased PEG lipid, the design of this lipid incorporates four ethylene glycol (EG) units between the glutathione and the remainder of the lipid. The EG units were added to help prevent non-specific protein binding to the lipid monolayer.

In a first step, DOPE (Avanti Polar Lipids) was reacted with SM-(PEG)$_4$ (Succinimidyl-([N-maleimidopropionamido]-4ethyleneglycol) ester) (Piercenet) (**2** in Figure 4.12), and the product, **5**, was purified with HPLC. Subsequently, the glutathione was then conjugated to **5** through a reaction of the maleimide group on **5** with the sulfhydryl group on glutathione. The final product, **8**, was purified by HPLC. The identities of all lipids were confirmed by mass spectrometry and NMR analysis.

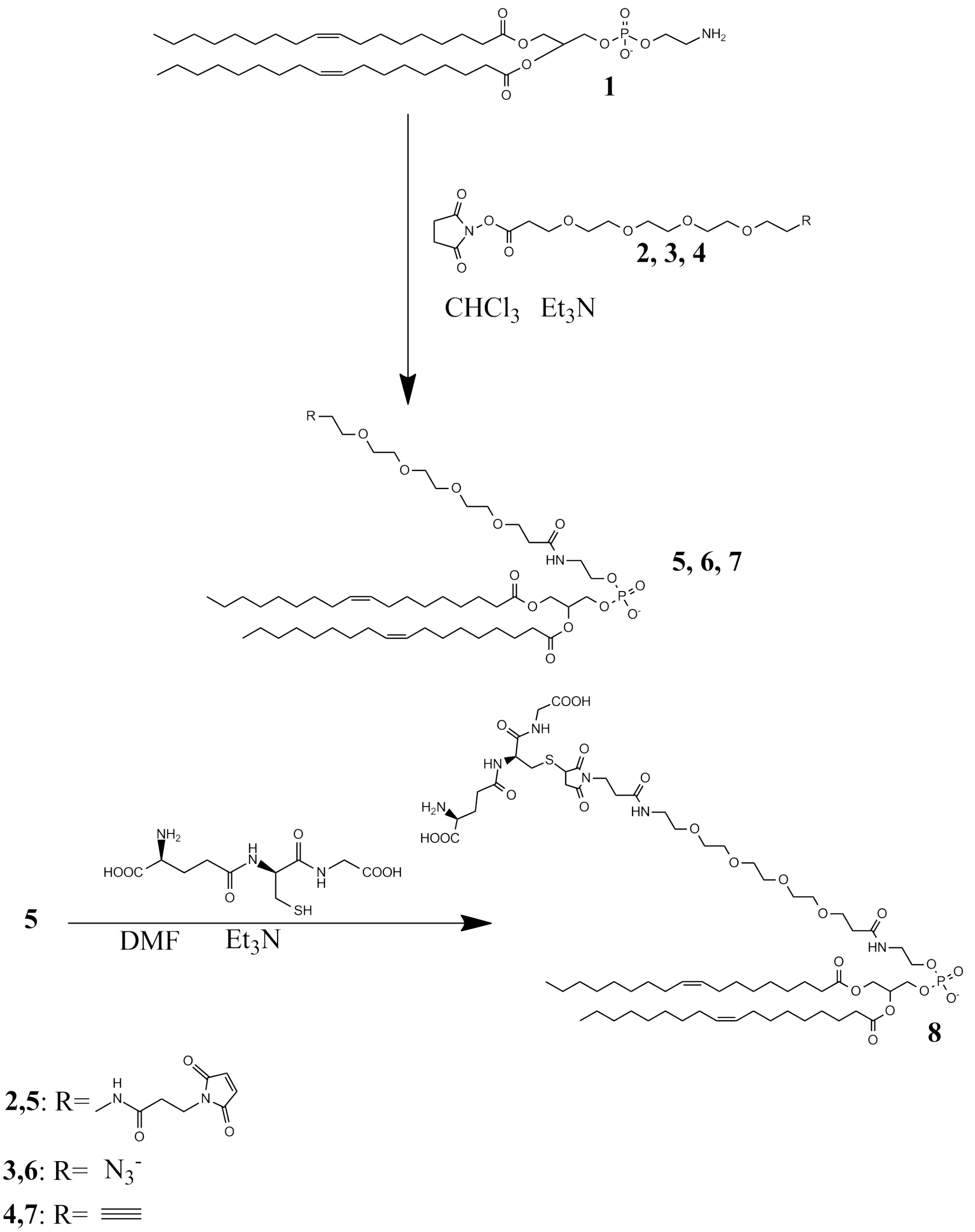


Figure 4.12: Reaction scheme for the synthesis of lipids with functionalized head groups.

To test if the newly synthesized lipids were effective in preventing non-specific protein binding, monolayers consisting of different percentages of the synthesized lipids, namely 0%, 50% and 100%, were cast over a solution of 0.1 mg/ml His-tagged Tf-TfR complex. In theory, the His-tagged Tf-TfR complex should not bind to the synthesized GSH-PEG$_4$ lipid. Indeed, as shown in Figure 4.13, when compared with a monolayer containing only DLPC, there was a significant reduction in binding of Tf-TfR complex to monolayers consisting of 50% and 100% GSH-PEG$_4$ lipid.

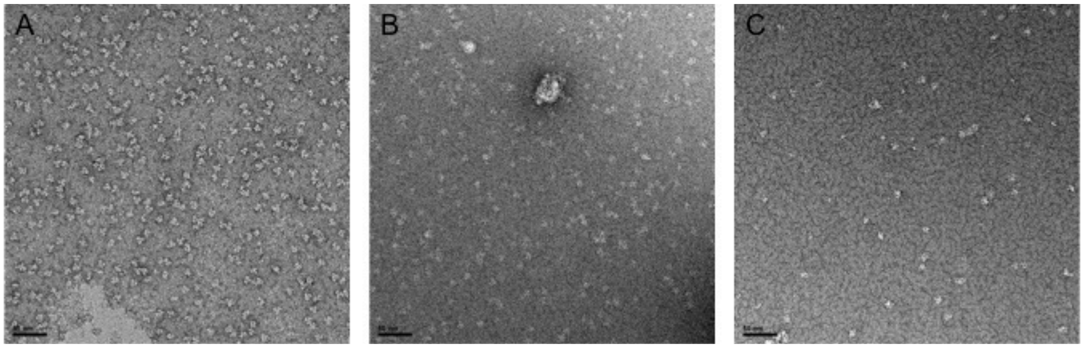


Figure 4.13: Non-specific adsorption of His-tagged Tf-TfR complex on a pure DLPC monolayer (A), a DLPC monolayer containing 50% GSH-PEG$_4$ lipid (B), and a pure GSH-PEG$_4$ lipid monolayer (C). The scale bars are 50 nm.

Discussion

Lipid monolayers have been used for many years to grow 2D protein crystals (Darst et *al.*, 1991; Chiu *et al.*, 1997), but the use of monolayers containing functionalized lipids as a tool to purify tagged proteins directly from impure solutions onto an EM grid has been a more recent innovation (Kelly *et al.,* 2008a; Kelly *et al.,* 2010a). My attempts to use monolayer purification and Affinity Grid now revealed an inherent problem with these techniques, namely the large extent of non-specific binding of proteins to the lipid monolayer that almost completely obscures any specific binding of the target protein. This observation is also made when lipid monolayers are used to prepare 2D crystals from a protein solution (Kubalek *et al.*, 1991; Thess *et al.*, 2000) except that it poses less of a problem since purified proteins are used in 2D crystallization. Thus, to make these techniques useful for single-particle EM, it is imperative to find ways to prevent or at least greatly reduce the extent of non-specific protein binding to lipid monolayers.

Non-specific adsorption of proteins to surfaces has been studied extensively by both physical chemists and material scientists. Because non-sticky surfaces are important in many areas, such as in microfluidic devices, in single-molecule studies, and in creating tools for drug delivery (Lasic and Martin, 2005; Lasic and Papahadjopoulos, 1998), many groups have attempted to create surfaces or modifications to surfaces that prevent non-specific protein adsorption. In 1991, Prime and Whitesides demonstrated with self-assembled monolayers that modification of lipid head groups with a PEG group is more effective in preventing protein adsorption compared with modifications with glucose and alkanes (Prime and Whitesides, 1991). This study also demonstrated that the effectiveness of preventing non-specific protein adsorption did not depend on the thickness of the PEG monolayer but depended on the extent of coverage of the surface with PEG. PEG has since been used in many applications. For instance, inclusion of PEG lipids into liposomes increases their half-life in circulation from minutes to hours or even days (Papahadjopoulos *et al.*, 1991). PEG has also been used to modify solid surfaces to control the density of adsorbed cells or proteins (Dori *et al.*, 2000). Most recently, the passivation of quartz surfaces with PEG has been used to combine conventional pull-down assays with single-molecule studies, which made it possible to directly visualize complexes such as mtorc1, MAVS and PKA from cell lysate using single-molecule immunofluorescence (Jain *et al.,* 2011).

Based on the widespread use of PEG in passivating surfaces, PEG lipids were an obvious choice in trying to minimize non-specific protein adsorption to lipid monolayers used in the monolayer purification and Affinity Grid techniques. Indeed, my results show that incorporation of PEG350 and PEG2000 lipids in the lipid monolayers did reduce the extent of non-specific protein adsorption (Figures 4.10 and 4.11). However, the most effective lipid in preventing non-specific binding was the GSH-PEG$_4$ lipid I synthesized (Figure 4.12). A recent atomic force microscopy study of lipid monolayers formed by different PEG lipids (PEG120, PEG750, PEG2000 and PEG5000) showed that the lipid with the shortest PEG chain, PEG120, formed the monolayer with the least defects (Dori *et al.*, 2000). This result may explain why the GSH-PEG$_4$ lipid, which contains only 4 PEG groups, is more effective in preventing non-specific adsorption than the PEG2000 lipid, which contains 45 PEG groups. However, it does not explain why the GSH-PEG$_4$ lipid is more effective than the PEG350 lipid, which contains only 7 PEG groups. It may be that the added GSH group in the GSH-PEG$_4$ lipid adds to the protein repulsive effect of this lipid. This possibility will be tested experimentally by directly comparing non-specific protein binding to lipid monolayers formed by PEG$_4$ and GSH-PEG$_4$ lipids.

It remains to be tested whether the use of PEG lipids can reduce non-specific protein adsorption to a level that does not obscure specific protein binding. If this is not the case, additional washing steps during grid preparation might help to remove non-specifically bound proteins. However, current lipid monolayers are too fragile to survive extensive washing steps. To increase resilience, amphiphilic block copolymers (Kita-Tokarczyk *et al.,* 2009) could be used, which form more robust monolayers. Also, the hydrocarbon portions of the lipids or block copolymers could be modified to include cross-linkable groups, such as acroyl, alkene (Conboy *et al.,* 2003) and diyne (Zhang *et al.,* 2003). Cross-linking would then drastically improve the robustness of the monolayers. Alternatively, it is conceivable that EM grids could be produced with covalently bound PEG that has been modified with functional groups that recruit tagged proteins. While this is not possible for carbon-coated grids, silicone-coated grids decorated with PEG are already commercially available (Dune Sciences).

Even if it will not be possible to develop a lipid monolayer-based technique that allows direct recruitment of target proteins from cell extracts, reducing non-specific adsorption would enhance the use of lipid monolayer for the assembly of complexes as was attempted in the visualization of the interaction of the Tf–TfR complex with GP1 (Kelly *et al.,* 2010a) and the TRAPPII complex with Ypt1 (Figure 4.4B). The unlabeled Tf–TfR complex was recruited to a Ni-NTA lipid-containing monolayer through a His-tagged version of GP1, but the resulting 3D reconstruction of the Tf–TfR complex showed no indication of bound GP1 (D. Kelly and T. Walz, unpublished results). Similarly, the unlabeled TRAPPII complex was recruited to a GSH lipid-containing monolayer through a GST-tagged version of Ypt1, but only one out of 50 class averages revealed Ypt1 bound to the TRAPPII complex (Figure 4.4C). The most likely explanation of these results is that the majority of the Tf–TfR and TRAPPII complexes were not specifically recruited to the monolayer through their respective ligands but instead adsorbed non-specifically to the lipid monolayer. If such non-specific adsorption could be eliminated or at least substantially reduced, 2D class averages and 3D reconstructions would be much more likely to reveal the bound ligand.

A future direction for lipid monolayers would be to extend their use beyond the visualization of protein-protein interactions by taking advantage of click chemistry. Click chemistry uses two functional groups, the alkyne and azide groups, neither of which reacts with other functional groups found in biological reactions, to covalently link two molecules. The reaction itself, a [3+2] cycloaddition of the azide and alkyne groups catalyzed by Cu(I), can be carried out under aqueous and mild conditions. I have already synthesized lipids functionalized with azide and alkyne groups (Figure 4.12) that can be incorporated into lipid monolayers. DNA or RNA with specific binding sequences can then be modified with the complementary group (commercially available) and covalently linked to the lipid monolayer simply by adding Cu(I). The monolayer thus decorated with DNA or RNA would then be ready to recruit the target protein.

In summary, if non-specific protein adsorption to lipid monolayers could be prevented or at least be significantly reduced, as my results suggest, the monolayer purification and

Affinity Grid techniques could still fulfill the promise they hold for fast and easy specimen preparation for single-particle EM.

Methods and Materials

Preparation of GST-tagged Ypt1

Full-length Ypt1 was cloned into pGEX6p-1 vector (Yip *et al.,* 2010) and overexpressed in BL21DE3 (pLys) *E. coli* cells. 2 l of culture were induced with 1 mM IPTG at OD_{600} = 0.6, and cells were harvested after 1 h. The cells were lysed by sonication in buffer containing 10 mM Tris, pH 8, 150 mM NaCl, 1 mM DTT, and spun down at 100,000 x g for 15 minutes. The clarified cell lysate was incubated with GSH-sepharose (GE Healthcare), and bound GST-Ypt1 was eluted with 10 mM glutathione. The final purification step was gel filtration with a Superose 12 column using a buffer containing 10 mM Tris, pH 8, 150 mM NaCl, 1 mM DTT. The identity of GST-Ypt1 was confirmed by SDS-PAGE.

Preparation of TRAPPII

The complex was purified as in Yip *et al.* (2010) with slight modifications. A yeast strain containing endogenously TAP-tagged Trs120 (Open Biosystems) was grown in 4 l of Yeast Extract Peptone Dextrose (YPD) medium to $OD_{600} < 10$. The cells were lysed in buffer containing 10 mM Tris, pH 8, 150 mM NaCl, 1 mM DTT and 1% 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS). With the exception of the detergent used, as mentioned above, the complex was purified according to the traditional tandem affinity purification protocol (Rigaut *et al.*, 1999). In addition, the concentration of CHAPS was steadily reduced to 0 by the end of the purification protocol. The identity of the complex was confirmed by negative-stain EM and LC/MS/MS (Taplin Mass Spectrometry Facility).

Preparation of yeast cell extract

Yeast strains containing overexpression construct of GST-Rpl5 (Open Biosystems) was grown overnight in 100 ml of galactose-containing media with -URA dropout supplement. Cells were harvested and lyzed in buffer containing 10 mM Tris, pH 8, 150 mM NaCl and protease inhibitor (Roche) with a bead beater. To remove insoluble cell debris, the cells were spun down at 14,000 rpm for 20 minutes at 4ºC on a 5415 R table-top centrifuge (Eppendorf). The protein concentration of the supernatant was quantified by measuring UV absorption at 280 nm.

Preparation of GSH lipid-based Affinity Grids

DOPC and GSH-DOPE were purchased from Avanti polar lipids. The lipids were dissolved in chloroform to 1 mg/ml. Mixtures of lipids in chloroform were prepared by mixing in 2% or 10% glutathione-DOPE by volume with 1 mg/ml of DOPC in chloroform. Teflon blocks containing 18 wells, with each well holding a volume of about 25 μl, were used to set up monolayer samples. To prepare the monolayer, 1 μl of lipid mixture in chloroform was overlaid across 25 μl of milliQ water. The Teflon blocks were incubated in a humidified environment at 4ºC overnight. The lipid monolayers were recovered by placing carbon-coated grids (stored for at least two weeks to ensure hydrophobicity of the carbon surface) on top of the monolayer. The grids were picked up with forceps and stored until use. Samples were added directly to the lipid monolayer-containing grids, incubated for 10 minutes, blotted and washed three times with buffer. Samples were then stained with 0.75% uranyl formate.

Preparation of His-tagged Tf-TfR complex

His-tagged Tf-TfR complex was prepared as described in Kelly *et al.,* 2008. The protein was quantified using the Bradford Assay (Biorad).

Preparation of insect cell extract

Sf9 cells were grown in shaker flasks at 28°C in Sf-900 II SFM serum-free medium (Invitrogen) containing 100 units/ml penicillin and 100 mg/ml streptomycin. The cells were harvested at a density of 4 x 10^6 viable cells per ml. Cells extracts were prepared by centrifugation of 50 ml of cell culture at 3,500 x g for 20 minutes at 4°C, and resuspending the cell pellets in 20 ml of lysis buffer containing 20 mM Tris, pH 8.0, 300 mM NaCl with sonication. Insoluble debris was removed by centrifugation at 100,000 x g for 30 minutes at 4°C. The supernatants were used for monolayer purification experiments. The total protein concentration was quantified by using the Bradford Assay.

Preparation of *E. coli* cell extract

A colony of ampicilin-resistant *E. coli* was inoculated into 50 ml of LB containing 100 μg/ml of ampicilin. The cells were harvested and lysed by sonication in buffer containing 20 mM Tris, pH 8.0, 150 mM NaCl. Cell debris was spun down at 14,000 rpm on a table-top centrifuge for 20 minutes at 4^{o}C. The supernatants were used for monolayer purification experiments. The total protein concentration was quantified by using the Bradford Assay. When needed, the salt concentration was adjusted to 300 mM or 500 mM NaCl by adding appropriate amounts of 5 M NaCl solution.

Adsorption of His-tagged proteins to Ni-NTA lipid-containing monolayers

DLPC and 1,2-dioleoyl-sn-glycero-3-[N-(5-amino-1-carboxypentyl)iminodiacetic acid] succinyl-nickel salt (Ni-NTA lipid) were purchased from Avanti Polar Lipids. Each lipid was reconstituted in chloroform to 1 mg/ml. Teflon blocks containing 18 wells, with each well holding a volume of ~25 μl, were used to set up monolayer samples. After placing 25 μl of protein solution or cell extract into a well, 1 μl of lipid mixture (DLPC containing the desired percentage of Ni-NTA lipid in chloroform) was added on top of the aqueous solution to form a monolayer at the air/water interface. The Teflon blocks were incubated in a sealed humid environment on ice for 15 min. Monolayer samples were recovered by placing a continuous carbon-coated EM grid on top of the monolayer. The grid was gently lifted off with forceps, blotted and stained with 0.75% uranyl formate.

Monolayer purification of His-tagged Tf-TfR complex from cell extracts

A volume of Tf-TfR complex solution (0.75 mg/ml) equivalent to 750 ng was added to 25 μl of Sf9 extract. For each monolayer trial, 25 μl of this mixture was used as the aqueous phase and overlaid with a DLPC monolayer containing 0%, 2% and 20% Ni-NTA lipid. 50 mM imidazole (final concentration) was added to the extracts before casting the lipid monolayers.

Recovering proteins from monolayer samples

Twenty monolayer samples were used to analyze proteins adsorbed to the Ni-NTA monolayers. The first monolayer sample was recovered with an EM grid, excess solution was blotted away, and the sample was incubated on the grid with 20 μl of 300 mM imidazole for 2 minutes. The 20-μl drop was removed from the grid and added to the next grid containing a monolayer sample. This procedure was repeated until all 20 samples were eluted into the same 20 μl drop of 300 mM imidazole.

Preparation of PEG lipids

PEG350-DOPE and PEG2000-DOPE were purchased as 10 mg/ml chloroform stocks from Avanti Polar lipids. Since the molecular weights of PEG-based lipids vary greatly from the filler lipids (DLPC or DLG), mol % was used instead of weight % to prepare lipid mixtures. Stock solutions of all the lipids were prepared at 1.6 μM $CHCl_3$ and the final concentration of the lipid mixtures was kept at 1.6 μM. The total volume of each lipid mixture was kept constant at 50 μl, and the volume of each lipid stock solution added to the final mixture of lipids was adjusted accordingly to reflect the accurate mol % composition of the PEG lipid.

Testing non-specific protein adsorption onto lipid monolayers

Other than exact lipid mixture used, protocol is as described under "Adsorption of His-tagged proteins to Ni-NTA lipid-containing monolayers".

Synthesis of lipids with functionalized head groups

Maleimide-PEG$_4$-DOPE: To 20 mg of DOPE dissolved in 1 ml of $CHCl_3$, 20 mg of succinimidyl-{[N-maleimidopropionamido]-4ethyleneglycol) ester (SM(PEG)$_4$) dissolved in 100 μl of chloroform($CHCl_3$) was added. 5 μl of triethylamine was added subsequently. The reaction mixture was allowed to go for 48 h under argon. The product was purified using HPLC starting with a gradient of 10% A (95% water : 5% methanol + 5 mM ammonium formate) and ending with 100% B (60% isopropanol, 35% methanol : 5% water + 5 mM ammonium formate). LC-MS: 1140.7. ^{1}H NMR ($CDCl_3$, 400 MHz) : δ 0.84-0.87 (t, 6H) δ 1.24-1.27(d, 40H) δ 1.97-2.24(d, 9H) δ 2.27-2.48(m, 5H) δ 2.48-2.56(m, 4H) δ 3.36-3.44(m, 6H) δ 3.57-3.60(d, 16H) δ 3.77-3.82(m, 4H) δ 3.92-3.93(t, 4H) δ 4.10-4.14(m, 1H) δ 4.31-4.37(m, 1H) δ 5.16-5.19(m, 1H) δ 5.30-5.33(m, 4H) δ 6.67(s, 2H) δ 7.98-8.06(m, 3H). NMR shifts are shown in the appendix.

GSH-PEG$_4$-DOPE: 20 mg of N-maleimidopropionamido]-4ethyleneglycol- dioleoyl-*sn*-glycero-3-phosphoethanolamide (**5** in figure 4.12) was dissolved in 2 ml of dimethylformamide (DMF). Triethylamine was added to 20 mg of glutathione in 2 ml of DMF until the pH was ~7. The reaction was run overnight under nitrogen. The product was purified using HPLC with a gradient of 10% A (95% water : 5% methanol + 5 mM ammonium formate) and ending with 100% B (100% methanol + 5 mM ammonium formate). LC-MS: 1447.8. ^{1}H NMR (CD_3OD, 400 MHz) :δ 0.88-0.92 (t, 6H) δ 1.24-1.27(t, 77H) δ 1.56-1.68(t, 9H) δ 1.97-2.18(m, 20H) δ 2.25-2.58(m, 21H) δ 3.36-3.44(m, 6H) δ 3.57-3.60(d, 20H) δ 3.77-3.82(m, 16H) δ 3.88-4.22(m, 14H) δ 4.10-4.14(m, 1H) δ 4.41-4.43(m, 3H) δ 5.28-5.39(m, 20H) δ 7.98-8.06(m, 1H). NMR shifts are shown in the appendix.

Azide-PEG$_4$-DOPE: To 20 mg of DOPE dissolved in 1 ml of $CHCl_3$, 20 mg of N-hydroxysuccinimide ester_tetraoxapentadecaneazide dissolved in 100 μl of $CHCl_3$ was added. 5 μl of triethylamine was added subsequently. The reaction mixture was allowed to go for 48 h under argon. The product was purified using HPLC starting with a gradient of 10% A (95% water : 5% methanol + 5 mM ammonium formate) and ending with 100% B (60% isopropanol : 35% methanol : 5% water + 5 mM ammonium formate). LC-MS: 1015. ^{1}H NMR ($CDCl_3$, 400 MHz): δ 0.84-0.87 (t, 6H) δ 1.24-1.27(d, 40H) δ 1.56(s, 4H) δ 1.97-2.00(d, 8H) δ 2.24-2.30(m, 5H) δ 2.50-2.53(m, 2H) δ 3.40-3.49(m, 5H) δ 3.58-3.68(d, 14H) δ 3.76-3.79(t, 2H) δ 3.91-3.95(q, 4H) δ 4.12-4.16(m, 1H) δ 4.34-4.38(m, 1H) δ 5.17-5.19(m, 1H) δ 5.30-5.35(m, 4H) δ 6.67(s, 2H) δ 8.19-8.22(t, 1H). NMR shifts are shown in the appendix.

Alkyne-PEG$_4$-DOPE: To 20 mg of DOPE dissolved in 1 ml of $CHCl_3$, 20 mg of N-hydroxysuccinimide ester_tetraoxapentadecanealkyne dissolved in 100 μl of $CHCl_3$ was added. 5 μl of triethylamine was added subsequently. The reaction mixture was allowed to go for 24 h under argon. The product was purified using HPLC starting with a gradient of 10% A (95% water : 5% methanol + 5 mM ammonium formate) and ending with 100% B (60% isopropanol : 35% methanol : 5% water + 5 mM ammonium formate). LC-MS: 1028.7. ^{1}H NMR ($CDCl_3$, 400 MHz): δ 0.84-0.87 (t, 6H) δ 1.24-1.27(d, 40H) δ 1.56(s, 4H) δ 1.97-2.08(d, 13H) δ 2.24-2.30(m, 4H) δ 2.47-2.50(m, 2H) δ 3.44-3.45(d, H) (m, 1H) δ 3.60-3.72(d, 14H) δ 3.77-3.80(t, 2H) δ 3.94-3.98(q, 4H) δ 4.13-4.18(m, 1H) δ 4.26-4.27(d, 2H) δ 4.35-4.39(m, 1H) δ 5.18-5.20(m, 1H) δ 5.30-5.35(m, 4H) δ 7.00(s, 3H) δ 8.17-8.18(s, 1H). NMR shifts are shown in the appendix.

Electron microscopy

Negatively stained specimens were imaged in an FEI Tecnai 12 electron microscope equipped with an LaB_6 filament and operated at an acceleration voltage of 120 kV. Images were recorded on a 2K x 2K CCD camera under low-dose conditions at a nominal magnification of 42,000x and a defocus value of about -1.5 μm.

Image processing

Using Boxer, the display program associated with the EMAN software package (Ludtke *et al.*, 1999), 3000 particles were selected from images of negatively stained TRAPPII particles captured on GST-based Affinity Grids. Using the SPIDER software package (Frank *et al.,* 1996), the particles were subjected to 10 cycles of multireference alignment. Each round of multireference alignment was followed by K-means classification specifying 50 output classes. The references used for the first multireference alignment were randomly chosen from the raw images.

Chapter 5: Conclusions

Membrane proteins, like TRPV and MP20, are responsible for most of the biological functions performed by cell membranes and mediate all the interactions between cells and their environment. TRPV channels sense temperature differences in the environment and allow an influx of Ca^{2+} ions into keratinocytes and neurons to communicate that information to the rest of the body (Venkatachalam and Montell, 2004). MP20, a member of the PMP22/MP20/EMP/claudin superfamily, has been demonstrated to be necessary for the proper formation of a transparent, refractive lens in the mammalian eye (Steele *et al.,* 1997), and mutations in MP20 lead to cataracts (Steele *et al.,* 1997; Pras *et al.*, 2002; Ponnam *et al.,* 2008). Given the importance of membrane proteins such as TRPV and MP20, it is essential to understand how these proteins perform their functions at the atomic level.

Obtaining structural data of membrane proteins can be very challenging, which is mainly due to their amphiphilic nature that makes them unstable in solution. Therefore, to study a membrane protein in solution, it is either necessary to remove its hydrophobic portion or to solubilize it with a lipid mimetic. TRPV channels consist of a transmembrane domain, with a topology similar to that of the Shaker K^+ channel, and two cytosolic domains, one at the N terminus and one at the C terminus (Venkatachalam and Montell, 2004). Since crystallizing the full-length TRPV channel has proved difficult and since the topology of the transmembrane domain is essentially known, we decided to focus on studying the cytoplasmic domains to understand the mechanism of TRPV channels. This strategy was previously applied to many other membrane proteins, including, for example, TAP transporters (Procko *et al.*, 2006; Procko *et al.*, 2008) and the human epidermal growth factor receptor HER3/ErbB3 (Jura *et al.*, 2009). However, this approach is only useful if the cytosolic domain has a role in the overall function of the protein and if the cytosolic domain can be studied in isolation from the rest of the membrane protein. MP20 is an example of a protein, in which case this strategy is not applicable. Most of MP20 is embedded within the membrane, and it is currently unclear whether and how the extramembranous domains contribute to its function. Hence, studies on MP20 needed to be performed on the full-length protein. MP20 was expressed, purified and reconstituted into proteoliposomes, which allowed the function of MP20 to be studied *in vitro*. This approach is commonly used to perform functional studies on membrane proteins, and examples include ABC transporters (Geertsma *et al.*, 2008) and influenza A virus M2 proton channels (Moffat *et al.*, 2008).

Using the two strategies described above, we were able to obtain biochemical and structural information for both the TRPV channels and MP20. In the case of TRPV channels, we used our crystal structure data of the cytosolic ankyrin repeat domains to understand the different sensitization mechanisms between thermoTRPVs, insights which led us to hypothesize that differences in the physiological function between these channels lie in their interactions with Ca^{2+}-calmodulin and ATP. We subsequently confirmed our structure-based predictions by electrophysiological studies, which showed that binding of Ca^{2+}-calmodulin to the N-terminal ankyrin repeat domains desensitizes

TRPV1, TRPV3 and TRPV4 to stimulus, whereas binding of ATP sensitizes TRPV1 and TRPV4 to stimulus but desensitizes TRPV3.

In the case of MP20, we were able to reconstitute recombinant MP20 into proteoliposomes, thus providing the means for future functional studies of MP20 *in vitro*. Using this assay, we already ruled out the notion that MP20 forms membrane junctions, alone or in conjunction with galectin-3. In addition, we optimized the conditions for large-scale expression of recombinant MP20 protein, thus setting the stage for systematic screening of crystallization conditions. Preliminary screens yielded crystals that were presumably formed by detergent, but many crystallization conditions remain untested and may eventually lead to the formation of MP20 crystals. Obtaining an atomic model of MP20 will not only provide molecular insights into its structure and function within the lens of mammalian eyes, but it may also enable us to understand the properties of other members of the PMP22/MP20/EMP20/claudin superfamily, many of which have been implicated in cancers and neuropathies (Van Itallie *et al.,* 2006).

My biochemical and structural studies of TRPV channels and MP20 met with many hurdles. Indeed, successful biochemical preparation of protein samples can be challenging, arduous and time-consuming. Despite the development of several new approaches to biochemistry and recent advances in structural biology techniques, obtaining structural information remains a challenge for many proteins. Single-particle EM may be the structural approach that has the least requirements for samples. The amount of material needed for structural analysis is small compared with other structural biology techniques. Sample heterogeneity also poses less of problem, and single-particle EM can indeed be used to capture multiple conformations of the target protein. Moreover, the monolayer purification technique (Kelly *et al.*, 2008a) and its derivative, the Affinity Grid (Kelly *et al.*, 2008b), were developed to further simplify specimen preparation for single-particle EM by combining purification and EM grid preparation into a single step. Ideally, both the monolayer purification technique and the Affinity Grid should make it possible to capture His-tagged protein complexes specifically out of cell lysate for visualization under the electron microscope. However, the methods have been limited by the large degree of non-specific protein adsorption to the lipid monolayers. Although the incorporation of PEG lipids into the monolayer appears to reduce non-specific protein adsorption, it remains to be tested whether this advance suffices to specifically capture only the His-tagged proteins from cell lysates. It is also currently unclear whether this technique will be applicable to membrane proteins. Nevertheless, in my future work I aim to further develop affinity techniques for specimen preparation in order to make the use of single-particle EM more widely amenable to both biochemists and cell biologists, and to fuel more rapid and deeper mechanistic understandings of protein function.

Bibliography

Abramoff MD, Magelhaes PJ, and Ram SJ (2004) Image processing with Image J. *Biophotonics International* 11: 36–42

Accardi A (2008) To ATP or not to ATP: this is the question *Journal of General Physiology* 131: 105–108

Alcalá J, Lieska N and Maisel H (1975) Protein composition of bovine lens cortical fiber cell membranes. *Experimental Eye Research* 21(6): 581-595

Artavanis-Tsakonas S, Matsuno K and Fortini ME (1995) Notch signaling *Science* 268(5208): 225-232

Ataullakhanov F, and Vitvitsky V (2002) What determines the ATP intracellular concentration. *Bioscience Reports* 22: 501–511

Berardi MJ, Shih WM, Harrison SC and Chou JJ (2011) Mitochondrial uncoupling protein 2 structure determined by NMR molecular fragment searching. *Nature* 476(7358): 109-113

Beck B, Lehen'kyi V, Roudbaraki M, Flourakis M, Charveron M, Bordat P, Polakowska R, Prevarskaya N, and Skryma R (2008) TRPC channels determine human keratinocyte differentiation: new insight into basal cell carcinoma. *Cell Calcium* 43: 492-505.

Binz HK, Stumpp MT, Forrer P, Amstutz P, and Pluckthun A (2003) Designing repeat proteins: well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins. *Journal of Molecular Bioloogy* 332: 489-503.

Boyd AW and Lackmann M (2001) Signals from Eph and ephrin proteins: a developmental tool kit. Sci STKE *2001*.

Blankenship RE (2008) *Molecular Mechanisms of Photosynthesis* (2nd edition). John Wiley & Sons Inc.

Bowie JU (2001) Stabilizing membrane proteins. *Current Opinion in Structural Biology* 11(4): 397-402

Boyer PD (1995) From human serum albumin to rotational catalysis by ATP synthase. *FASEB Journal* 9(7): 559–561

Bryan J, Mun˜oz A, Zhang X, Du¨fer M, Drews G, Krippeit-Drews P and Aguilar-Bryan L (2007) ABCC8 and ABCC9: ABC transporters that regulate K+ channels. *Pflugers Archives* 453: 703–718

Caterina MJ (2007) Transient receptor potential ion channels as participants in thermosensation and thermoregulation. *American Journal of Physiology Regulatory Integrative Comparative Physiology* 292(1): R64-R76

Celia H, Wilson-Kubalek E, Milligan RA and Teyton L (1999) Structure and function of a membrane-bound murine MHC class I molecule. *Proceedings of the National Academy of Sciences of the United States of America* 96(10): 5634-5639

Chang Q, Gyftogianni E, van de Graaf SFJ, Hoefs S, Weidema FA, Bindels RJM, and Hoenderop JGJ (2004) Molecular determinants in TRPV5 channel assembly. *Jounal of Biological Chemistry* 279: 54304-54311.

Chaves AV, Baah J, Wang Y, McAllister TA, and Benchaar C (2012) Effects of cinnamon leaf, oregano and sweet orange essential oils on fermentation and aerobic stability of barley silage. *Journal of Science Food and Agriculture* 92: 906-915.

Chen T, Li X, Yang Y, Erdene AG and Church RL (2003) Does lens intrinsic membrane protein MP19 contain a membrane-targeting signal? *Molecular Vision* 9: 735-746

Cheng Y and Walz T (2009) The advent of near-atomic resolution in single-particle electron microscopy. *Annual Reviews in Biochemistry* 78: *723-742*

Chiu W, Avila-Sakar AJ and Schmid MF (1997) Electron crystallography of macromolecular periodic arrays on phospholipid monolayers. *Advances in Biophysics* 34: 161-172

Chung MK, Lee H, Mizuno A, Suzuki M, and Caterina MJ (2004) 2-aminoethoxydiphenyl borate activates and sensitizes the heat-gated ion channel TRPV3. *Journal of Neuroscience* 24: 5177–5182

Chung MK, Gu¨ler AD, and Caterina MJ (2005) Biphasic currents evoked by chemical or thermal activation of the heat-gated ion channel, TRPV3. *Journal of Biological Chemistry* 280: 15928–15941

Clapham DE (2003) TRP channels as cellular sensors. *Nature* 426: 517-524.

Clapham DE, Julius D, Montell C, and Schultz G (2005) International Union of Pharmacology. XLIX. Nomenclature and structure-function relationships of transient receptor potential channels. *Pharmacological Review* 57: 427-450.

Clapham DE (2007) Calcium signaling. *Cell* 131(6): 1047-1058

Collins TJ (2007) ImageJ for microscopy. *BioTechniques* 43: 25–30

Conboy JC, Liu S, O'Brien DF and Saavedra SS (2003) Planar supported bilayer polymers formed from bis-diene lipids by Langmuir-Blodgett deposition and UV irradiation. *Biomacromolecules* 4(3): 841-849

Coudray N, Hermann G, Caujolle-Bert D, Karathanou A, Erne-Brand F, Buessler JL, Daum P, Plitzko JM, Chami M, Mueller U, Kihl H, Urban JP, Engel A and Rémigy HW (2011) Automated screening of 2D crystallization trials using transmission electron microscopy: a high-throughput tool-chain for sample preparation and microscopic analysis. *Journal of Structural Biology* 173(2): 365-374

Dattilo M, Penington NJ, and Williams K (2008) Inhibition of TRPC5 channels by intracellular ATP.*Molecular Pharmacology* 73: 42–49

Darst SA, Ribi HO, Pierce DW and Kornberg RD (1988) Two-dimensional crystals of Escherichia coli RNA polymerase holoenzyme on positively charged lipid layers. *Journal of Molecular Biology* 203(1): 269-273

Darst SA, Kubalek EW, Edwards AM and Kornberg RD (1991) Two-dimensional and epitaxial crystallization of a mutant form of yeast RNA polymerase II. *Journal of Molecular Biology* 221(1): 347-257

den Dekker E, Hoenderop JGJ, Nilius B and Bindels RJM (2003) The epithelial calcium channels, TRPV5 & TRPV6: from identification towards regulation. *Cell Calcium* 33: 497-507.

Derler I, Hofbauer M, Kahr H, Fritsch R, Muik M, Kepplinger K, Hack ME, Moritz S, Schindl R, Groschner K, and Romanin C (2006) Dynamic but not constitutive association of calmodulin with rat TRPV6 channels enables fine tuning of Ca^{2+}-dependent inactivation. *Journal of Physiology* 577: 31-44.

Dori Y, Bianco–Peled H, Satija SK, Fields GB, McCarthy JB and Tirrell M (2000) Ligand accessibility as means to control cell response to bioactive bilayer membranes. *Journal of Biomedical Material Research* 50(1): 75-81

Erler I, Hirnet D, Wissenbach U, Flockerzi V, and Niemeyer BA (2004) Ca^{2+}-selective transient receptor potential V channel architecture and function require a specific ankyrin repeat. *Journal of Biological Chemistry* 279: 34456-34463

Ervin LA, Ball LE, Crouch RK and Schey KL (2005) Phosphorylation and glycosylation of bovine lens MP20**.** *Investigative Ophthalmology and Visual Science* 46(2): 627-635

Faham S and Bowie JU (2002) Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. *Journal of Molecular Biology* 316(1): 1-6

Ferreiro DU, Cervantes CF, Truhlar SME, Cho SS, Wolynes PG and Komives EA (2007)

Stabilizing IkappaBalpha by "consensus" design. *Journal of Molecular Biology* 365: 1201-1216.

Fixemer T, Wissenbach U, Flockerzi V, and Bonkhoff H (2003) Expression of the Ca^{2+}-selective cation channel TRPV6 in human prostate cancer: a novel prognostic marker for tumor progression. *Oncogene* 22: 7858-7861

Frank J, Radermacher M, Penczek P, Zhu J, Li Y, Ladjadj M and Leith A (1996) SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *Journal of Structural Biology* 116(1): 190-199

Gaudet, R (2008) A primer on ankyrin repeat function in TRP channels and beyond. *Molecular Biosystems* 4: 372–379

Gipson B, Zeng X, Zhang ZY and Stahlberg H (2007) 2dx--user-friendly image processing for 2D crystals. *Journal of Structural Biology* 157(1): 64-72

Gipson B, Zeng X and Stahlberg H (2007) 2dx_merge: data management and merging for 2D crystal images. *Journal of Structural Biology* 160(3): 375-384

Gonen T, Grey AC, Jacobs MD, Donaldson PJ and Kistler J (2001) MP20, the second most abundant lens membrane protein and member of the tetraspanin superfamily, joins the list of ligands of galectin-3. *BMC Cell Biology* 2: 17

Gonen T, Cheng Y, Kistler J and Walz T (2004) Aquaporin-0 membrane junctions form upon proteolytic cleavage. *Journal of Molecular Biology* 342(4): 1337-1345

Grey AC, Jacobs MD, Gonen T, Kistler J and Donaldson PJ (2003) Insertion of MP20 into lens fibre cell plasma membranes correlates with the formation of an extracellular diffusion barrier. *Experimental Eye Research* 77(5): 567-574

Grandl J., Hu H, Bandell M, Bursulaya B, Schmidt M, Petrus M and Patapoutian A (2008) Pore region of TRPV3 ion channel is specifically required for heat activation. *Nature Neuroscience* 11: 1007–1013

Grigorieff N and Harrison SC (2011) Near-atomic resolution reconstructions of icosahedral viruses from electron cryo-microscopy. *Current Opinion in Structural Biology* 21(2): 265-273

Güler AD, Lee H, Iida T, Shimizu I, Tominaga M and Caterina M (2002) Heat-evoked activation of the ion channel, TRPV4.*Journal of Neuroscience* 22: 6408–6414

Gupta PD, Johar K and Vasavada A (2004) Causative and preventive action of calcium in cataractogenesis. *Acta Pharmacologica Sinica* 25(10): 1250-1256

Haswell ES, Phillips R and Rees DC (2011) Mechanosensitive channels: what can they do and how do they do it? *Structure* 19(10): 1356-1369

Hellwig N, Albrecht N, Harteneck C, Schultz G, and Schaefer M (2005) Homo- and heteromeric assembly of TRPV channel subunits. *Journal of Cell Science* 118: 917-928.

Hoenderop JGJ, Voets T, Hoefs S, Weidema F, Prenen J, Nilius B and Bindels RJM (2003) Homo- and heterotetrameric architecture of the epithelial Ca^{2+} channels TRPV5 and TRPV6. *EMBO J* 22: 776-785.

Hite RK, Raunser S and Walz T (2007) Revival of electron crystallography. *Current Opinion in Structural Biology* 17(4): 389-395

Hite RK, Schenk AD, Li Z, Cheng Y and Walz T (2010) Collecting electron crystallographic data of two-dimensional protein crystals. *Methods in Enzymology* 481: 251-282

Hong M, Zhang Y and Hu F (2012) Membrane Protein Structure and Dynamics from NMR Spectroscopy *Annual Reviews of Physical Chemistry* 63(1): 1-24

Hu HZ, Gu Q, Wang C, Colton CK. Tang J, Kinoshita-Kawada M, Lee LY, Wood JD. and Zhu MX (2004) 2-aminoethoxydiphenyl borate is a common activator of TRPV1, TRPV2, and TRPV3. *Journal of Biological Chemistry* 279: 35741-35748

Hu H, Grandl J, Bandell M, Petrus M and Patapoutian A (2009) Two amino acid residues determine 2-APB sensitivity of the ion channels TRPV3 and TRPV4. *Proceedings of the National Academy of Science of the United States of America* 106: 1626-1631

Iacovache I, Biasini M, Kowal J, Kukulski W, Chami M, van der Goot FG, Engel A and Rémigy HW (2010) The 2DX robot: a membrane protein 2D crystallization swiss army knife. *Journal of Structural Biology* 169(3): 370-378

Jain A, Liu R, Ramani B, Arauz E, Ishitsuka Y, Ragunathan K, Park J, Chen J, Xiang YK and Ha T (2010) Probing cellular protein complexes using single-molecule pull-down. *Nature* 473(7348): 484-488

Jarvis LJ and Louis CF (1995) Purification and oligomeric state of the major lens fiber cell membrane proteins. *Current Eye Research* 14(9): 799-808

Jin X, Touhey J, and Gaudet R (2006) Structure of the N-terminal ankyrin repeat domain of the TRPV2 ion channel. *Journal of Biological Chemistry* 281: 25006–25010

Karp G (2008) *Cell and molecular biology: concepts and experiments* (5th edition). John Wiley & Sons Inc.

Keeler J (2005) *Understanding NMR Spectroscopy* Wiley-Blackwell

Kelly DF, Dukovski D and Walz T (2008) Monolayer purification: a rapid method for isolating protein complexes for single-particle electron microscopy. *Proceedings of the National Academy of Science of the United States of America* 105: 4703-4708.

Kelly DF, Abeyrathne PD, Dukovski D and Walz T (2008) The Affinity Grid: a prefabricated EM grid for monolayer purification. *Journal of Molecular Biology* 382: 423-433.

Kelly DF, Dukovski D and Walz T (2010) Strategy for the use of Affinity Grids to prepare non-His-tagged macromolecular complexes for single-particle electron microscopy. *Journal of Molecular Biology* 400: 675-661.

Kelly DF, Dukovski D and Walz T (2010) A practical guide to the use of monolayer purification and affinity grids. *Methods in Enzymology* 481: 83-107

Kelly DF, Lake RJ, Middlekoop TC, Fan HY, Artavanis-Tsakanos S and Walz T (2010) Molecular structure and dimeric organization of the Notch extracellular domain as revealed by electron microscopy. *PLoS One* 5: e10532.

Kijac AZ, Li Y, Sligar SG and Rienstra CM (2007) Magic-angle spinning solid-state NMR spectroscopy of nanodisc-embedded human CYP3A4. *Biochemistry* 46(48): 13696-13703

Kita-Tokarczyk K, Itel F, Grzelakowski M, Egli S, Rossbach P and Meier W (2009) Monolayer interactions between lipids and amphiphilic block copolymers. Langmuir 25(17): 9847-9856

Koplas PA, Rosenberg RL, and Oxford GS (1997) The role of calcium in the desensitization of capsaicin responses in rat dorsal root ganglion neurons. *Journal of Neuroscience* 17: 3525-3537

Kubalek EW, Le Grice SF and Brown PO (1994) Two-dimensional crystallization of histidine-tagged, HIV-1 reverse transcriptase promoted by a novel nickel-chelating lipid. *Journal of Structural Biology* 113(2): 117-123

Kussie PH, Gorina S, Marechal V, Elenbaas B, Moreau J, Levine AJ and Pavletich NP (1996) Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science* 274: 948-953.

Lambers TT, Weidema AF, Nilius B, Hoenderop JGJ and Bindels RJM (2004) Regulation of the mouse epithelial Ca^{2+} channel TRPV6 by the Ca^{2+}-sensor calmodulin. *Journal of Biological Chemistry* 279: 28855-28861.

Lameris AL, Monnens LA, Bindels RJ, and Hoenderop JGJ (2012) Drug-induced alterations in Mg2+ homoeostasis. *Clinical Science (London)* 123: 1-14.

Landau EM and Rosenbusch JP (1996) Lipidic cubic phases: A novel concept for the crystallization of membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America* 93(25): 14532-14535

Landau M, Mayrose I, Rosenberg Y, Glaser F, Martz E, Pupko T, and Ben-Tal N (2005) ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Research* 33: W299-W302

Landouré G, Zdebik AA, Martinez TL, Burnett BG, Stanescu HC, Inada H, Shi Y, Taye AA, Kong L, Munns CH, Choo SS, Phelps CB, Paudel R, Houlden H, Ludlow CL, Caterina MJ, Gaudet R, Kleta R, Fischbeck KH and Sumner CJ (2010) Mutations in TRPV4 cause Charcot-Marie-Tooth disease type 2C. *Nature Genetics* 42: 170-174

Lasic DD and Martin FJ (2005) *Stealth Liposomes* CRC Press

Lasic DD and Papahadjopoulos D (1998) *Medical Applications of Liposomes* Elsevier Science

Lee H, and Caterina MJ (2005) TRPV channels as thermosensory receptors in epithelial cells. *Pflugers Archives* 451: 160–167

Leitz AJ, Bayburt TH, Barnakov AN, Springer BA and Sligar SG (2006) Functional reconstitution of Beta2-adrenergic receptors utilizing self-assembling Nanodisc technology. *Biotechniques* 40(5): 601-606

Lévy D, Mosser G, Lambert O, Moeck GS, Bald D and Rigaud JL (1999) Two-dimensional crystallization on lipid layer: A successful approach for membrane proteins. *Journal of Structural Biology* 127(1): 44-52

Liedtke W, Choe Y, Martí-Renom, MA, Bell AM., Denis CS, Sali A, Hudspeth AJ, Friedman JM, and Heller S (2000) Vanilloid receptor-related osmotically activated channel (VR-OAC), a candidate vertebrate osmoreceptor. *Cell* 103: 525-535

Lishko PV, Procko E, Jin X, Phelps CB and Gaudet R (2007) The ankyrin repeats of TRPV1 bind multiple ligands and modulate channel sensitivity. *Neuron* 54: 905–918

Long SB, Tao X, Campbell EB and MacKinnon R (2007) Atomic structure of a voltage-dependent K+ channel in a lipid membrane-like environment. *Nature* 450(7168): 376-382

Lowe AR and Itzhaki LS (2007) Rational redesign of the folding pathway of a modular protein. *Proceedings of the National Academy of Science of the United States of America* 104: 2679-2684.

Lum L and Beachy PA (2004) The Hedgehog response network: sensors, switches, and routers. *Science* 304(5678): 1755-1759

Ludtke SJ, Baldwin PR and Chiu W (1999) EMAN: semiautomated software for high-resolution single-particle reconstructions. *Journal of Structural Biology* 128(1): 82-97

Maafi M, Mahedero MC, and Aaron JJ (1997) Fluorimetric properties of a 2-

hydroxypropyl-beta-cyclodextrin: 9-methyl-benzo[a]phenothiazine inclusion complex in aqueous media. Analytical usefulness. *Talanta* 44: 2193-2199.

Malbon CC, Wang H and Moon RT (2001) Wnt signaling and heterotrimeric G-proteins: strange bedfellows or a classic romance? *Biochemical and Biophysical Research Communications* 287(3): 589-593

McCleverty CJ, Koesema E, Patapoutian A, Lesley SA, and Kreusch,A (2006) Crystal structure of the human TRPV2 channel ankyrin repeat domain. *Protein Science* 15: 2201–2206

Meyer S., Savaresi S., Forster IC, and Dutzler R (2007) Nucleotide recognition by the cytoplasmic domain of the human chloride transporter ClC-5. *Nature Structural and Molecular Biology* 14: 60–67

Montell C (2003) The venerable inveterate invertebrate TRP channels. *Cell Calcium* 33: 409-417

Montell C, Birnbaumer L and Flockerzi V (2002) The TRP channels, a remarkably functional family. *Cell* 108: 595-598

Morris RJ, Perrakis A and Lamzin VS (2003) ARP/wARP and automatic interpretation of protein electron density maps. *Methods of Enzymology* 374: 229-244.

Mosavi LK, Cammett TJ, Desrosiers DC and Peng ZY (2004) The ankyrin repeat as molecular architecture for protein recognition. *Protein Science* 13: 1435-1448.

Mosavi LK, Minor DL and Peng ZY (2002) Consensus-derived structural determinants of the ankyrin repeat motif. *Proceedings of the National Academy of Sciences of the United States of America* 99: 16029-16034.

Myers BR, Bohlen CJ, and Julius D (2008) A yeast genetic screen reveals a critical role for the pore helix domain in TRP channel gating. *Neuron* 58: 362–373

Murshudov GN, Vagin AA and Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallography D Biological Crystallography* 53: 240-255.

Neeper M.P, Liu Y, Hutchinson TL, Wang Y, Flores CM, and Qin N (2007) Activation properties of heterologously expressed mammalian TRPV2: evidence for species dependence. *Journal of Biological Chemistry* 282: 15894-15902

Niemeyer BA, Bergs C, Wissenbach U, Flockerzi V and Trost, C (2001) Competitive regulation of CaT-like-mediated Ca^{2+} entry by protein kinase C and calmodulin. *Proceedings of the National Academy of Sciences of the United States of America* 98: 3600-3605.

Nijenhuis T, Hoenderop JG and Bindels RJ (2005) TRPV5 and TRPV6 in Ca^{2+} (re)absorption: regulating Ca^{2+} entry at the gate. *Pflugers Archives* 451: 181- 192

Nilius B, Prenen J, Tang J, Wang C, Owsianik G, Janssens A, Voets T, and Zhu MX (2005) Regulation of the Ca^{2+} sensitivity of the nonselective cation channel TRPM4. *Journal of Biological Chemistry.* 280: 6423–6433

Nilius B and Voets T (2005) TRP channels: a TR(I)P through a world of multifunctional cation channels. *Pflugers Archives* 451: 1-10.

Nilius B, Weidema F, Prenen J, Hoenderop JGJ, Vennekens R, Hoefs S, Droogmans G, and Bindels RJM. (2003) The carboxyl terminus of the epithelial Ca^{2+} channel ECaC1 is involved in Ca^{2+}-dependent inactivation. *Pflugers Archives* 445: 584-588.

Nilius B, Owsianik G, and Voets T (2008) Transient receptor potential channels meet phosphoinositides. *EMBO Journal* 27: 2809–2816

Notterpek L, Roux KJ, Amici SA, Yazdanpour A, Rahner C and Fletcher BS (2001) Peripheral myelin protein 22 is a constituent of intercellular junctions in epithelia. *Proceedings of the National Academy of Sciences of the United States of America* 98(25): 14404-14409

Owsianik G, D'Hoedt D, Voets T and Nilius B (2006) Structure-function relationship of the TRP channel superfamily. *Review of Physiology and Biochemical Pharmacology* 156: 61-90.

Palczewski K. (2012) Chemistry and Biology of Vision. *Journal of Biological Chemistry* 287(3): 1612-1619

Papahadjopoulos D, Allen TM, Gabizon A, Mayhew E, Matthay K, Huang SK, Lee KD, Woodle MC, Lasic DD and Redemann C (1991) Sterically stabilized liposomes: improvements in pharmacokinetics and antitumor therapeutic efficacy. *Proceedings of the National Academy of Science of the United States of America* 88(24): 11460-11464

Patton C, Thompson S, and Epel D (2004) Some precautions in using chelators to buffer metals in biological solutions.*Cell Calcium* 35: 427–431

Pebay-Peyroula E, Rummel G, Rosenbusch JP and Landau EM. (1997) X-ray structure of bacteriorhodopsin at 2.5 angstroms from microcrystals grown in lipidic cubic phases. *Science* 277(5332): 1676-1681

Pelletier I and Sato S. (2002) Specific recognition and cleavage of galectin-3 by Leishmania major through species-specific polygalactose epitope. *Journal of Biological Chemistry* 277(20): 17663-17670

Philippsen A, Schenk AD, Stahlberg H and Engel A (2003) Iplt--image processing library and toolkit for the electron microscopy community. *Journal of Structural Biology* 144(1-2): 4-12

Philippsen A, Schenk AD, Signorell GA, Mariani V, Berneche S, Engel A (2007) Collaborative EM image processing with the IPLT image processing library and toolbox. *Journal of Structural Biology* 157(1): 28-37

Peier AM, Reeve AJ, Andersson DA, Moqrich A, Earley TJ, Hergarden AC, Story GM, Colley S, Hogenesch JB, McIntyre P, Bevan S, and Patapoutian A (2002) A heat-sensitive TRP channel expressed in keratinocytes. *Science* 296: 2046–2049

Perrakis A, Morris R, and Lamzin VS (1999) Automated protein model building combined with iterative structure refinement. *Nature Structural Biology* 6: 458-463.

Perrakis A, Sixma TK, Wilson KS and Lamzin VS (1997) wARP: improvement and extension of crystallographic phases by weighted averaging of multiple-refined dummy atomic models. *Acta Crystallography D Biological Crystallography* 53: 448-455.

Phelps CB, Procko E, Lishko PV, Wang RR and Gaudet R (2007). Insights into the roles of conserved and divergent residues in the ankyrin repeats of TRPV ion channels. *Channels (Austin)* 1: 148-151.

Phelps CB, Huang RJ, Lishko PV, Wang RR and Gaudet R (2008) Structural analyses of the ankyrin repeat domain of TRPV6 and related TRPV ion channels. *Biochemistry* 47: 2476–2484

Phelps CB, Wang RR, Choo SS and Gaudet R (2010) Differential regulation of TRPV1, TRPV3, and TRPV4 sensitivity through a conserved binding site on the ankyrin repeat domain. *Journal of Biological Chemistry* 285: 731-740

Pike JW, Meyer MB, Watanuki M, Kim S, Zella LA, Fretz JA, Yamazaki M, and Shevde NK (2007) Perspectives on mechanisms of gene regulation by 1,25-dihydroxyvitamin D3 and its receptor. *Journal of Steroid Biochemistry and Molecular Biology* 103: 389-395.

Plisson C, Drucker M, Blanc S, German-Retana S, Le Gall O, Thomas D and Bron P (2003) Structural characterization of HC-Pro, a plant virus multifunctional protein. *Journal of Biological Chemistry* 278(26): 23753-27361

Ponnam SP, Ramesha K, Tejwani S, Matalia J and Kannabiran C (2008) A missense mutation in LIM2 causes autosomal recessive congenital cataract. *Molecular Vision* 14: 1204-1208

Popot JL (2010) Amphipols, nanodiscs, and fluorinated surfactants: three nonconventional approaches to studying membrane proteins in aqueous solutions. *Annual Reviews in Biochemistry* 79: 737-775

Potterton L, McNicholas S, Krissinel E, Gruber J, Cowtan K, Emsley P, Murshudov GN, Cohen S, Perrakis A and Noble M (2004) Developments in the CCP4 molecular-graphics project. *Acta Crystallography D Biological Crystallography* 60: 2288-2294.

Pras E, Levy-Nissenbaum E, Bakhan T, Lahat H, Assia E, Geffen-Carmi N, Frydman M, Goldman B and Pras E (2002) A missense mutation in the LIM2 gene is associated with autosomal recessive presenile cataract in an inbred Iraqi Jewish family. *American Journal of Human Genetics* 70(5): 1363-1367

Prime KL and Whitesides GM (1991) Self-assembled organic monolayers: model systems for studying adsorption of proteins at surfaces. *Science* 252: 1164-1167.

Prosser RS, Evanics F, Kitevski JL and Al-Abdul-Wahid MS (2006) Current applications of bicelles in NMR studies of membrane-associated amphiphiles and proteins. *Biochemistry* 45(28): 8453-8465.

Puk O, Ahmad N, Wagner S, de Angelis MH and Graw J (2011) Microphakia and congenital cataract formation in a novel Lim2(C51R) mutant mouse. *Molecular Vision* 17: 1164-1171

Raschle T, Hiller S, Yu TY, Rice AJ, Walz T and Wagner G (2009) Structural and functional characterization of the integral membrane protein VDAC-1 in lipid bilayer nanodiscs. *Journal of American Chemical Society* 131(49): 17777-17779

Raunser S and Walz T (2009) Electron crystallography as a technique to study the structure on membrane proteins in a lipidic environment. *Annual Reviews in Biophysics* 38: 89-105

Ribi HO, Reichard P and Kornberg RD (1987) Two-dimensional crystals of enzyme-effector complexes: ribonucleotide reductase at 18-A resolution. *Biochemistry* 26(24): 7974-7979

Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Séraphin B (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nature Biotechnology* 17(10): 1030-1032

Roux KJ, Amici SA and Notterpek L (2004) The temporospatial expression of peripheral myelin protein 22 at the developing blood-nerve and blood-brain barriers. *The Journal of Comparative Neurology* 474(4): 578-588

Schenk AD, Castaño-Díez D, Gipson B, Arheit M, Zeng X and Stahlberg H. (2010) 3D reconstruction from 2D crystal image and diffraction data. *Methods in Enzymology* 482: 101-129

Schmidt AG, Yang PL and Harrison SC (2010) Peptide inhibitors of dengue-virus entry target a late-stage fusion intermediate. *Plos Pathology* 6(4): e1000851

Schwarz EC, Wissenbach U, Niemeyer BA, Strauss B, Philipp SE, Flockerzi V and Hoth M (2006) TRPV6 potentiates calcium-dependent cell proliferation. *Cell Calcium* 39: 163-173.

Shi Y, De Maria AB, Wang H, Mathias RT, FitzGerald PG and Bassnett S (2011) Further analysis of the lens phenotype in Lim2-deficient mice. *Investigative Ophthalmology and Visual Science* 52(10): 7332-7339

Smith GD, Gunthorpe MJ, Kelsell RE, Hayes PD, Reilly P, Facer P, Wright JE, Jerman, JC, Walhin JP, Ooi L, Egerton, J, Charles KJ, Smart D, Randall AD, Anand P, and Davis, JB (2002) TRPV3 is a temperature-sensitive vanilloid receptor-like protein. *Nature* 418: 186–190

Steele EC Jr, Kerscher S, Lyon MF, Glenister PH, Favor J, Wang J and Church RL (1997) Identification of a mutation in the MP19 gene, Lim2, in the cataractous mouse mutant To3. *Molecular Vision* 3: 5

Steele EC Jr, Wang JH, Lo WK, Saperstein DA, Li X and Church RL (2000) Lim2(To3) transgenic mice establish a causative relationship between the mutation identified in the lim2 gene and cataractogenesis in the To3 mouse mutant. *Molecular Vision* 6: 85-94

Sternfeld L, Anderie I, Schmid A, Al-Shaldi H, Krause E, Magg T, Schreiner D, Hofer, HW, and Schulz I (2007) Identification of tyrosines in the putative regulatory site of the Ca^{2+}channel TRPV6. *Cell Calcium* 42: 91-102.

Strotmann R, Harteneck C, Nunnenmacher K, Schultz G, and Plant TD (2000) OTRPC4, a nonselective cation channel that confers sensitivity to extracellular osmolarity. *Nature Cell Biology* 2: 695–702

Strotmann R, Schultz G, and Plant TD (2003) Ca^{2+}-dependent potentiation of the nonselective cation channel TRPV4 is mediated by a C-terminal calmodulin binding site. *Journal of Biological Chemistry* 278: 26541–26549

Tadross MR, Dick IE, and Yue DT (2008) Mechanism of local and global Ca^{2+} sensing by calmodulin in complex with a Ca^{2+} channel. *Cell* 133: 1228–1240

Taylor JS, Vigneron DB, Murphy-Boesch J, Nelson SJ, Kessler HB, Coia L, Curran W, and Brown TR (1991) Free magnesium levels in normal human brain and brain tumors: 31P chemical-shift imaging measurements at 1.5 T. *Proceedings of the National Academy of Science of the United States of America* 88: 6810–6814

Thébault S, Cao G, Venselaar H, Xi Q, Bindels RJ and Hoenderop JG (2008) Role of the alpha-kinase domain in transient receptor potential melastatin 6 channel and regulation by intracellular ATP. *Journal of Biological Chemistry* 283: 1999–2007

Thess A, Hutschenreiter S, Hofmann M, Tampé R, Baumeister W and Guckenberger R. (2002) Specific orientation and two-dimensional crystallization of the proteasome at metal-chelating lipid interfaces. *Journal of Biological Chemistry* 277(39): 36321-36328

Tominaga M and Caterina MJ. (2004) Thermosensation and pain. *Journal of Neurobiology* 61: 3–12

Uzgiris EE and Kornberg RD (1983) Two-dimensional crystallization technique for imaging macromolecules, with application to antigen--antibody--complement complexes. *Nature* 301(5896): 125-129

Van de Graaf SFJ, Chang Q, Mensenkamp AR, Hoenderop JGJ, and Bindels RJM (2006) Direct interaction with Rab11a targets the epithelial Ca^{2+} channels TRPV5 and TRPV6 to the plasma membrane. *Molecular and Cellular Biology* 26: 303-312

Van Itallie CM and Anderson JM (2006) Claudins and epithelial paracellular transport. *Annual Reviews in Physiology* 68: 403-429

Venkatachalam K and Montell C (2007) TRP channels. *Annual Review in Biochemistry* 76: 387–417

Vennekens R, Hoenderop JG, Prenen J, Stuiver M, Willems PH, Droogmans G, Nilius B and Bindels RJ (2000) Permeation and gating properties of the novel epithelial Ca^{2+} channel. *Journal of Biological Chemistry* 275: 3963-3969.

Vennekens R, Owsianik G and Nilius B (2008) Vanilloid transient receptor potential cation channels: an overview. *Current Pharmaceutical Design* 14: 18–31

Vold RR Prosser RS and Deese AJ (1997) Isotropic solutions of phospholipid bicelles: a new membrane mimetic for high-resolution NMR studies of polypeptides. *Journal of Biomolecular NMR* (1997) 9(3): 329-335

Vriens J, Owsianik G, Voets T, Droogmans G and Nilius B (2004) Invertebrate TRP proteins as functional models for mammalian channels. *Pflugers Archives* 449: 213-226

Wallin E and von Heijne G (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sciences* 7(4): 1029-1038

Watanabe H, Davis JB, Smart D, Jerman, JC, Smith GD, Hayes P, Vriens J, Cairns W, Wissenbach U, Prenen J, Flockerzi V, Droogmans G, Benham CD and Nilius B (2002) Activation of TRPV4 channels (hVRL-2/mTRP12) by phorbol derivatives. *Journal of Biological Chemistry* 277: 13569–13577

Watanabe H, Vriens J, Prenen J, Droogmans G, Voets T and Nilius B (2003) Anandamide and arachidonic acid use epoxyeicosatrienoic acids to activate TRPV4 channels. *Nature* 424: 434–438

Xu H, Ramsey IS, Kotecha SA, Moran MM, Chong JA, Lawson D, Ge P, Lilly J, Silos-Santiago I, Xie Y, DiStefano PS, Curtis R and Clapham, DE (2002) TRPV3 is a calcium-permeable temperature-sensitive cation channel. *Nature* 418: 181–186

Xiao R, Tang J, Wang C, Colton CK, Tian J and Zhu MX (2008) Calcium plays a central role in the sensitization of TRPV3 channel to repetitive stimulations. *Journal of Biological Chemistry* 283: 6162-6174

Xu H, Delling M, Jun JC and Clapham DE (2006) Oregano, thyme and clove-derived flavors and skin sensitizers activate specific TRP channels. *Nature Neuroscience* 9: 628–635

Yip CK, Berscheminski J and Walz T (2010) Molecular architecture of the TRAPPII complex and implications for vesicle tethering. *Nature Structural and Molecular Biology* 17(11): 1298-1304

Zella LA, Meyer MB, Nerenz RD, Lee SM, Martowicz ML and Pike JW (2010).Multifunctional enhancers regulate mouse and human vitamin D receptor gene transcription. *Molecular Endocrinology* 24: 128-147

Zhang X, Borda MJ, Schoonen MA and Strongin DR (2003) Pyrite oxidation inhibition by a cross-linked lipid coating. *Geochemical Transactions* 4: 8

Zhou ZH (2008) Towards atomic resolution structural determination by single-particle cryo-electron microscopy. *Current Opinion in Structural Biology* 18(2): 218-228

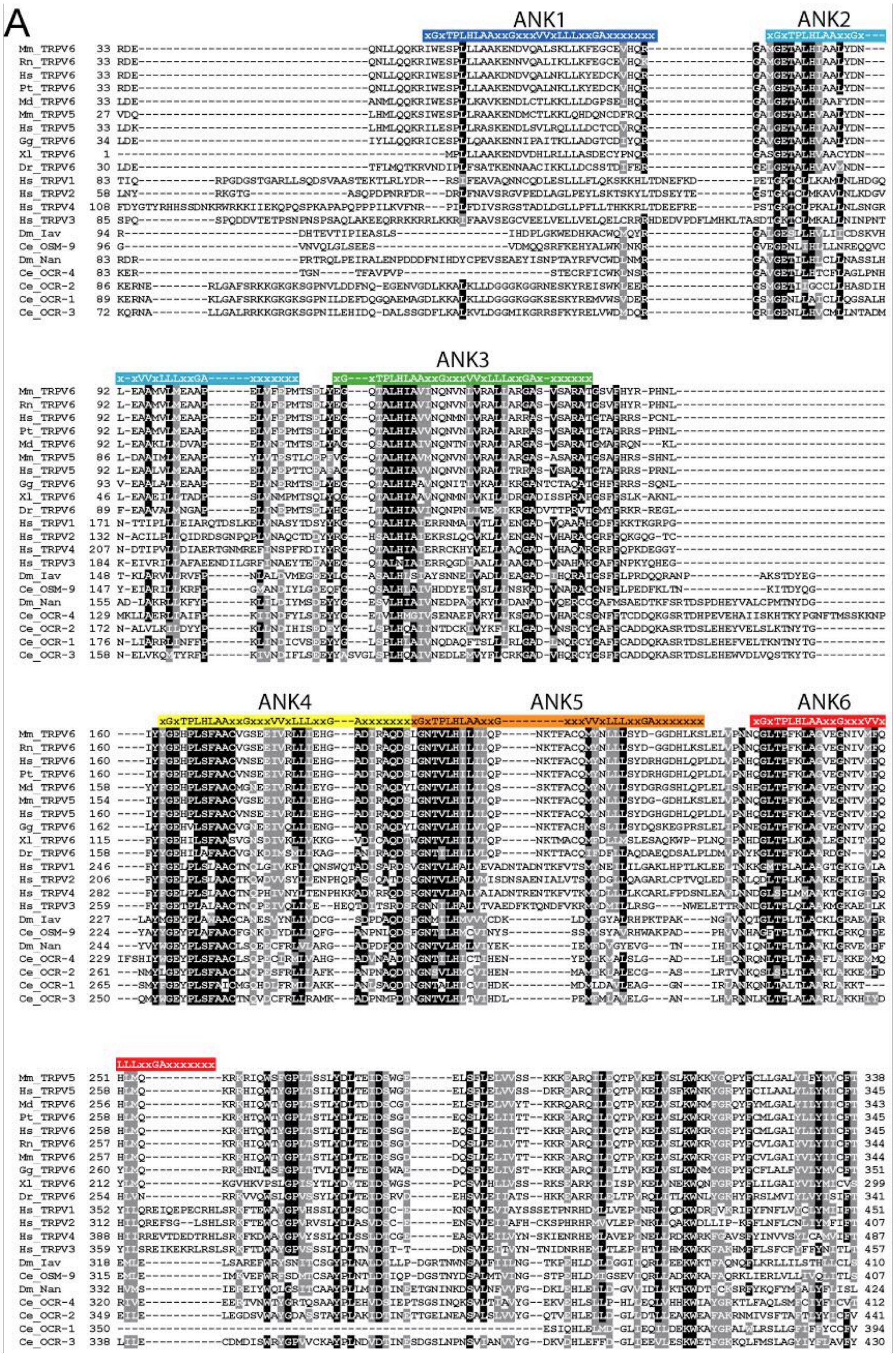
Zhu MX (2005) Multiple roles of calmodulin and other Ca^{2+}-binding proteins in the functional regulation of TRP channels. *Pflugers Archives* 451: 105–115

Zweifel ME, Leahy DJ, Hughson FM and Barrick D (2003) Structure and stability of the ankyrin domain of the Drosophila Notch receptor. *Protein Science* 12: 2622-2632.

Appendix

Supplementary Material for Chapter 2.2

Figure 2.2.5



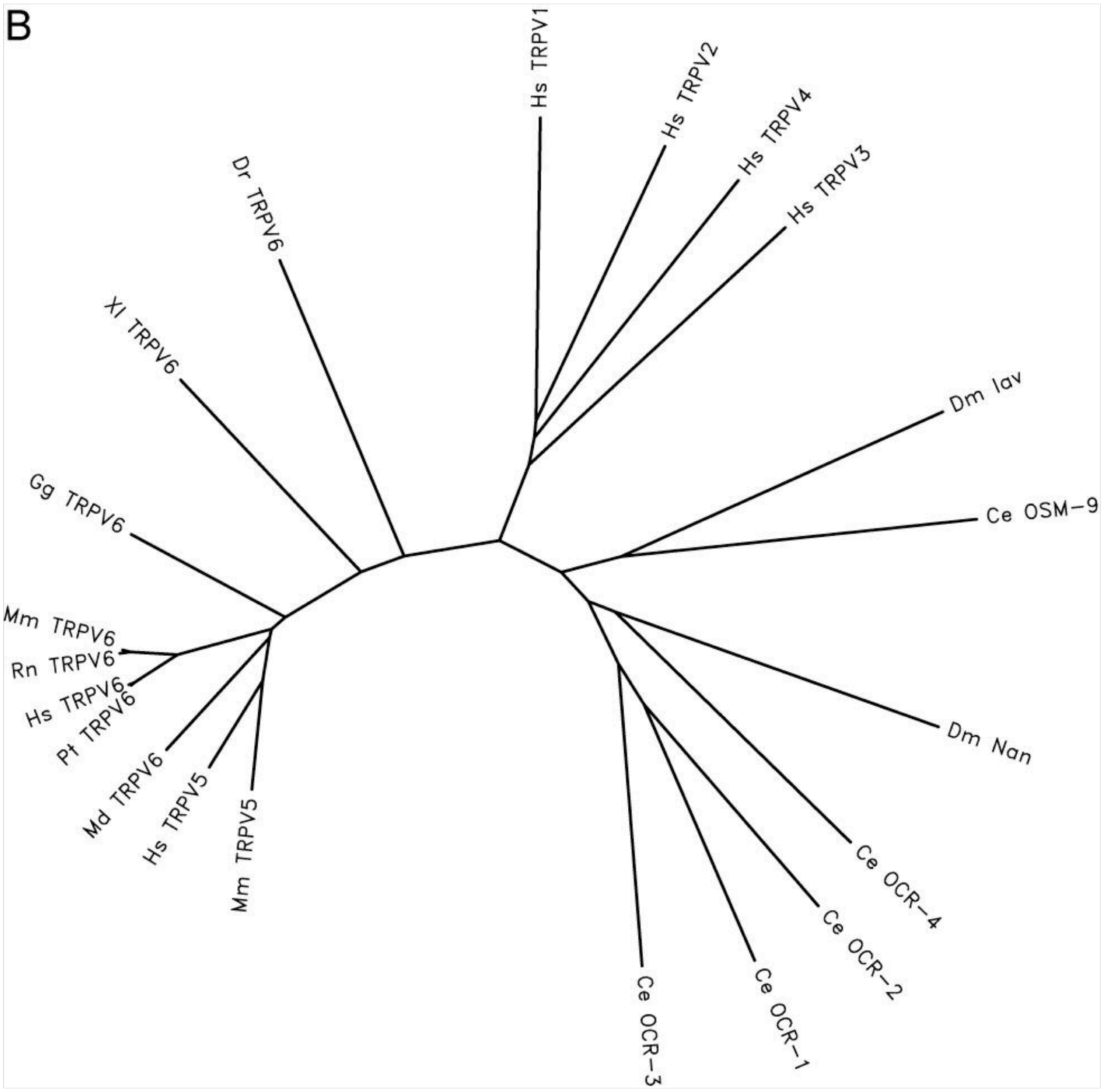
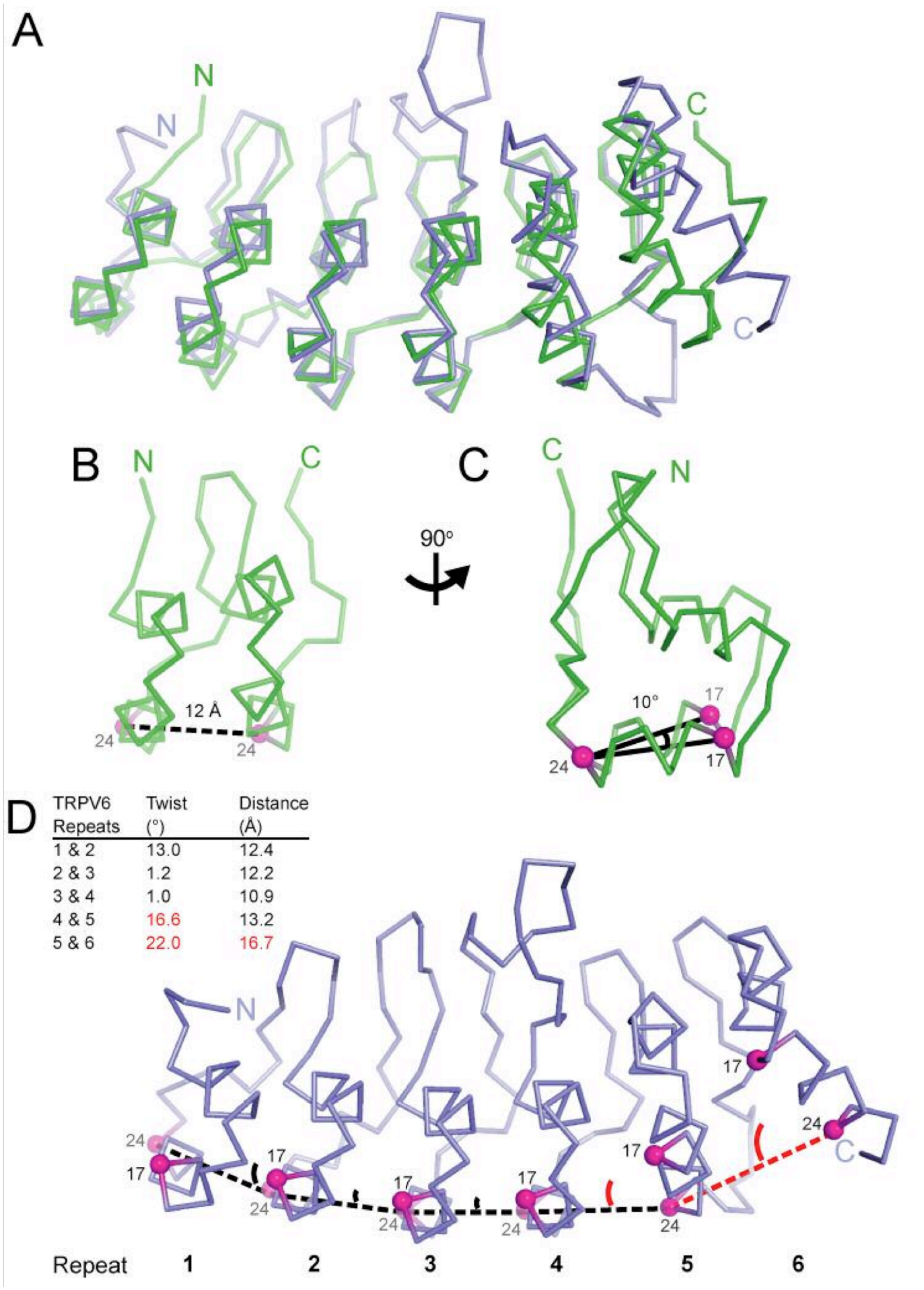


Figure 2.2.5 (continued): Alignment of TRPV family proteins. A) Sequence alignment of vertebrate an invertebrate TRPV proteins and B) corresponding Unrooted Phylogenetic Tree. The human sequence was used for mammalian TRPVs, except TRPV6, for which sequences from fish, amphibian, bird and mammals are included. The location of ankyrin repeats 1-6 are indicated above the alignment. Species used: Ce *Caenorhabditis elegans* (flatworm), Dm *Drosophila melanogaster* (fly), Dr *Danio rerio* (zebra fish), Gg *Gallus gallus* (chicken), Hs *Homo sapiens* (human), Mm *Mus musculus* (mouse), Pt *Pan*

Figure 2.2.5 (Continued) *troglodytes* (chimpanzee), Rn *Rattus norvegicus* (rat) and Xl *Xenopus laevis* (frog). The sequence analyses were performed with ClustalW and PHYLIP and colored with Boxshade.



TRPV6 Repeats	Twist (°)	Distance (Å)
1 & 2	13.0	12.4
2 & 3	1.2	12.2
3 & 4	1.0	10.9
4 & 5	16.6	13.2
5 & 6	22.0	16.7

Figure 2.2.6: Measurement of deviations from canonical ankyrin repeat geometry in TRPV6. A) Overlay of the Cα backbones of TRPV6-ARD (blue) and the central six repeats of AnkyrinR (green, amino acid residues 502-693 from PDB ID 1N11). B) The distance between repeats was measured between the Cαs at position 24 of the ankyrin repeat consensus, located at the C terminus of the outer helix. C) The twist between

Figure 2.2.6 (Continued) repeats was measured as the dihedral angle between the Cαs at positions 17 and 24 in the outer helices of consecutive ankyrin repeats. In B and C, residues 502-567 from AnkyrinR are shown as a Cα trace with the Cαs positions 24 (B) and 17 and 24 (C) shown as magenta spheres. Measurements were taken for the following ankyrin repeat protein structures in the PDB: AnkyrinR (1N11), p16 (1BI7), p18 (1IHB), p19 (1AP7), IκBα (1IKN), Bcl3 (1N1A), BABP-β (1AWC), Notch (1YYH), Gankyrin (1OUH), and three designed ARDs of three, four and five repeats (1N0Q, 1N0R and 1MJO, respectively). From these structures the average twist (dihedral) was 6.0° ± 3.3° and the average distance was 12.1 Å ± 0.9 Å. Errors listed are standard deviations. D) Inter-repeat twists and distances for TRPV6-ARD are indicated (arcs for twist and dashed lines for distance) and listed in the table on the right. Measurements that deviate significantly from the average are colored red. Positions 17 and 24 are labeled for reference, and the Cα positions are shown as magenta spheres.

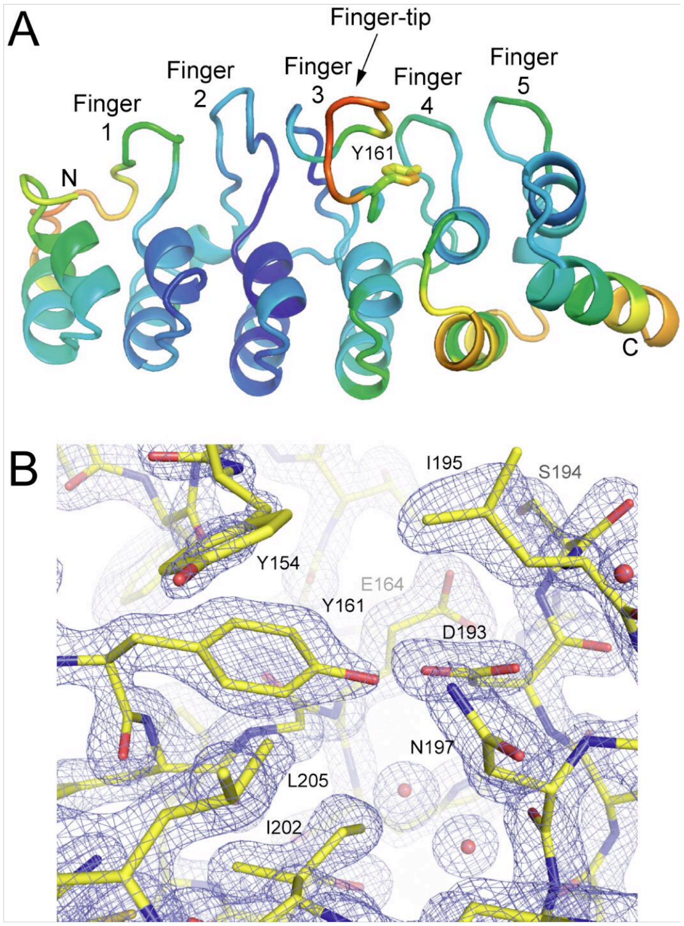


Figure 2.2.7: A) Structure of the TRPV6-ARD colored according to *B*-factors. *B*-factor values range from 25.9 Å^2 (blue) to 52.2 Å^2 (red). The side chain of Y161 (average side chain atom *B*-factor 41.6 Å^2) is shown as sticks. The average *B*-factor for residues at the base of Finger 3 (both main chain and side chain atoms), 142-152 and 161-164, is 32.8 Å^2, below the average over the entire structure, 35.4 Å^2. The average *B*-factor for the tip of Finger 3 is 46.1 Å^2, the highest of any region of the ARD. B) $2F_o$-F_c electron density map contoured at 2σ over the region of the TRPV6-ARD around Y161. Protein atoms are shown as sticks (C yellow, O red, N blue) and water molecules are shown as spheres. Note the relatively poor electron density for the side chain of Y154 (top left) located within the tip of Figure 2.2.7 compared to the others in the region. A large

Figure 2.2.7 (Continued) conformational change would be required to render Y161 accessible to a kinase. The view is rotated 90° horizontally relative to A.

Supplementary Material for Chapter 2.3

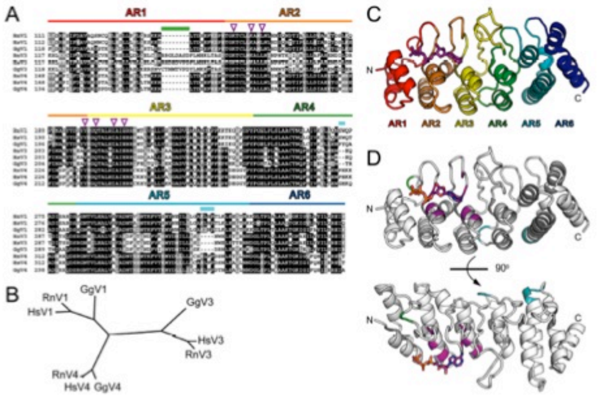


Figure 2.3.9: A) Alignment of the amino acid sequences of the ankyrin repeat domains of TRPV1 from chicken (Gg, NP_989903), human (Hs, NP_542436), and rat (Rn, NP_114188), TRPV3 from chicken (Gg, XP_001235155), human (Hs, NP_659505) and rat (Rn, NP_001020928), and TRPV4 from chicken (Gg, NP_990023), human (Hs, NP_067638) and rat (Rn, NP_076460). The alignment was colored with identical residues in black and similar residues in grey. Amino acids that contact ATP in the crystal structure of the TRPV1-ARD (PDB ID 2PNN) are indicated by purple arrows above the sequences. Colored bars above the sequence indicate the individual ankyrin repeats. An insertion and two deletions in the TRPV3-ARD compared to TRPV1 and TRPV4 are indicated by light green and teal boxes, respectively. B) Unrooted phylogenetic tree showing the relationship between the ARD sequences in A. The alignment and phylogenetic analysis were performed with ClustalW and PHYLIP as part of the SDSC Biology Workbench. C) Ribbon diagram of the structure of the TRPV1-ARD (2PNN) with individual ankyrin repeats colored according to (A). The ATP is

Figure 2.3.9 (Continued) shown as purple sticks. D) The location of the insertion and two deletions in the TRPV3-ARD is mapped onto the structure of TRPV1- ARD. The ARD is shown as a ribbon diagram with the location of ATP-binding residues in purple, and the insertion and deletions in the TRPV3-ARD colored as in (A). The ATP is shown as sticks with carbon atoms colored purple.

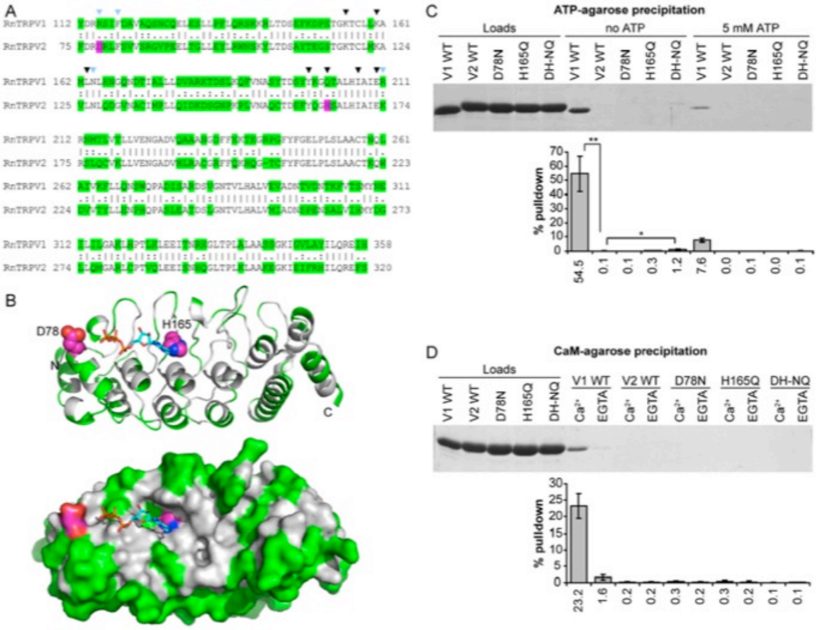


Figure 2.3.10: Surface mutations on the TRPV2-ARD do not promote ATP or calmodulin binding. A) Alignment of the rat TRPV1 and rat TRPV2 ARD sequences. Differing residues are highlighted green. Black arrowheads indicate residues within 4 Å of the ATP in the TRPV1-ARD structure; blue arrowheads point to other proximal residues. The two residues mutated in TRPV2 are highlighted pink. B) Structure of the TRPV2-ARD (PDB code 2ETB) with the ATP molecule (sticks) bound to TRPV1-ARD shown for reference. The coloring follows (A), and the two mutated sidechains are shown as spheres on a ribbon diagram at the top, whereas the molecular surface is shown at the bottom. C) Coomassie-stained gel of wild-type and mutant TRPV2-ARD loaded (left) and bound to ATP-agarose in the absence (middle) or presence (right) of competing free ATP. D) Coomassie-stained gel shows wild-type and mutant TRPV2-ARD loaded (left) and bound to CaM-agarose in the presence of Ca^{2+} or EGTA. In C) and D), wild-type TRPV1-ARD was used as a positive control and the average percentage of protein

Figure 2.3.10 (Continued) recovered (+/- standard deviation) is plotted below, and the average % pull-down value is indicated under the histogram. In C), the statistical significance of the difference in binding to ATP-agarose between TRPV1-ARD and TRPV2-ARD, and TRPV2-ARD and the DH-NQ TRPV2-ARD double mutant over three experiments are noted, with $p<0.02$ and $p<0.01$ indicated by * and **, respectively, using a one-tailed modified Student t-test.

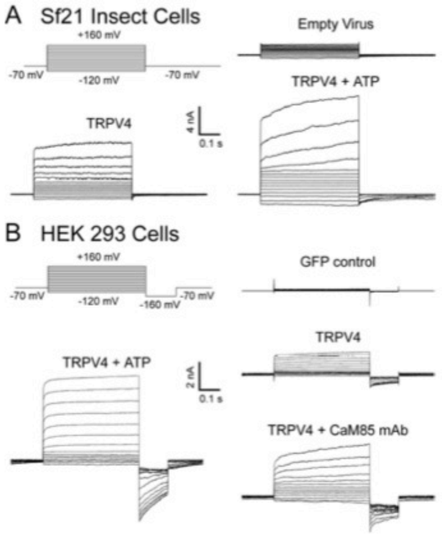


Figure 2.3.11: A) Voltage step protocol and sample recordings from insect cells infected with empty virus (top right) or virus carrying TRPV4 in the absence (lower left) or presence (lower right) of intracellular ATP. B) Voltage step protocol and sample recordings from HEK293 cells transfected with a TRPV4-expression plasmid and perfused with 5 μM 4αPDD. Recordings in the absence (middle right) or presence of intracellular ATP (left) or anti-CaM antibody (bottom right). A control cell transfected with a GFP-expression plasmid is included at the top right.

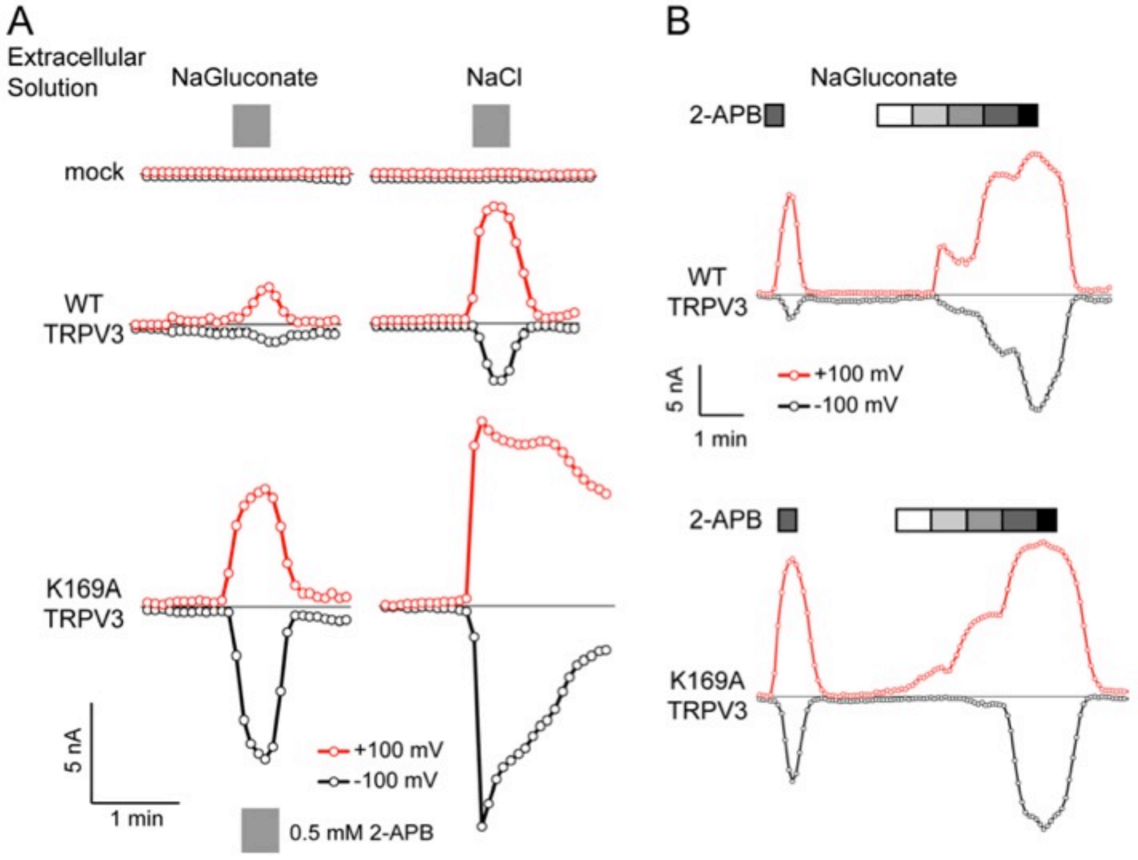


Figure 2.3.12: TRPV3 response to extracellular agonists is decreased and inactivation is faster in NaGluconate extracellular solution. A) Response of mock-infected Sf21 insect cells and insect cells expressing wild-type and K169A TRPV3 to 0.5 mM 2-APB in extracellular solutions were the primary anion is either gluconate (left) or chloride (right). For all recordings BAPTA was used as the intracellular calcium chelator. Shown are currents at +100 (grey circles) or -100 mV (black circles) extracted from linear voltage ramps from a control cell (top) and cells with intracellular ATP (bottom). Applications of 0.5 mM 2-APB are indicated by grey bars and zero current by black lines. B) Sample dose response recordings for wild type (top) and K169A (bottom). Cells were pulsed once with 2-APB to insure that the TRPV3 was pre-sensitized by BAPTA in the intracellular solutions. Recordings were carried out in

Figure 2.3.12 (Continued) NaGluconate extracellular solution to allow for timely inactivation from the test application of 2-APB. Application of increasing concentrations of 2-APB are indicated; for wild type: 0.05 mM – white, 0.25 mM – light grey, 0.5 mM – grey, 2 mM – dark grey and 4 mM – black; due to the higher sensitivity of K169A the 2-APB concentrations are half of wild type (0.025, 0.125, 0.25, 1 and 2 mM).

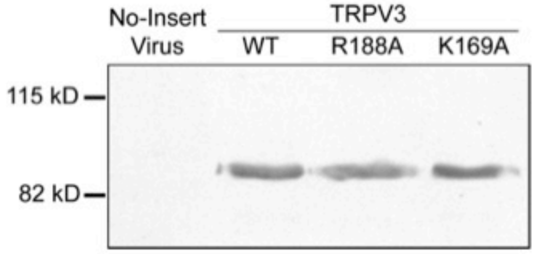
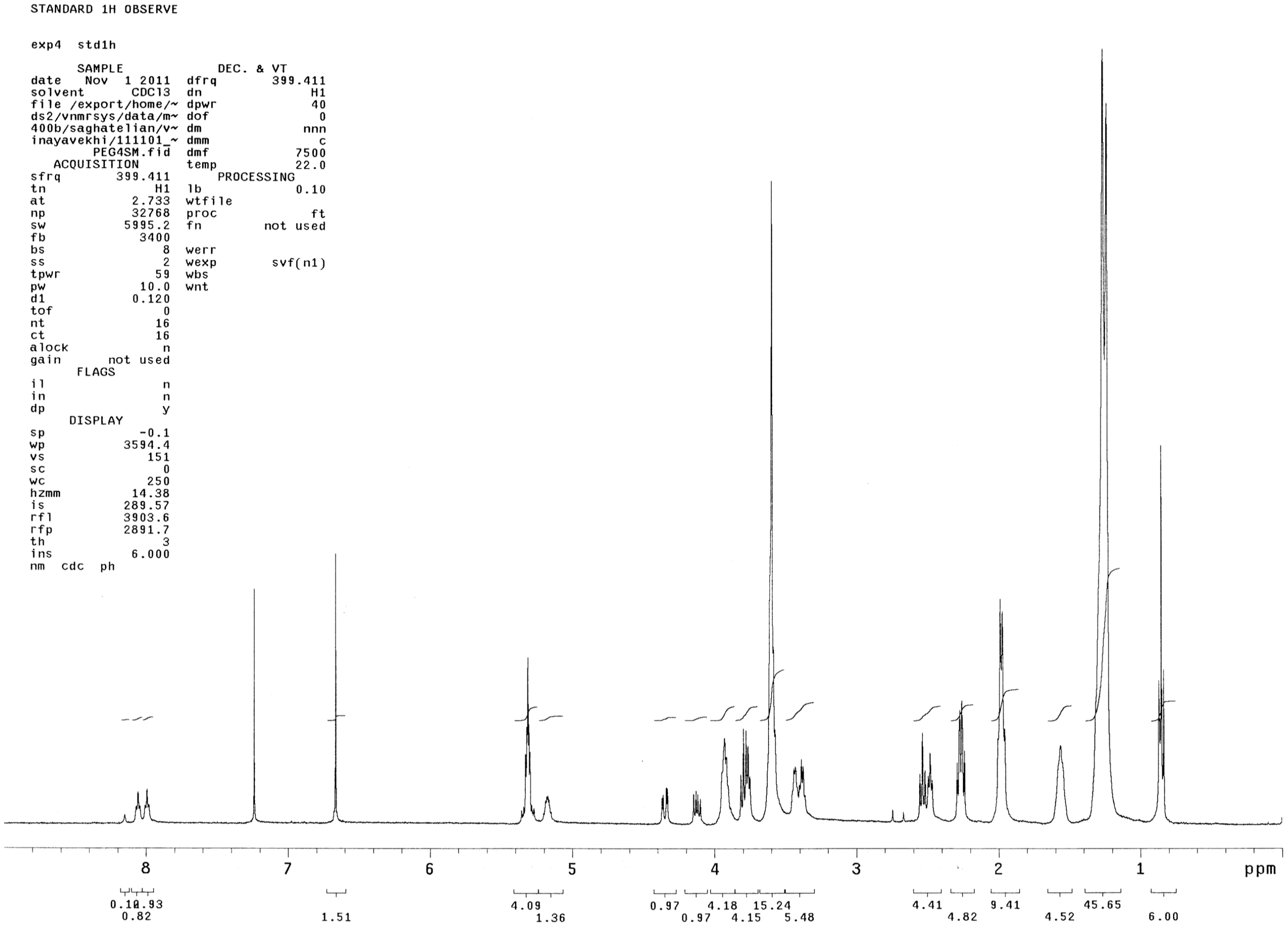


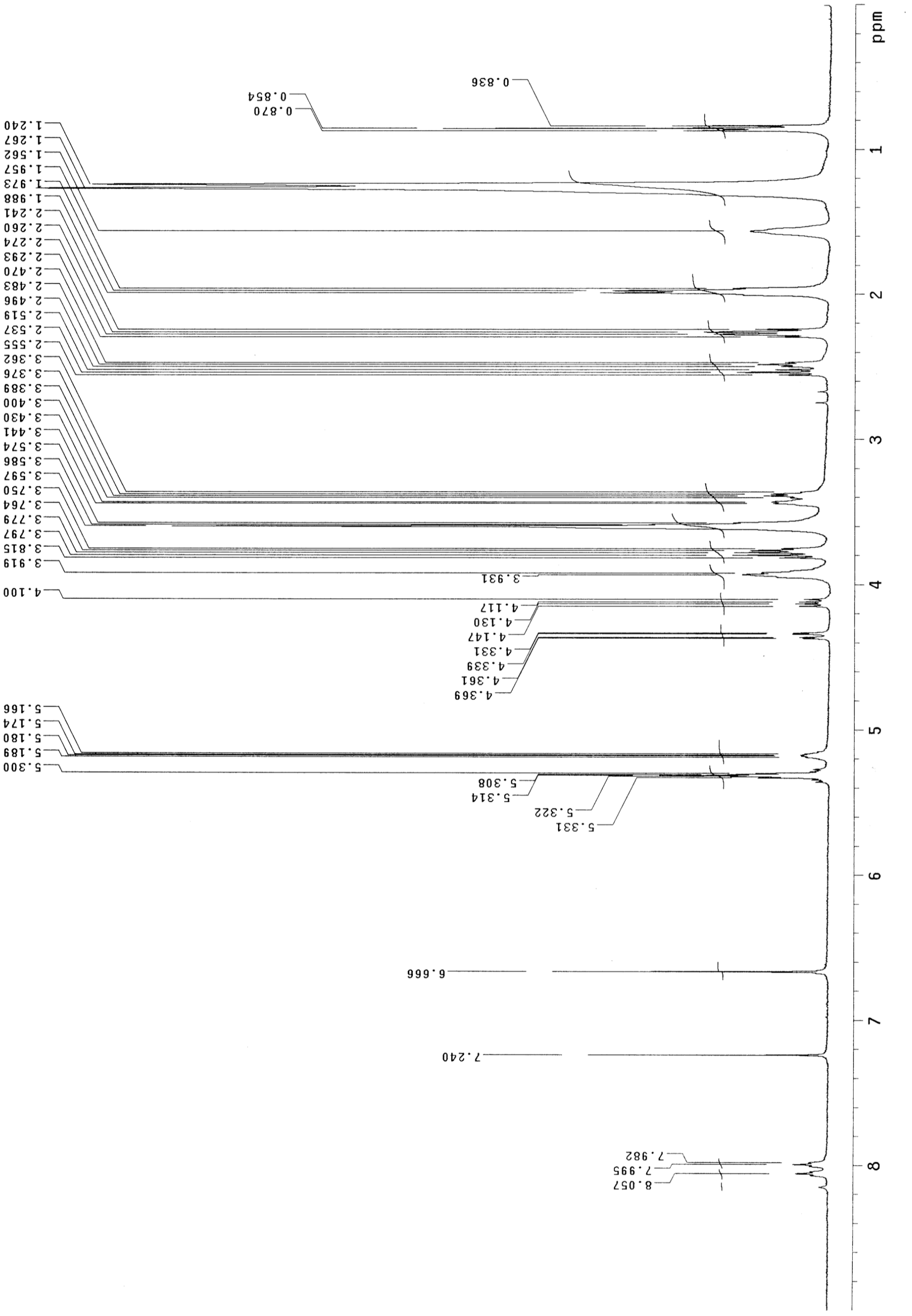
Figure 2.3.13: Wid-type, R188A and K169A TRPV3 are expressed to the same level in baculovirus-infected insect cells. Cells were harvested 48 hours after infection with baculovirus carrying FLAG-tagged TRPV3 wild-type, R188A and K169A or a no-insert control virus and samples were subjected to 10% SDS-PAGE/Western blot using the M2 anti-FLAG-alkaline phosphatase conjugated monoclonal antibody (Sigma). The position of molecular weight standards are indicated on the left.

Supplementary Material for Chapter 4

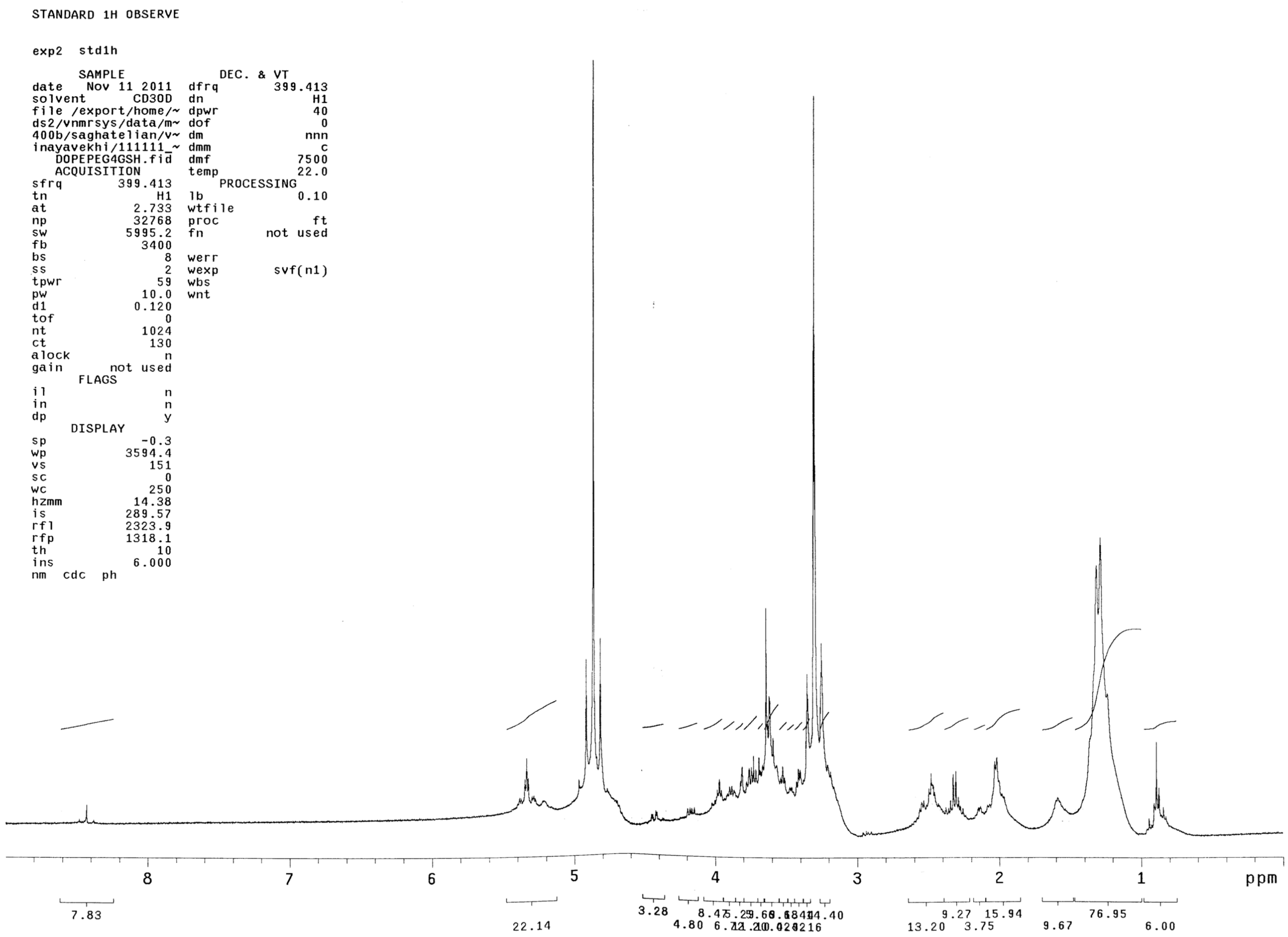
NMR Spectrum of Maleimide-PEG$_4$-DOPE



```
STANDARD 1H OBSERVE

exp4  std1h

      SAMPLE             DEC. & VT
date   Nov  1 2011  dfrq        399.411
solvent       CDCl3  dn               H1
file /export/home/~ dpwr              40
ds2/vnmrsys/data/m~ dof                0
400b/saghatelian/v~ dm               nnn
inayavekhi/111101_~ dmm                c
        PEG4SM.fid  dmf             7500
   ACQUISITION      temp            22.0
sfrq        399.411      PROCESSING
tn               H1  lb             0.10
at            2.733  wtfile
np            32768  proc             ft
sw           5995.2  fn         not used
fb             3400
bs                8  werr
ss                2  wexp        svf(n1)
tpwr             59  wbs
pw             10.0  wnt
d1            0.120
tof               0
nt               16
ct               16
alock             n
gain       not used
      FLAGS
il                n
in                n
dp                y
     DISPLAY
sp             -0.1
wp           3594.4
vs              151
sc                0
wc              250
hzmm          14.38
is           289.57
rfl          3903.6
rfp          2891.7
th                3
ins           6.000
nm  cdc  ph
```

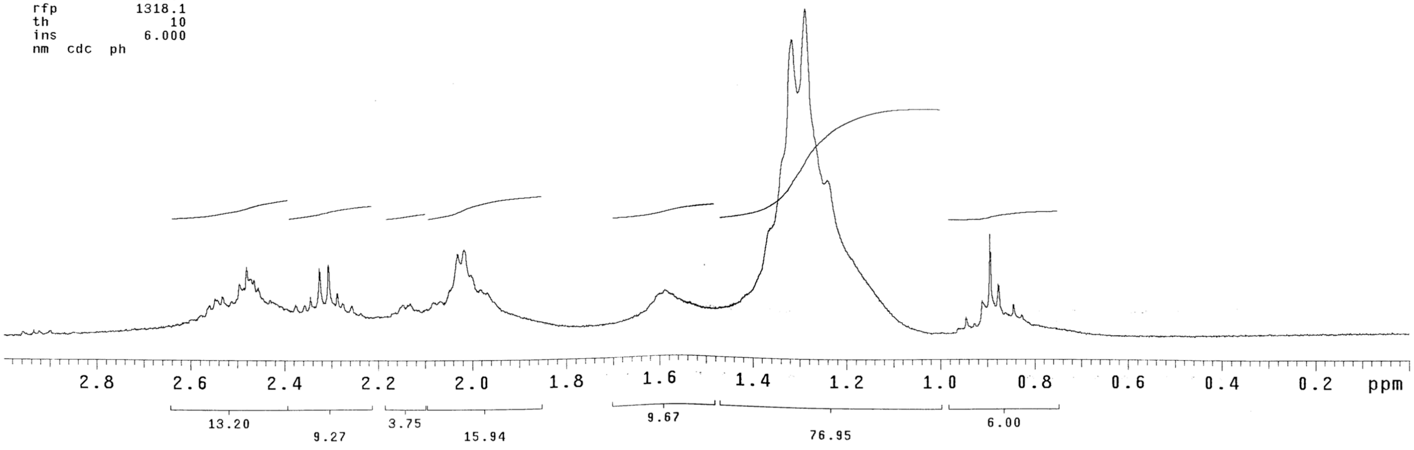

0.836
0.854
0.870
1.240
1.267
1.562
1.957
1.973
1.988
2.241
2.260
2.274
2.293
2.470
2.483
2.496
2.519
2.537
2.555
3.362
3.376
3.389
3.400
3.430
3.441
3.574
3.586
3.597
3.750
3.764
3.779
3.797
3.815
3.919
3.931
4.100
4.117
4.130
4.147
4.331
4.339
4.361
4.369
5.166
5.174
5.180
5.189
5.300
5.308
5.314
5.322
5.331
6.666
7.240
7.982
7.995
8.057
ppm
1
2
3
4
5
6
7
8

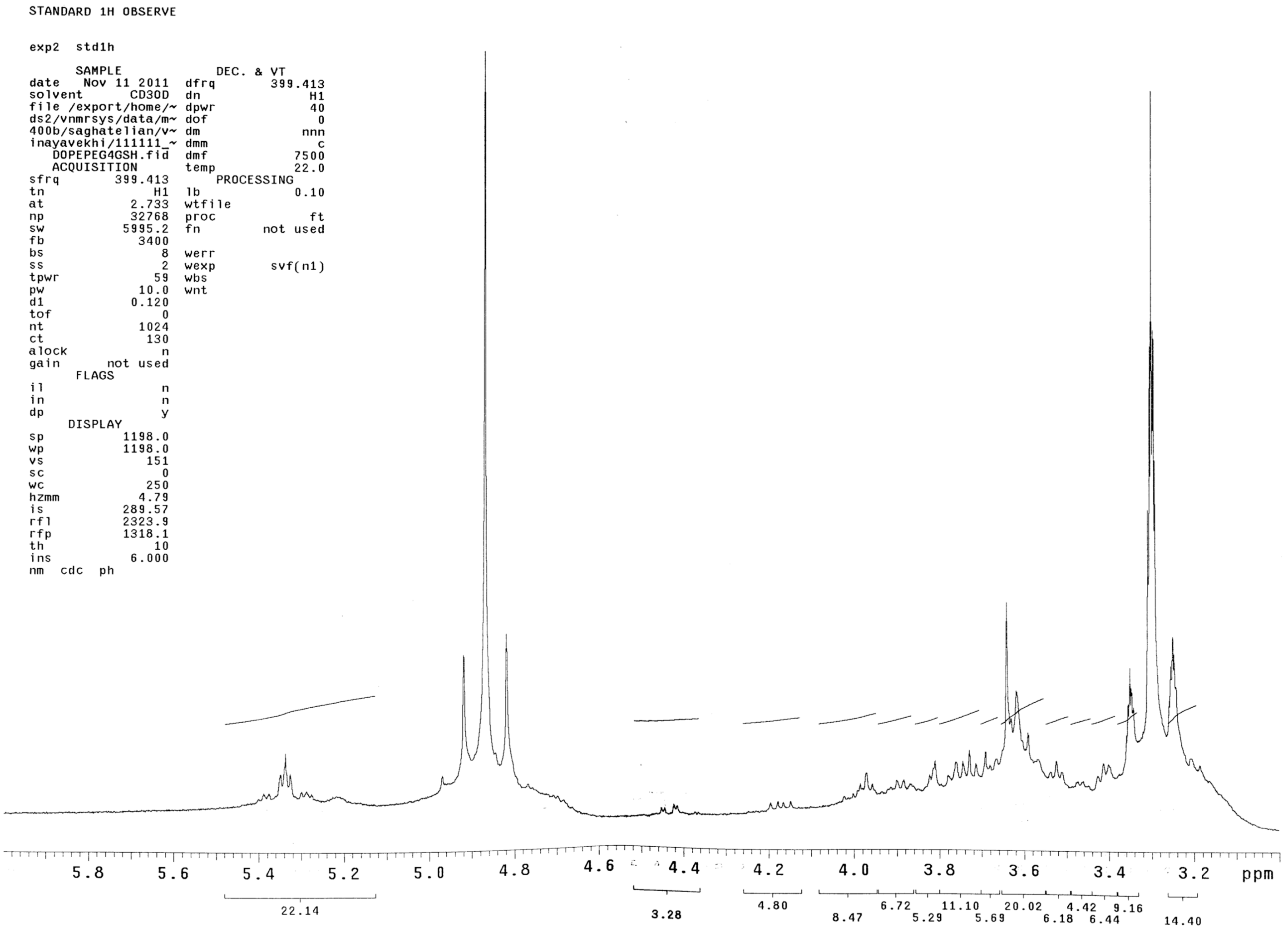
NMR Spectrum of GSH-PEG$_4$-DOPE



```
STANDARD 1H OBSERVE

exp2  std1h

      SAMPLE              DEC. & VT
date   Nov 11 2011  dfrq       399.413
solvent      CD3OD  dn              H1
file /export/home/~ dpwr            40
ds2/vnmrsys/data/m~ dof              0
400b/saghatelian/v~ dm             nnn
inayavekhi/111111_~ dmm              c
   DOPEPEG4GSH.fid  dmf           7500
   ACQUISITION      temp          22.0
sfrq       399.413      PROCESSING
tn              H1  lb            0.10
at           2.733  wtfile
np           32768  proc            ft
sw          5995.2  fn        not used
fb            3400
bs               8  werr
ss               2  wexp       svf(n1)
tpwr            59  wbs
pw            10.0  wnt
d1           0.120
tof              0
nt            1024
ct             130
alock            n
gain      not used
      FLAGS
il               n
in               n
dp               y
     DISPLAY
sp            -0.3
wp          3594.4
vs             151
sc               0
wc             250
hzmm         14.38
is          289.57
rfl         2323.9
rfp         1318.1
th              10
ins          6.000
nm  cdc  ph
```

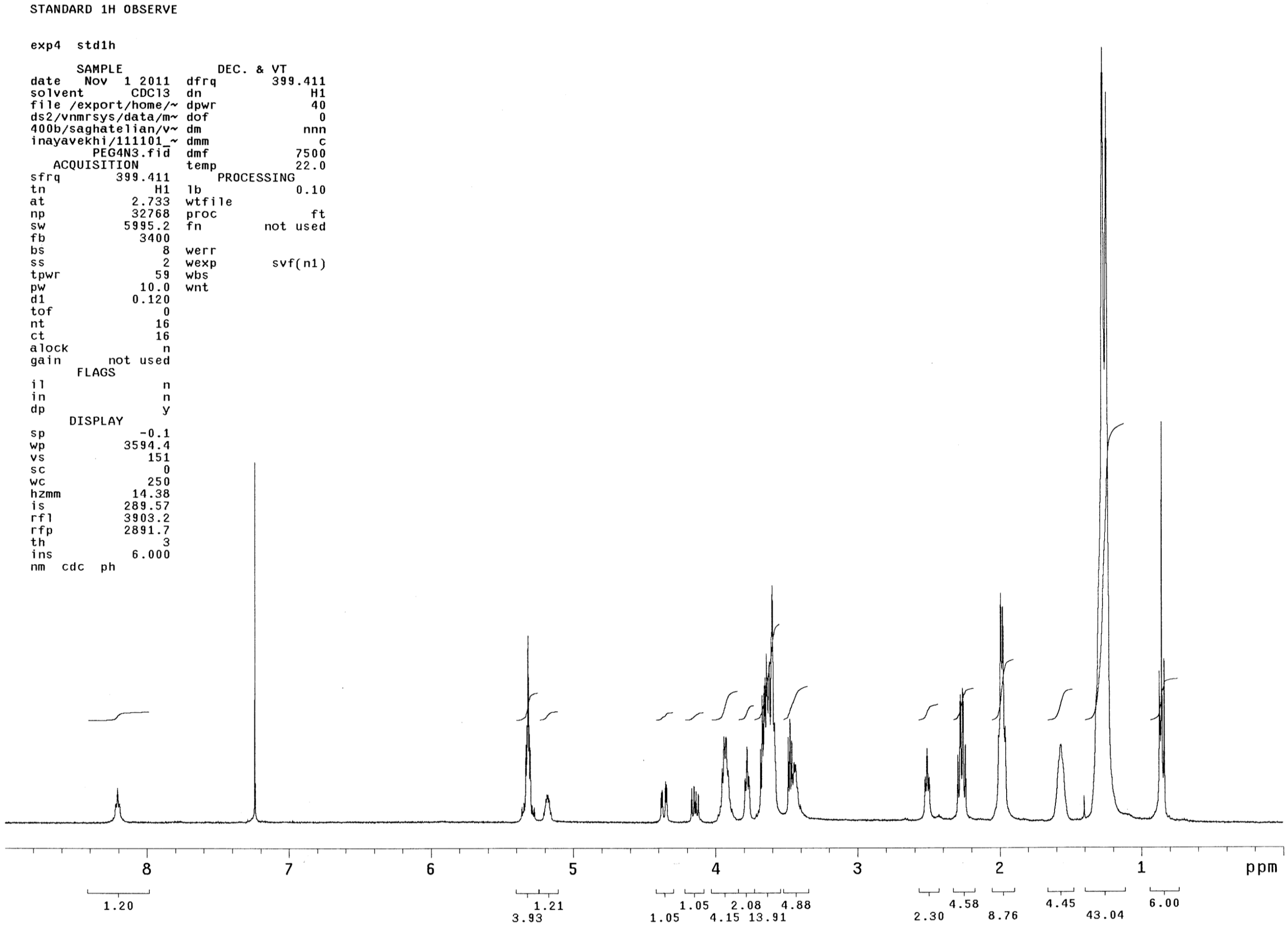
```
STANDARD 1H OBSERVE


exp2  std1h

      SAMPLE              DEC. & VT
date   Nov 11 2011  dfrq        399.413
solvent      CD3OD  dn               H1
file /export/home/~ dpwr             40
ds2/vnmrsys/data/m~ dof               0
400b/saghatelian/v~ dm              nnn
inayavekhi/111111_~ dmm               c
   DOPEPEG4GSH.fid  dmf            7500
   ACQUISITION      temp           22.0
sfrq       399.413      PROCESSING
tn              H1  lb             0.10
at           2.733  wtfile
np           32768  proc             ft
sw          5995.2  fn         not used
fb            3400
bs               8  werr
ss               2  wexp        svf(n1)
tpwr            59  wbs
pw            10.0  wnt
d1           0.120
tof              0
nt            1024
ct             130
alock            n
gain      not used
      FLAGS
il               n
in               n
dp               y
     DISPLAY
sp            -0.3
wp          1198.0
vs             151
sc               0
wc             250
hzmm          4.79
is          289.57
rfl         2323.9
rfp         1318.1
th              10
ins          6.000
nm  cdc  ph
```

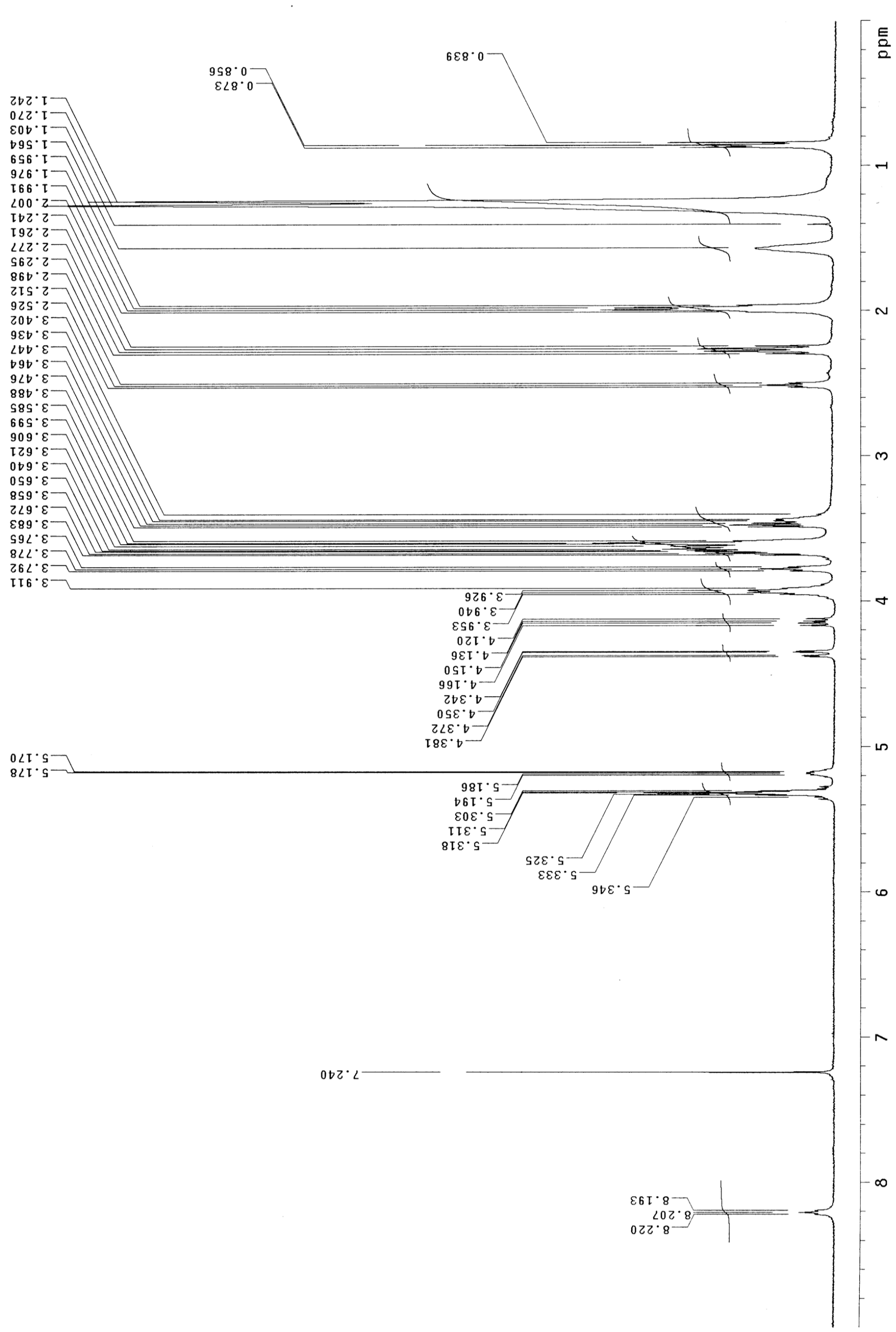




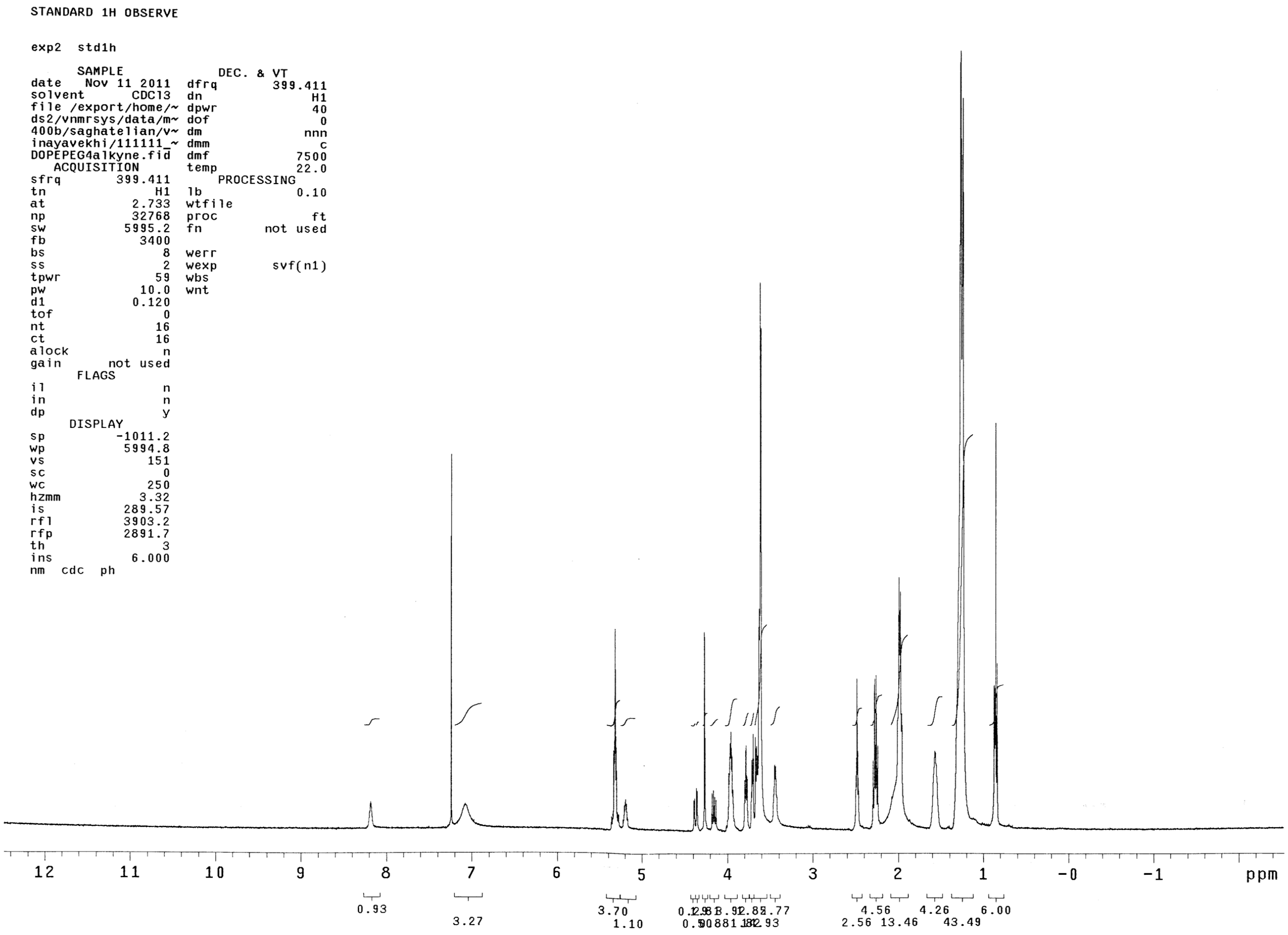
```
STANDARD 1H OBSERVE

exp2  std1h

      SAMPLE              DEC. & VT
date   Nov 11 2011  dfrq       399.413
solvent      CD3OD  dn              H1
file /export/home/~ dpwr            40
ds2/vnmrsys/data/m~ dof              0
400b/saghatelian/v~ dm             nnn
inayavekhi/111111_~ dmm              c
   DOPEPEG4GSH.fid  dmf           7500
   ACQUISITION      temp          22.0
sfrq       399.413       PROCESSING
tn              H1  lb            0.10
at           2.733  wtfile
np           32768  proc            ft
sw          5995.2  fn        not used
fb            3400
bs               8  werr
ss               2  wexp       svf(n1)
tpwr            59  wbs
pw            10.0  wnt
d1           0.120
tof              0
nt            1024
ct             130
alock            n
gain      not used
      FLAGS
il               n
in               n
dp               y
     DISPLAY
sp          1198.0
wp          1198.0
vs             151
sc               0
wc             250
hzmm          4.79
is          289.57
rfl         2323.9
rfp         1318.1
th              10
ins          6.000
nm  cdc  ph
```

NMR spectrum of Azide-PEG$_4$-DOPE

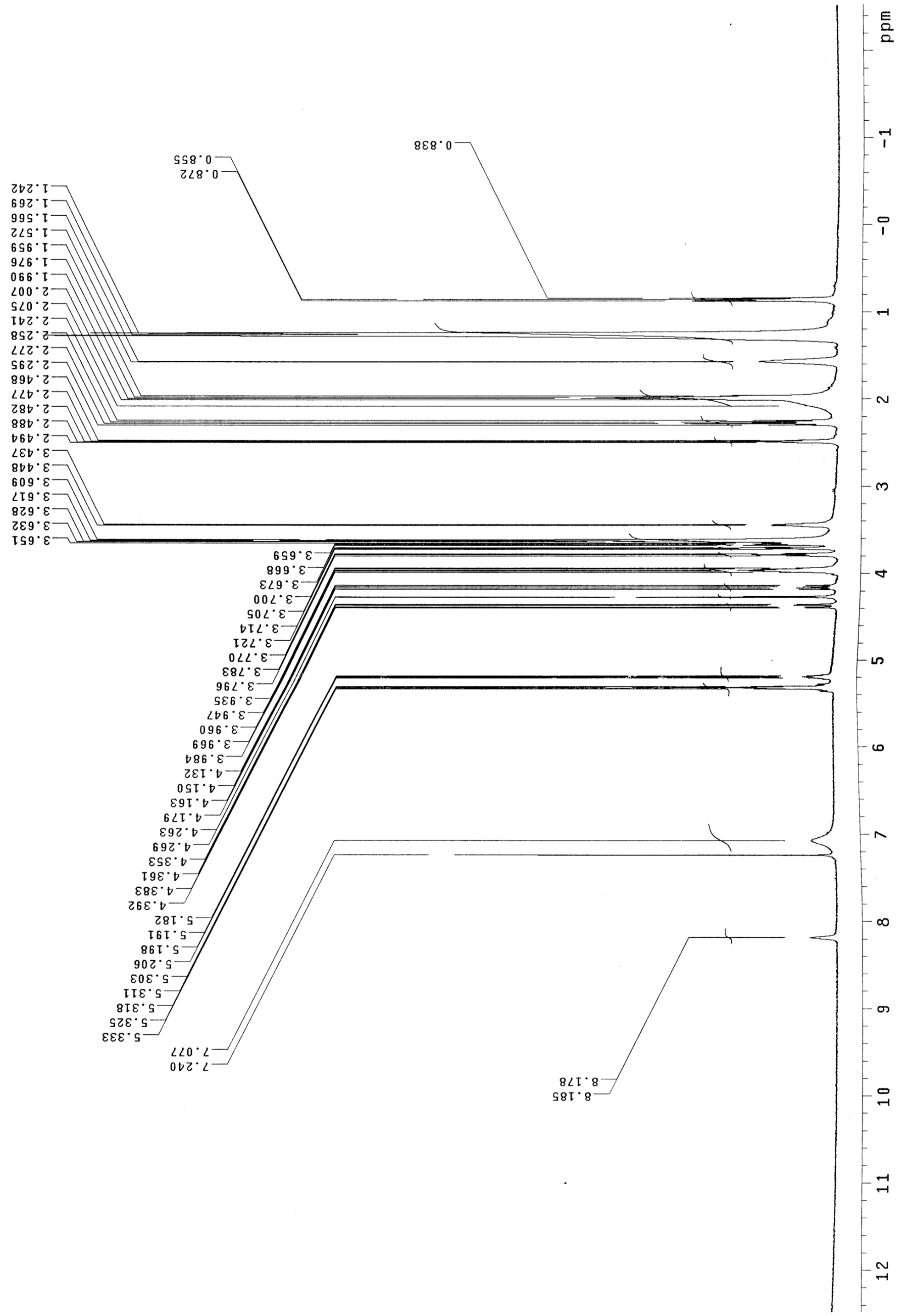



0.839
0.856
0.873
1.242
1.270
1.403
1.564
1.959
1.976
1.991
2.007
2.241
2.261
2.277
2.295
2.498
2.512
2.526
3.402
3.436
3.447
3.464
3.476
3.488
3.585
3.599
3.606
3.621
3.640
3.650
3.658
3.672
3.683
3.765
3.778
3.792
3.911
3.926
3.940
3.953
4.120
4.136
4.150
4.166
4.342
4.350
4.372
4.381
5.170
5.178
5.186
5.194
5.303
5.311
5.318
5.325
5.333
5.346
7.240
8.193
8.207
8.220
ppm
1
2
3
4
5
6
7
8

NMR spectrum of Alkyne-PEG$_4$-DOPE



```
STANDARD 1H OBSERVE

exp2  std1h

      SAMPLE              DEC. & VT
date   Nov 11 2011  dfrq       399.411
solvent      CDCl3  dn              H1
file /export/home/~ dpwr            40
ds2/vnmrsys/data/m~ dof              0
400b/saghatelian/v~ dm             nnn
inayavekhi/111111_~ dmm              c
DOPEPEG4alkyne.fid  dmf           7500
   ACQUISITION      temp          22.0
sfrq        399.411      PROCESSING
tn               H1 lb            0.10
at            2.733 wtfile
np            32768 proc            ft
sw           5995.2 fn        not used
fb             3400
bs                8 werr
ss                2 wexp       svf(n1)
tpwr             59 wbs
pw             10.0 wnt
d1            0.120
tof               0
nt               16
ct               16
alock             n
gain       not used
      FLAGS
il                n
in                n
dp                y
     DISPLAY
sp          -1011.2
wp           5994.8
vs              151
sc                0
wc              250
hzmm           3.32
is           289.57
rfl          3903.2
rfp          2891.7
th                3
ins           6.000
nm  cdc  ph
```


0.838
0.855
0.872
1.242
1.269
1.566
1.572
1.959
1.976
1.990
2.007
2.075
2.241
2.258
2.277
2.295
2.468
2.477
2.482
2.488
2.494
3.437
3.448
3.609
3.617
3.628
3.632
3.651
3.659
3.668
3.673
3.700
3.705
3.714
3.721
3.770
3.783
3.796
3.935
3.947
3.960
3.969
3.984
4.132
4.150
4.163
4.179
4.263
4.269
4.353
4.361
4.383
4.392
5.182
5.191
5.198
5.206
5.303
5.311
5.318
5.325
5.333
7.077
7.240
8.178
8.185
12
11
10
9
8
7
6
5
4
3
2
1
−0
−1
ppm