Structure/Function Studies on Vps9 domain-containing Ankyrin Repeat Protein (VARP)

This dissertation is submitted for the degree of Doctor of Philosophy

Harriet Elizabeth Mary Crawley-Snowdon

MRC Laboratory of Molecular Biology and Darwin College

University of Cambridge

June 2019

Declaration and Statement of Length

This dissertation is the result of my own work and includes nothing that is the outcome of work done in collaboration except as declared in the Acknowledgements and specified in the text.

This dissertation is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Acknowledgements and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Acknowledgements and specified in the text.

This thesis does not exceed the word limit as prescribed by the Faculty of Biology.

Abstract

Structure/Function Studies on Vps9 domain-containing Ankyrin Repeat Protein (VARP)

Harriet Crawley-Snowdon

Vps-9 domain ankyrin repeat protein (Varp) is a widely expressed multi-domain protein of 1050 amino acids that resides in the cytosol. Varp has been shown to be a Rab21 GEF, a Rab32/38 effector and also helps to regulate VAMP7 fusion activity. Varp has more recently been proposed to bind to the Golgi tether golginA4, the kinesin motor Kif5a and the retromer complex subunit Vps29. The retromer core complex, comprising Vps35, Vps26, and Vps29, utilises different SNX proteins to mediate both Endosome to Golgi traffic, and Endosome to Plasma membrane traffic. The interaction of Varp with Vps29 localises Varp to endosomal membranes. Varp has two ‘zinc knuckle’ sequences (residues 369-460 and 692-746) that have been shown to bind zinc and it is these small sequences that mediate the interaction with Vps29.

The Varp:Vps29 complex structure was determined by a combined X-Ray Crystallographic/NMR method. This thesis demonstrates that one zinc knuckle sequence of Varp is sufficient to form an interaction with Vps29, with an affinity in the micromolar range. Single point mutations in Vps29 are sufficient to affect the binding affinity of Varp. Additionally, as some reported functional interactions have not yet been localised to any particular region of Varp, the as-yet-unstudied N-terminal domain of Varp (residues 1-136) was investigated using a Yeast-2-Hybrid screening approach, to try to identify any interaction partners that bind in this region. Many potential interaction partners were discovered, however neither Kif5a nor golginA4 was detected binding in this region.

Acknowledgements

I would firstly like to extend most grateful thanks to my supervisor David Neuhaus, who has been extremely supportive, helpful and unendingly patient. David Neuhaus also assisted the work in this thesis considerably by devising and implementing within the program XPLOR-NIH the approach used for generating the structures, and, together with Ji-Chun Yang, executing most of the calculations. I would also like to thank Ji-Chun Yang for his help with NMR queries, assignment queries and supportive conversations. Special thanks go to David Owen who provided valuable and ongoing support, informative discussions, and constructive feedback. My University supervisor Paul Luzio was also a great help and provided much valuable advice, as well as expressing great interest in my work and allowing me to include unpublished Yeast-2-Hybrid screening data generated in his group.

I wish to thank various people for their contributions to this project. Information and practical guidance provided by Sally Gray regarding Yeast-2-Hybrid methods was invaluable for this research and was gratefully received, and in addition, Sally performed the directed Y2H screening experiments discussed in Section 6.3 to confirm the biological implications of the structural data that I generated. Nathan Zaccai was very helpful in aiding my understanding of both the theory and practicalities of SPR experiments; Nathan also allowed me to present his SPR findings, based on my own earlier experiments, in this thesis as discussed in part of Section 5.6. Lena Wartosch aided my research with important discussions and practical advice and her friendship and great company were much appreciated. Lena Wartosch also provided her *in vivo* data relating to structure-directed mutants of Vps29, discussed in part of Section 5.7

I would also like to thank lab members past and present for their company, enjoyable tea breaks and supportive conversations, including Laura Easton, Katy Hedgethorne, Tom Ogden, Wilfred Wu, Antoni Wrobel and Lauren Jackson. The staff at the LMB NMR facility were also always a great help, so my thanks go out to you all. My thanks also go to the administrative and support staff at the LMB for all their work behind the scenes. I also want to thank the

Laboratory of Molecular Biology for funding this project and allowing me to undertake my further studies.

Lastly, I want to extend thanks to both my family and friends. Thank you to my father, Leon Snowdon, and my siblings for their support and encouragement, homemade dinners, and occasional lodging space. To my late mother, Karen Crawley, thank you for pushing me to be my best. Thank you to the entire Godden family, who have treated me as one of their own and special thanks go to Liam Godden, who has always been by my side. Thanks to the Marlions for the completely ridiculous conversations and the fun times. Finally, for the unending laughs, encouragement, chocolate, phone calls, sanity checks, and hugs I cannot thank Gemma Raishbrook, Ferrelyn Tugwell, Tabitha Mardle and Skye Maddox enough.

Table of Contents

List of Figures

List of Tables

Chapter One

1 Introduction

The trafficking of transmembrane proteins through the endocytic system is a subject of much study due to the complexity of the systems that cells employ to engulf, sort, transport, export, store, and degrade a variety of different molecules. It is therefore necessary to delve into the molecular mechanisms of these trafficking pathways to deepen our understanding of this fundamental cellular process. Intracellular trafficking is undertaken and regulated by a vast network of interacting proteins.

In 2006, Vps9 domain-containing Ankyrin Repeat Protein (**Varp** or ANKRD27) was identified in a Yeast-2-Hybrid screen as being a guanine nucleotide exchange factor for Rab21 [1]. Rab21 is one of a family of Rab GTPases that are organelle specific 'switches' that underpin the process of intracellular trafficking by recruiting effector proteins with a variety of roles. From further investigation into the function of Varp, it became evident that Varp interacts with many classes of proteins that have roles throughout the trafficking process [2-10], especially the evolutionarily conserved retromer complex which mediates trafficking within the endocytic system.

In order to understand the interaction between Varp and the retromer sorting complex component Vps29 [7], the work presented in this thesis employs NMR studies on both Varp and Vps29 as well as various biochemical and biophysical techniques on the complex to increase our understanding of this interaction. The information gained is used to direct *in vivo* experiments carried out by collaborators. In addition, further studies were conducted to attempt to elucidate the function (if any) of the N-terminal region of Varp. This chapter introduces the background to this project, including a brief general introduction to intracellular trafficking, followed by a discussion of the role of both the retromer sorting complex and Varp, and then a brief overview of Nuclear Magnetic Resonance (NMR) Spectroscopy as it has been used in this project.

1.1 Intracellullar trafficking

The compartmentalization and segregation of many processes in a eukaryotic cell is essential to proper functioning and overall survival. Essential processes such as protein glycosylation, proteolytic processing and waste digestion all require specific environments to function most efficiently. To achieve this, membrane-enclosed compartments (referred to collectively as organelles) create a lumen that can be very different from the cytoplasm in composition and properties. Because the presence of specific transmembrane and peripheral proteins and lipids define the function of individual organelles, these macromolecules can be employed as 'markers' to identify membrane-enclosed compartments. For example, it is accepted that the conversion of Rab5 GTPase to Rab7 GTPase on endosomal membranes marks the maturation of early endosomes to late endosomes [11].

In most instances, however, material must move continuously through various compartments within the cell. For example, newly made proteins must be modified and transported to their cellular site of activity (Biosynthetic pathway), and internalized extracellular molecules may be either recycled or degraded (Endocytic pathway). One of the key challenges a cell must overcome is to regulate the flow of material through different organelles, whilst maintaining the lipid composition of the membrane and the assortment of functional proteins required within the organelle to enable the specific process it exists to perform. These functional macromolecules must be spatially and temporally regulated to ensure the fidelity of the system. The strategies employed by the cell to maintain organelle diversity and ensure timely cargo delivery to the correct organelle are a subject of great interest particularly as it becomes apparent that defects in endosomal systems can result in a variety of disorders. This includes both Alzheimer's and Parkinson's diseases [12-14], additionally, defects in a variety of trafficking related proteins cause Charcot-Marie-Tooth type disorders [15, 16], whilst errors in DNA encoding the clathrin adaptor AP-3 can result in Hermansky-Pudlak Syndrome [17].

Proteins in the cell are subject to intracellular trafficking throughout their life, from synthesis to degradation, and although the paths they take at different stages can be distinct depending on the fate of the protein, there may be overlaps in destinations along the way. The biosynthetic pathway is the route by which intracellular proteins are post-translationally modified and transported to their site of function, which can either be intracellular or extracellular, whereas the endocytic pathway exists primarily to internalize molecules from the extracellular space and plasma membrane from whence they can be recycled or degraded, however, there is cross talk between the various routes. Key studies to identify details of the endocytic pathway have been undertaken on the internalization and movement of receptors and their ligands from the cell-surface.

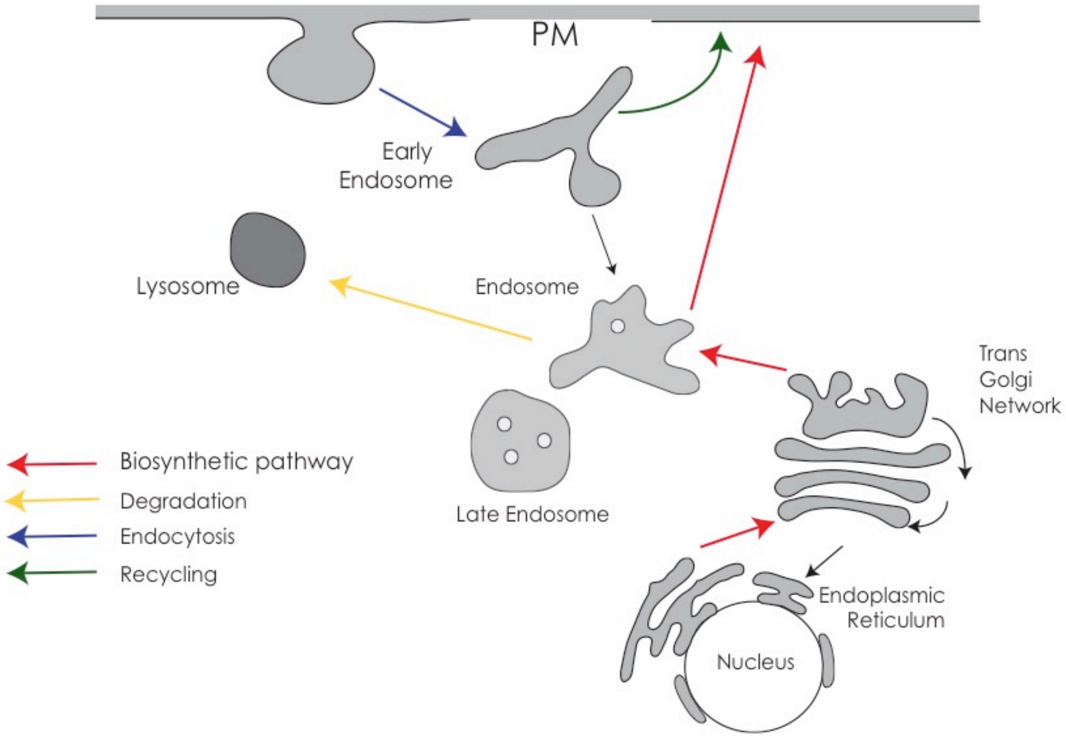


Figure 1: Diagram depicting the various routes included in the biosynthetic, recycling, degradative and endocytic pathways within the cell.

1.2 Vesicular-tubular Trafficking

The endocytic trafficking system relies on the formation and transport of membrane-bounded vesicular or tubular structures that act as carriers. In general carrier transport can be described by four sequential steps, cargo selection and carrier formation, budding and transport, tethering and docking, and finally membrane fusion, as depicted in Figure 2. A diverse network of proteins is employed throughout the trafficking pathway to ensure fidelity of the system at every stage. This includes cargo sorting machinery, coat proteins, tethering proteins, Rab GTPases and membrane fusion-mediating SNARE (SNAP (Soluble NSF(N-ethylmaleimide-sensitive factor) Attachment Protein) Receptor) proteins which all work together to mediate the intracellular destination of a transport vesicle [18]. The model of vesicle trafficking shown in Figure 2 very briefly demonstrates some of the roles of these proteins. Due to the complexity of the system the various connections between the different types of protein involved have yet to be mapped out entirely.

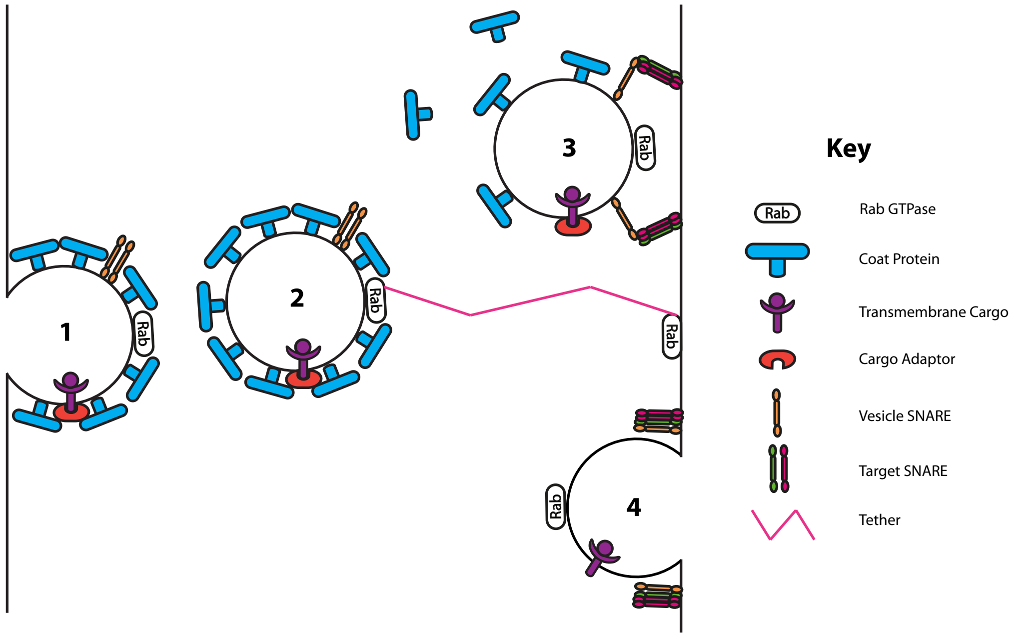


Figure 2: A schematic and simplified diagram of vesicle trafficking. Vesicle trafficking within the cell relies on a variety of different proteins working in unison. 1. Vesicles invaginate to sequester transmembrane cargo bound by a cargo adaptor. These cargo adaptors can bind to coat proteins, which then deform the membrane into vesicles. Vesicles that have budded then travel to the appropriate destination using the cytoskeleton. 2. Rab GTPase proteins and tethering complexes seem to mediate the interaction between vesicles and their correct destination membrane. Incoming vesicles are recognized by long-range tethering proteins. 3. Once captured, vesicle SNARE proteins interact with target SNARE proteins on the target membrane to form a trans-SNARE complex. This is thought to overcome the repulsive energy barrier of the two membranes and allow them to fuse. 4. The trans-SNARE complex is ‘zippered’ shut causing membrane fusion and delivery of the transmembrane protein cargo contained within the vesicle.

1.2.1 Cargo selection and carrier formation

Cargo selection and carrier formation are intrinsically linked, as both are contingent upon one another. Coat proteins have been described for most of the intracellular trafficking routes currently known and are responsible for the membrane deformation needed to generate membrane-enclosed carriers. The coat specific adaptor or cargo recognition proteins act to identify and bind cargo, and as a result of their coat-specific interactions, sequester cargo into transport carriers. Different combinations of coat proteins and cargo adaptor proteins perform each of the possible trafficking routes in the cell.

Clathrin was the first coat protein complex to be discovered, and the role of clathrin in both endocytosis and intracellular trafficking has subsequently been heavily studied and comprehensively reviewed [19, 20]. Since then, other coat proteins mediating intracellular trafficking have been described, including COPI, COPII and retromer [21-23]. The diagram in Figure 3 depicts a simple outline of the intracellular trafficking pathways mediated by some specific coat proteins. Clathrin has been shown to interact with three cargo-binding ‘adaptor protein complexes’ termed AP-1, AP-2 and AP-3. However, AP-3 does not absolutely require clathrin binding for function, as demonstrated in yeast where AP-3 lacks the clathrin interaction region, called the ‘clathrin box’ [20] . There are additional adaptor protein complexes, AP-4 and AP-5 and the more distantly related TSET, but their contribution to clathrin trafficking is open to question; AP-4 does not bind to clathrin and AP-5 is not universally conserved and neither are required for viability in animals [20, 24] . Clathrin mediated endocytosis (CME) from the plasma membrane is undertaken by the well-studied coat protein/adaptor complex pair, clathrin and adaptor complex AP-2, whilst clathrin coated vesicles (CCVs) arising from specific intracellular organelles are formed with AP-1 (TGN) and in some instances AP-3 (Endosome) [25]. It should be noted however, that clathrin and AP-2 are not alone responsible for CME, there is a large number of accessory proteins that are also needed [26]. Many different types of proteins have been identified as being part of the endocytic process, with diverse roles including expanding the cargo binding repertoire, carrier formation, membrane fission and fusion. For example, ARH and arrestins bind transmembrane receptors LDLR and GPCRs respectively, as cargo for endocytosis [27, 28] whilst specific proteins are required to select SNARE protein cargo for trafficking: CALM traffics VAMP2, VAMP3 and VAMP8, and Hrb traffics VAMP7 [29, 30]. Proteins such as epsin, SNX9, and Amphiphysin act to create the necessary membrane curvature, by helical insertion and/or molecular crowding [31, 32]. To promote CCV formation, the regulatory protein FCHO recruits some of the machinery for CCV formation, including eps15 which can act as a scaffolding protein that stabilises the forming clathrin coated pit [31, 33]. Subsequently, there are a variety of proteins that co-ordinate to finalise the process, including dynamin which controls fission of the CCV [34] and both HSC70 and auxilin which act to dissemble the clathrin coat [35]. There are also proteins that regulate the process, such as the kinase AAK1, and by homology the closely related kinase BMP2K which has been demonstrated to colocalise with the AP-2 α subunit [36, 37]. There is some functional redundancy in this process, for example, SNX9 not only has a role in membrane curvature, but has been reported to select cargo and link CME to the actin cytoskeleton [31, 38].

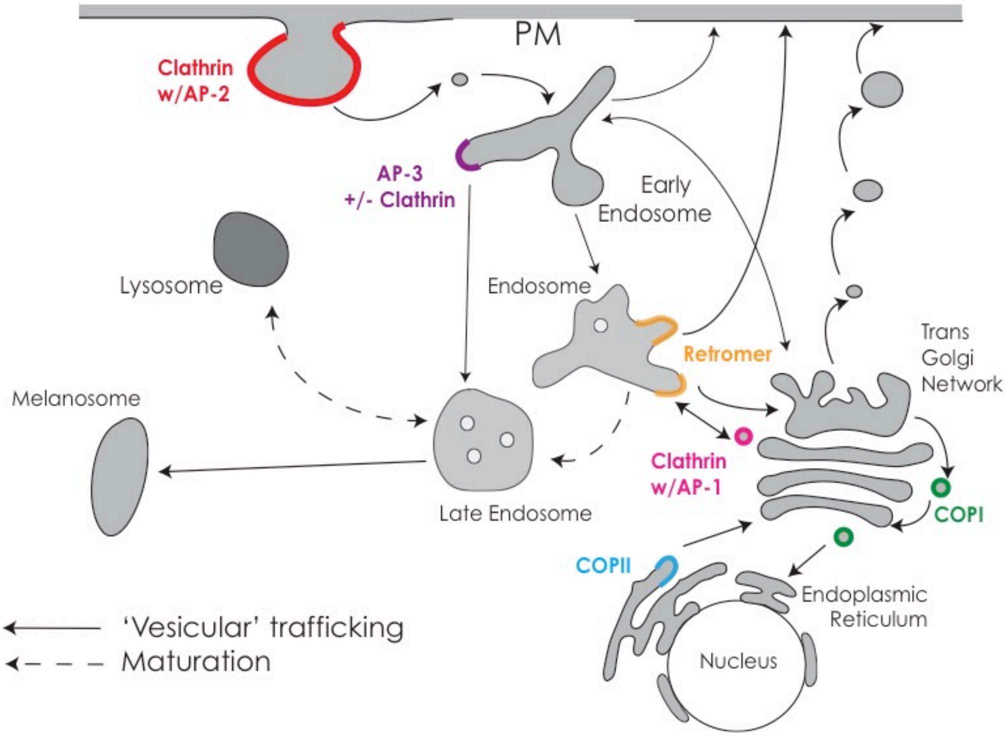


Figure 3: A diagram showing the role of some different coat and adaptor proteins within the cell.

In the secretory pathway, the clathrin related proteins COPI and COPII are critical as they are the coat complexes that mediate exchange between the Endoplasmic Reticulum (ER) and the Golgi apparatus; the COPI coat is also required for trafficking between the Golgi stacks [39]. It has also been reported that COPI may have a role in endosome maturation and the formation of multivesicular bodies (MVBs), although this is somewhat controversial and the loss of MVBs in Arf1GTPase deficient cells is possibly due to other activities of Arf1 besides the recruitment of COPI [40].

There is some question over whether retromer is a true 'coat' as it may generate tubules rather than vesicles [41]. However, for the purpose of this discussion, it will be considered as a coat. The retromer complex functions to return proteins from the endosome to the TGN, which can be considered as a 'retrograde' trafficking route in the context of the biosynthetic pathway. More recently, retromer has been implicated in the recycling of internalised receptors to the Plasma Membrane (PM). Retromer-mediated recycling could thus prevent lysosomal degradation of critical receptors. Indeed, Arighi et al demonstrated that depletion of retromer results in lysosomal swelling and increased lysosomal degradation of CIMPR [42]. The function and architecture of retromer will be discussed in more detail in Section 1.3.

1.2.2 Membrane Identity and Tethering

The identity of both the desired donor and acceptor membranes must be accurately recognised by the proteins involved in cargo transport to ensure the accurate delivery of the correct cargo to the correct target. This recognition must occur in the face of the membrane needing to change identity between leaving the parent and becoming incorporated into its target, hence there is also a temporal element involved. Generally, there are molecular markers that one can use to describe the identity of intracellular organelles, tubules and vesicles. Key determinants of membrane identity seem to be the phospholipids of which the membrane itself is comprised and the proteins that are either integral or associated with the membrane [43-45]. Small Rab GTPase proteins are localised to specific membranes throughout the cell, and can therefore act as markers for organelles within the cell.

1.2.2.1 Rab GTPases

Since Zerial and colleagues initially described the distinct localisation of Rab GTPases to specific membrane-bound organelles within the cell [43], Rab GTPases have been widely recognised as identifiers of various intracellular compartments. Currently, 66 human Rabs have been identified (11 Rab-analogue Ypt proteins in *S.cerevisae*). The localisation, if not the function, of many of these Rab proteins is well documented.

Commonly thought of as 'molecular switches', Rab GTPases cycle between an 'active', GTP-bound, form and an 'inactive', GDP-bound form. Takai et al demonstrated that Rab3A-GDP (previously known as smg p25) was chaperoned in the cytosol by a novel protein, which they named GDP dissociation inhibitor (GDI), in reference to the fact that by binding Rab3A-GDP, GDI can prevent the release of GDP by the Rab [46]. Whilst Rab3A could bind to membranes in both GTP- and GDP- bound forms, it was demonstrated that GDI extracts only Rab3A-GDP from the membrane [47].

Subsequently, it has been discovered that all Rab GTPases are chaperoned in the cytosol, by one of two GDI proteins in humans, or just one GDI protein in yeast. In 2006, Pylypenko et al solved the structure of Ypt1 (the yeast homolog of Rab1) in complex with GDI, and thus established that GDI binds not only to the active 'switch' region of Ypt1, but also the membrane-binding prenyl group attached to Rab GTPases [48].

The binding of GDI to Rab GTPase geranylgeranyl moieties prevents Rab-GDP from using these lipid moieties to insert into the membrane. Once the GDI is removed, the inactive Rab-GDP can bind to the appropriate membrane. It has been shown that both Rab5-GDP and Rab9-GDP bind to membranes and are subsequently converted *in situ* to the active, GTP bound form, and it is now generally thought that this mechanism holds true for all Rab GTPases [49, 50]. How then do Rab GTPases in their GDP-bound form become released from GDI and become capable of binding the membrane? Based on the discovery of a possible GDI – displacement factor (GDF), Yip3 (PRA1 in humans), it was speculated that GDFs may be a universal requirement to remove GDI and thus allow Rab-GDP to bind to membranes [51, 52]. This has not, however, yet been demonstrated and whilst the specific location of Rab GTPases on intracellular membranes is well known, the process regulating this localisation is still somewhat unclear.

Studies have shown that Rab GTPases can be mislocalised by anchoring their cognate guanine nucleotide exchange factors (GEFs) to inappropriate intracellular membranes [53-55]. GEFs work to 'activate' Rab GTPases by exchanging GDP for GTP, thereby preventing GDI proteins extracting them from the membrane. A host of different proteins have GEF activity, some of which have common domains - including Vps9 domains that interact specifically with Rab5 family members. Therefore the current general theory is that Rab GTPases have the ability to anchor in many membranes, but it is the conversion of membrane-inserted Rab-GDP to Rab-GTP by GEFs that keep Rab GTPases anchored in the appropriate membrane [56].

Given their ability to control Rab GTPase membrane localisation, GEFs must also be subject to regulation. The current model for this suggests that GEFs exist in an autoinhibited state that is liberated upon interaction with regulatory proteins. For example, whilst the Rab5-binding Vps9 domain of Rabex5 has some innate GEF activity, it is much more strongly activated in the presence of the Rab5 effector protein Rabaptin-5 [57]. Active GTP-bound Rab GTPases bind a variety of functional proteins, termed ‘effector’ proteins. In some instances, an effector for one active GTP-bound Rab GTPase will be a GEF for another subsequent Rab GTPase. This is the basis for what is known as a ‘Rab cascade’ whereby one Rab protein establishes the localisation for subsequent acting Rabs. A well-known illustration of such a cascade mechanism is the conversion of Rab5 (Vps21 in yeast) to Rab7 (Ypt7 in yeast) via the action of GEFs and effectors during early to late endosome maturation [11, 58].

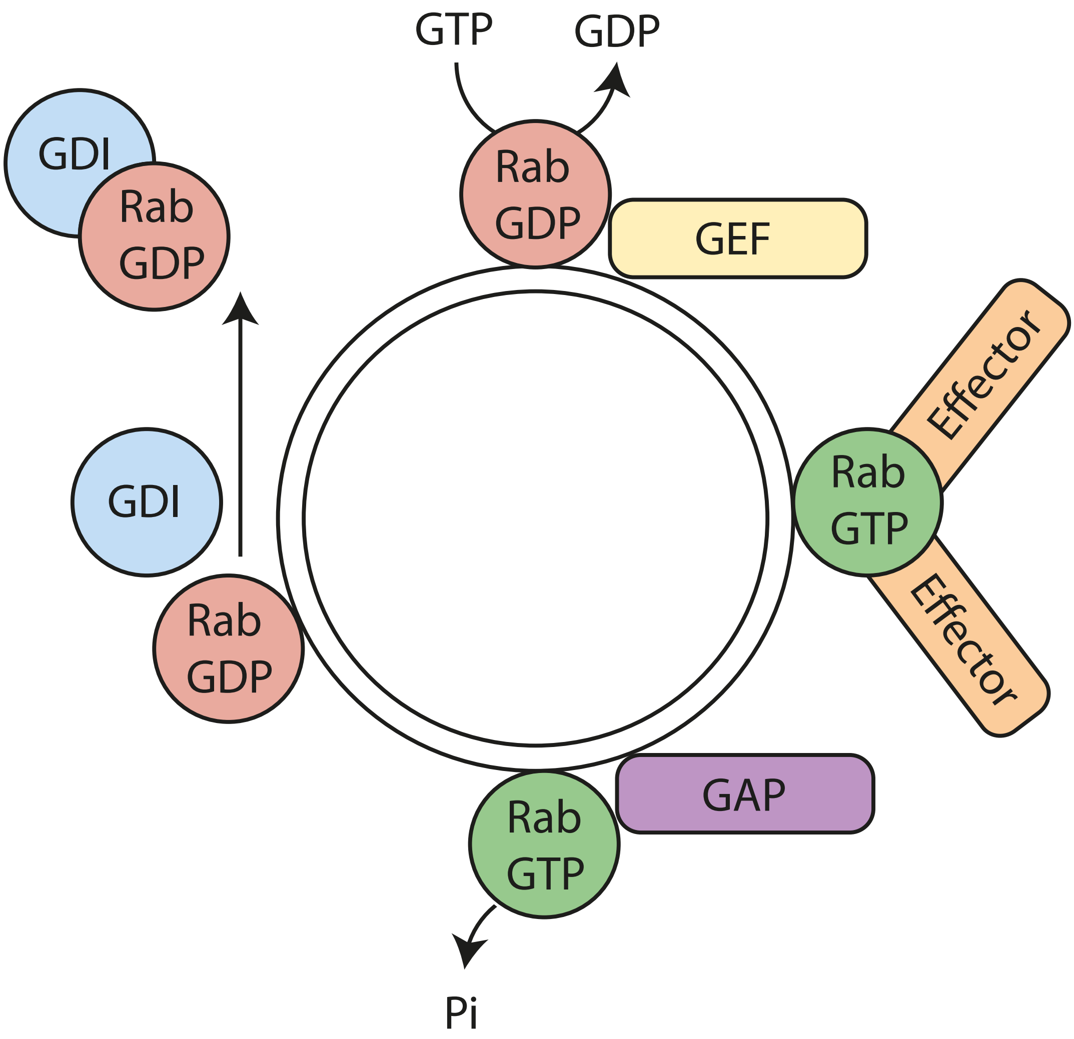


Figure 4: A diagram illustrating the Rab GTPase activation/inactivation cycle. GDP Dissociation Inhibitor (GDI) chaperones Rab-GDP in the cytosol. Once bound to the membrane, the action of Rab Guanine nucleotide Exchange Factors (GEFs) converts Rab-GDP to Rab-GTP. Effector proteins bind to the active Rab-GTP to carry out a number of roles within the cell. To terminate activity, Rab-GTP can be inactivated by GTPase Activating Proteins (GAPs). GDI can then remove inactive Rab-GDP from the membrane.

In terms of intracellular trafficking, there must also be a mechanism for active Rab-GTP to be inactivated to Rab-GDP e.g. once the Rab has reached the destination membrane and is no longer required. The intrinsic GTPase activity of Rab GTPases is rather slow, however interaction with GTPase activating proteins (GAPs) greatly increases the rate of GTP hydrolysis [59-61]. The Rab GAPs consist mainly of TBC-family (Tre-2/Bub2/Cdc16) domain-containing proteins (or GTPase activating protein for Ypt proteins, Gyps, in yeast); a comprehensive review of TBC proteins can be found in [62] . Along with recruiting the Rab Sec4-GEF Sec2, Ypt32 in yeast is known to utilise two GAPs as effector proteins, Gyp1 and Gyp6, GAPs for Ypt1 and Ypt6 respectively [63-65]. This demonstrates how one Rab (Ypt32) can activate a subsequent Rab (Sec4) whilst also enabling the deactivation of previous Rabs (Ypt1 and Ypt6) once they are no longer needed. The inactive, GDP bound forms of Rab GTPases are subsequently extracted from the membrane by GDI proteins. The process of Rab activation/inactivation and membrane binding and removal, as orchestrated by GDI, GDF, GEF and GAPs, can therefore be seen as a cycle, as depicted in Figure 4. A detailed review of the control of Rab GTPase activity by GEFs, GAPs and GDI proteins, which is beyond the scope of this thesis, can be found in [56]. Due to the interplay between Rabs and the variety of proteins that interact with them, the specific targeting of Rabs to the correct intracellular membranes is currently thought to be due to a range of proteins, including Guanine nucleotide exchange factors (GEFs), GTPase-activating proteins (GAPs) and Rab effector proteins.

Effector proteins are varied and diverse, with some Rabs having been discovered to bind more than 25 effectors, likely with many more yet to be recognised [66]. These effector proteins have a variety of different functions which help drive the formation of transport vesicles and enable linkage to motor proteins to allow vesicle movement, as well as recognizing docking factors present on acceptor membranes [67]. Effector proteins most commonly bind to Rabs in their GTP-bound (active) state, although in a few cases some effectors do bind the GDP-bound (inactive) state, while some even do not have a preference for the nucleotide bound form [68]. There are also a number of effector proteins that bind two Rab GTPases [68]. For example, the HOPS (homotypic fusion and protein sorting) complex has two Rab7 binding surfaces, on two different subunits at separate ends of the multisubunit complex [69-71]. This means that the HOPS complex can link two separate membranes to assist in membrane tethering, and this may well prove to be a common mode of tethering. Many molecules involved in membrane tethering can be classed as Rab effectors.

1.2.2.2 Membrane Tethers and Tethering Factors

Molecules that act to tether two membranes together are generally categorised into two groups on the basis of their structure, the coiled-coil proteins and the multisubunit tethering complexes (MTCs). The biological impact of a knockout, or loss-of function mutation of some membrane tethers can in some instances be compensated for, however these mutations/knockouts can also be embryonically lethal [72-74].

The long helical coiled-coil group of tethering proteins are generally only found on the cytosolic side of Golgi membranes, however some coiled-coil tethers can be found elsewhere, e.g. the Rab5 effector EEA1 which is present on early endosomes [75, 76]. The coiled-coil tethering proteins residing on the Golgi are referred to as golgins, and have a specific localisation within the Golgi stack [77]. Golgins are known to be able to extend up to 300nm from the membrane surface, thus allowing a large sampling space for interaction with incoming vesicles [77]. Wong and Munro were able to demonstrate the role of golgins in providing specificity in membrane trafficking by repositioning golgins to the mitochondria and subsequently detecting cargo destined for the Golgi rerouted to the mitochondria [78]. More recently the Munro group has shown that a TBC- family member, TBC1D23, utilises a Rab GTPase GAP domain (which is thought in this case to be catalytically inactive) to bind to golgin-97 or golgin- 245 and the WASH complex on endosome-derived vesicles, thus acting as a golgin-vesicle adaptor [79]. Simultaneously, Navarro Negredo et al showed that TBC1D23 interacts with a WDR11/Fam91A1/C17orf75 complex on AP-1 derived vesicles destined for the Golgi via golgin-245 tethering [80].

Many golgins have been demonstrated to bind to small GTPases, including Rab GTPases; these interactions are thought to be important in tether recruitment [81]. However, coiled coil tethering molecules may bind Rabs at many sites throughout their length, not all of which are Golgi resident Rab GTPases, for instance GCC185 can bind to Rab9A that is present on endosome-derived vesicles [82]. Early Endosomal antigen 1 (EEA1) is a well-studied non-Golgi coiled coil tether that acts to tether vesicles from the plasma membrane to the early endosome, where EEA1 is bound to Rab5 and PI-3-P in the membrane [75, 83]. It is thought that EEA1 can act as a scaffold to hold all the factors required for fusion together because EEA1 forms a large complex that includes another Rab5 effector Rabaptin-5 and GEF Rabex-5, as well as syntaxin 13, a Q-SNARE required for membrane fusion [84].

In general, there are two types of Multisubunit Tethering Complexes, or MTCs, those that act in the secretory system (called Complexes Associated with Tethering Containing Helical Rods - CATCHR) and those that act in the endolysosomal system (Non-CATCHR), as shown in Table 1. Although included under non-CATCHR complexes, the putative tethering proteins of the TRAPP (TRAnsport Protein Particle) family have as-yet-unproven direct tethering function, as discussed in detail in [85]. It should be noted however that roles for TRAPP proteins have been have been demonstrated in both secretory and endolysosomal trafficking, as well as autophagy [86-88].

Multisubunit Tethering Complexes (MTCs)		
	Function/s	**Reference**
CATCHR		
COG	Intra-Golgi trafficking, endosome to Golgi	[89]
Dsl1p	Golgi to ER	[89]
GARP	Endosome to TGN	[89]
Exocyst	Recycling endosomes/Golgi to PM	[89]
Non-CATCHR		
HOPS	Endosome to vacuole/lysosome	[90]
CORVET	Endosome/endosome fusion	[90]
TRAPPI	ER to Golgi	[89]
TRAPPII	Intra-Golgi, Endosome to Golgi	[89]
TRAPPIII (yeast)	Autophagosome formation	[88]

Table 1: A table describing multisubunit tethering complexes and their functions.

The non-CATCHR complexes involved in endolysosomal trafficking, HOPS and CORVET, are both heterohexamers that share a common core of four subunits, Vps11, Vps16, Vps18 and Vps33. CORVET contains two Rab5-interacting subunits, Vps3 and Vps8, whilst the HOPS complex contains two Rab7-interacting subunits, Vps39 and Vps41 [69, 91-94]. The conversion of Rab5 to Rab7 on maturing endosomes thus means that CORVET and HOPS act sequentially.

As previously mentioned, HOPS can simultaneously bind two Rab7 molecules at distinct ends of the complex and it has been shown that HOPS binds to Rab GTPases on opposing membranes to promote fusion [95]. Additionally, HOPS has also been shown to interact with fusion-critical SNARE proteins on opposing membranes [96]. Whilst it has long been known that the HOPS complex binds SNARE proteins, the molecular details of these interactions have proven difficult to determine. It has recently become clear that HOPS can bind the N-terminal region of some SNAREs independent of the Vps33 subunit, however the subunit Vps33 (common to both HOPS and CORVET) is critical in binding SNARE motifs specifically [97]. Vps33 is a Sec1/Munc18 (SM) like protein that interacts with Q-SNAREs and also a quaternary SNARE complex on the vacuole [97]. In 2015, Baker et al reported the structure of Vps33 in complex with Qa-SNARE Vam3 and the R-SNARE Nyv1, and demonstrated the simultaneous SNARE binding by size-exclusion chromatography [96]. The non-overlapping sites of R- and Qa-SNARE binding in this study support this [96]. The CORVET complex is very much less well studied than HOPS, therefore how the Vps33 subunit is utilised in the CORVET complex is not currently clear, as presumably different SNARE complexes would be required in CORVET-mediated fusion. A recent review of the regulation of SNARE complex formation by MTCs more generally can be found in [98].

1.2.3 Membrane Fusion by SNARE proteins

SNAREs (Soluble N-ethylmalemide sensitive factor Attachment protein REceptors) were first characterised as interaction partners of SNAP (Soluble N-ethylmalemide sensitive factor Attachment Protein), which itself interacts with NSF (N-ethylmalemide Sensitive Factor) [99-101] . Specific SNARE proteins, present on donor and acceptor membranes, interact to form a heterotetrameric coiled coil complex. It is thought that the formation of *trans*-SNARE complexes, as they are termed, provides enough energy to overcome the repulsion of two membranes, bringing them close enough to fuse and thus cause lumenal mixing and delivery of cargo. Direct studies of the fusion process have proven difficult however. SNARE proteins are, like many of the proteins involved in membrane trafficking, membrane specific and highly regulated so as to prevent promiscuous and untimely membrane fusion. Whilst *in vitro* assays have demonstrated that SNARE protein complex formation alone is sufficient for synthetic liposome membrane fusion, experimentally reproducing the submillisecond vesicle fusion seen in especially in pre-synaptic neurons has proved problematic [102, 103] .

SNAREs are mostly integral membrane proteins that contain a SNARE motif of approximately 60-70aa, a C-terminal transmembrane (TM) helix and in most cases, an N-terminal domain. A subfamily of SNAREs, termed the SNAP-25 family, do not contain a TM domain, and have two SNARE motifs connected by a linker [104]. It seems that SNAP-25 family proteins are targeted to the membrane via palmitoylated cysteines, however, there is evidence that SNAP-25 palmitoylation is not required for maintaining SNAP-25 localisation close the membrane, presumably post SNARE complex formation [105]. The highly conserved SNARE Ykt6 is also post-translationally palmitoylated to ensure proper Golgi apparatus localisation [106].

The *trans*-SNARE complex required for fusion is formed of a complex of four SNARE motifs, requiring either one SNAP-25 family member and two integral membrane SNAREs, or four integral membrane-type SNARE proteins. The SNARE proteins are further categorised by the residue in the centre of the SNARE motif, either Arginine (R) or Glutamine (Q), with the Q-SNAREs being further subdivided on the basis of the sequence identity of the SNARE motif into Qa-, Qb- and Qc- proteins [107]. Every known *trans*-SNARE complex is formed of a heterotetramer of Qa-, Qb-, Qc- and R-SNARE, including those formed with SNAP-25 family members, as SNAP-25 family proteins all contain one Qb- and one Qc-SNARE [104]. The *trans*-SNARE complex of a four-helix bundle is then 'zippered' which is sufficient to cause membrane fusion with subsequent *cis*-SNARE complex formation, when all the SNARE proteins are contained within one membrane [108, 109]. Similar to Rab GTPases, SNAREs can be thought of as acting in a cycle, as depicted in Figure 5.

Most SNAREs contain an extended region N-terminal to the SNARE motif, with a few exceptions [110]. These N-terminal domains are categorised into 5 subtypes, including the three helical Habc domain found in syntaxins, and the profillin-like longin domain. Longin domains are relatively common, with seven superfamilies of proteins containing longin domains, many of which are proteins involved in trafficking, including AP protein complex subunits such as AP-2 μ2 and σ2 [111, 112]. Examples of both Habc domains and longin domains have been demonstrated to bind back onto the concomitant SNARE motif, thus creating an auto-inhibited 'closed-conformation' which prohibits trans-SNARE complex formation. In this manner, SNARE proteins can be trafficked through the secretory and endocytic pathway without any improper fusion events occurring. In some instances, this is further mediated by accessory proteins.

The extreme N-terminus of some SNARE proteins, referred to as the N-peptide, has been shown to bind to a class of SNARE-interacting proteins called the Sec1/Munc18-like (SM) proteins [113, 114]. Indeed, it has been shown that the N-peptide is critical for the correct apical localisation of Syntaxin 3 (Stx3) in polarised epithelial cells, and also to be a determinant for which isoform of Munc18 is bound [115]. There has been some question over whether Sec1/Munc18 proteins act to prevent or promote membrane fusion. There are seven known SM proteins in humans, with four in yeast. The Munc18-1 (aka Munc18a): Syntaxin1a complex structure demonstrates that Munc18-1 acts to hold SNARE Syntaxin1a in a closed, non-fusogenic conformation, suggesting an inhibitory role [114, 116, 117]. However, as previously discussed, the Vps33 subunit of the HOPS complex, which is an SM protein, can bind two SNAREs, Vam3 and Nyv1, which are thought to be on opposing membranes [96]. Critically, the overlaid structures demonstrate that the two SNARE proteins, once bound, are orientated N- to C- terminus, with the critical R- and Q- residues in alignment [96]. It is known that Munc18-1 can bind the SNARE proteins syntaxin1a (Stx1a) and VAMP2 individually, and recent experiments have shown that, using a model informed by the Vps33:Vam3:Nyv1 structure , Munc18-1, an 'open' form of Stx1a and VAMP2 interact to form a template complex for SNARE complex formation [118]. This study also demonstrated the Vps33:Vam3:Nyv1 template complex, and also a Munc18-3:Stx4:VAMP2 template complex formation as precursors to SNARE complex formation [118].Therefore, if this were true for SM proteins generally, they could be considered to have a role in promoting fusion which is consistent with the finding that SM proteins are required in fusion events, as described more fully in a review by Südhof and Rothman [119]. It has been shown that the HOPS complex can catalyse the interdependent assembly of all four SNARE motifs, not by direct binding, but perhaps by presenting two SNAREs in a conformation conducive to four helix bundle formation [120]. Additionally, it has been shown that HOPS is needed for pore opening in a capacity beyond just SNARE binding activity [121].

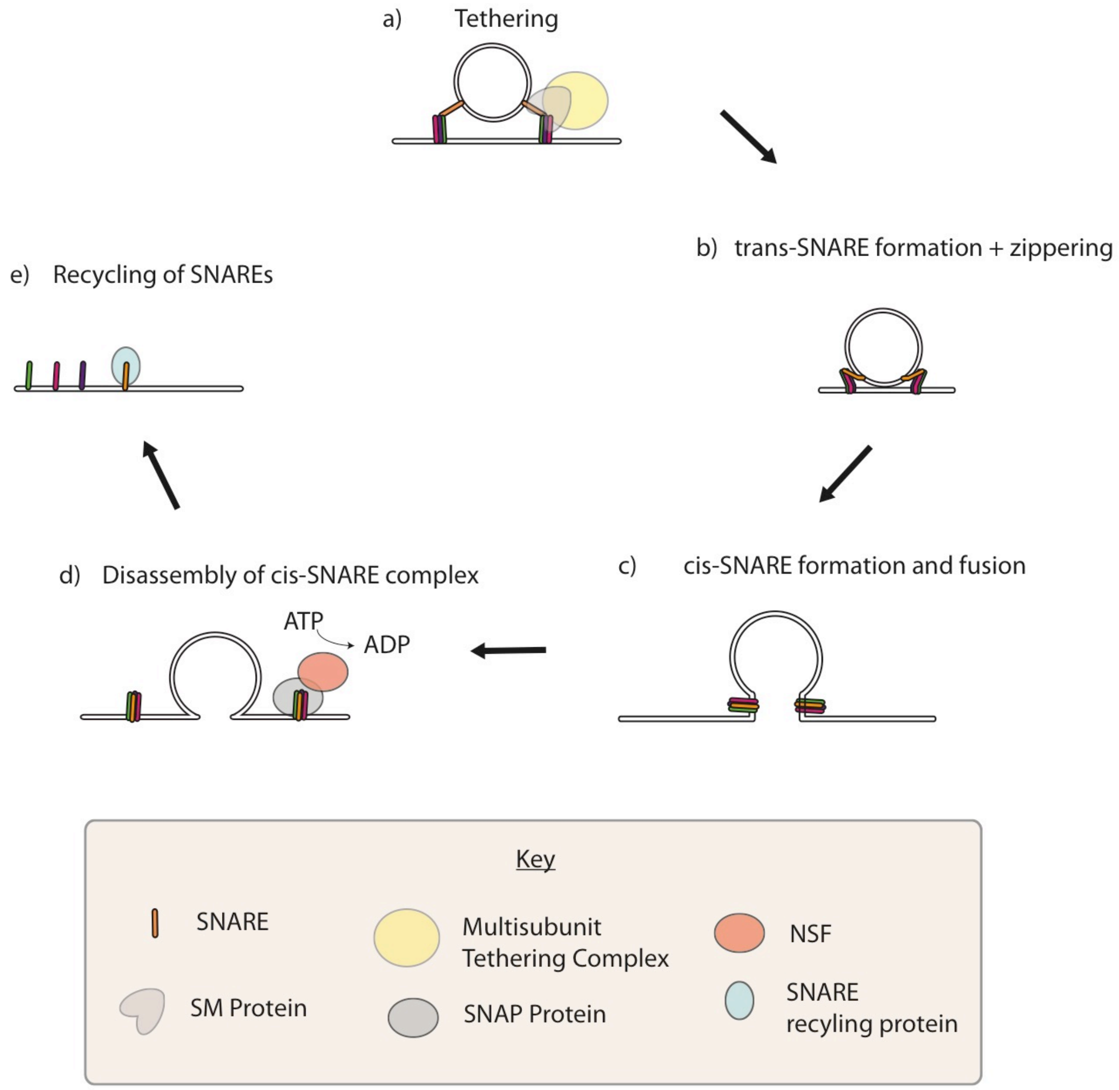


Figure 5: A diagram depicting the SNARE cycle. a) Tethering, as mediated by MTCs or golgins, brings the vesicle close to the acceptor membrane. b) The *trans*-SNARE complex forms, with SNARE proteins interacting via N-terminus of the SNARE motif. This is then 'zipped' to form stage c. c) *Cis*-SNARE complex formation is thought to release enough energy to fuse opposing membranes. d) The *cis*-SNARE complex is dissembled via the action of α/β/γ-SNAP proteins and NSF via the hydrolysis of ATP. e) Individual SNARE proteins are free to be recycled back to their origin membrane.

After membrane fusion and *cis*-SNARE complex formation, NSF and α/β/γ-SNAP proteins (Sec18 and Sec17 in yeast respectively) act to disassemble *cis*-SNARE complexes to release the individual SNARE proteins for reuse, with energy provided from ATP hydrolysis by NSF [100, 122, 123]. However, seemingly in contrast to this well-documented function, it has recently been shown that α-SNAP can also act to stabilise the zippered four helix bundle and thus enhance SNARE assembly and membrane fusion [124]. Additionally, the p97/p47 complex has been shown to act in a similar fashion to the NSF/α-SNAP complex, at the Golgi and the ER. Protein p97 is an ATPase, whilst p47 binds to Golgi-resident SNARE syntaxin 5 (Stx5), however unlike NSF/α-SNAP, this complex requires a cofactor VCIP135 to dissociate SNARE complexes via p97 ATP hydrolysis [125].

SNARE proteins not originating from the fusion membrane must be trafficked back to their origin. This may either be accomplished by recycling single, fusogenically inactive SNARE proteins back to the site of origin for re-inclusion into carriers or as part of an inactive, post-fusion *cis*-SNARE complex [126].

1.2.3.1 VAMP7

VAMP7 is one of three ancestral eukaryotic R-SNARE proteins, and is implicated in many important processes throughout the cell, including membrane remodelling and neurite outgrowth in dendritic cells [127], and it is the most relevant SNARE protein to the work discussed in this thesis. VAMP7 contains a longin domain and mediates fusion of endosomes and endolysosomes with lysosomes and also the fusion of endolysosomes and their derived carriers with the plasma membrane [128]. The longin domain can bind back onto the SNARE domain and in so doing inhibits the SNARE domain from forming *trans*-SNARE complexes [129, 130].

As a result of this, after acting to fuse vesicles with the plasma membrane, VAMP7 must be recycled back to endosomes/endolysosomes. Post-fusion, VAMP7 is returned to endolysosomes by sequential interactions with trafficking proteins, firstly from the plasma membrane to endosomes by AP-2/Clathrin coated vesicles via an accessory clathrin adaptor, Hrb [30], and from early endosomes to later structures via the delta subunit of the AP-3 complex[126]. Both Hrb and the AP-3 delta subunit have been shown to interact with the longin domain of VAMP7 [30, 131]. These trafficking interactions both require VAMP7 to be contained within an inactive *cis*-SNARE complex, thus VAMP7 is prevented from engaging in any fusion reactions until reaching an appropriate compartment [30, 126]. More recently, it was discovered by Schafer at al that VAMP7, no longer in a complex, is bound in an autoinhibited, closed conformation by Varp, until such a time fusion is required [129]. It has been shown that a network of VAMP7, Varp and retromer are all required to traffic the glucose transporter GLUT1 from endosomes to the plasma membrane, which would involve VAMP7 fusogenic capabilities at the plasma membrane [7].

1.3 The Retromer Complex.

As a result of studies in yeast by Seaman et al [23], the retromer complex was discovered as a sorting complex for endosome to Golgi cargo retrieval, and thought to act as a coat complex. Subsequently, studies on the retromer complex have indicated involvement in neurodegenerative diseases such as Alzheimer's disease and Parkinson's disease [12-14].

In yeast, the retromer complex has been said to be a stable heteropentamer comprising the proteins Vps29p, Vps35p, Vps26p, and a sorting nexin with BAR domain (SNX-BAR) dimer of Vps5p, and Vps17p [23]. In mammalian cells the architecture is not so simple. Vps29, Vps35 and Vps26 form a 'core complex' (previously referred to as the cargo selective complex) trimer, whilst the Vps5p homologues SNX1/2 and Vps17p orthologues SNX5/6 [132] seem to be more loosely associated with the core complex, at least in vitro [133, 134]. The composition of the retromer complex in mammals has been extensively studied, with many reviews on the subject [135-137]. Indeed, in mammals, it has become apparent that the core complex and the retromer SNX-BAR proteins can act independently of one another [138-141]. Due to this, the heteropentameric model of a SNX-BAR dimer and a core complex trimer is no longer valid for the vast majority of eukaryotes. The retromer core complex can associate not only with the membrane deforming SNX-BAR proteins SNX1/2 and 5/6, but also a variety of other proteins, including non-BAR SNX proteins like SNX3 (Grd19p/Snx3p homologue) and SNX27 [142, 143]. Additionally, the retromer core complex interacts with the WASH complex and also accessory proteins such as TBC1D5 and Varp (discussed in more detail later). A recent and comprehensive review of retromer sorting complex architecture and function is given in [144].

1.3.1 Retromer Core complex

1.3.1.1 Architecture of the Core complex

The retromer core is a trimer of Vps35, Vps26 and Vps29. In mammalian cells, retromer Vps26 exists in 2 isoforms, Vps26A and Vps26B which share 69% identity and 82% similarity, with marked differences in the C-terminal residues [145]. It has been shown by immunoprecipitation and Yeast 2 Hybrid experiments that Vps35 binds both Vps29 and Vps26, with Vps29 binding at the C-terminus of Vps35 and Vps26 binding at the N-terminal [137, 146, 147]. Structural studies confirm that Vps29 binds to the C-terminal portion of Vps35, which wraps around a large part of the Vps29 protein [137]. More recently, the structure of the Vps35 N-terminal region in complex with Vps26A has been directly determined, albeit with SNX3 to stabilize the crystal [148]. Although structures of both the N-terminal (14-796) and the C-terminal (471-781) portions of Vps35 exist, the structure of full length (1-781) Vps35 was not available until very recently [149]. Figure 6A shows the three retromer core complex subunits from *Chaetomium thermophilum* bound to form the retromer core complex, whilst Figure 6B shows a SNX-BAR dimer of two Vps5 molecules.

Both Vps26A and Vps26B have an arrestin fold, although Vps26 has no sequence homology to previously known arrestin proteins [150, 151]. Arrestin proteins are known to play a part in both regulating GPCR cell signalling and acting as adaptor proteins for clathrin-mediated endocytosis of GPCRs [28, 152]. There are also multiple isoforms of Vps29, the main ones being isoform 1 and 2, different only in the N-terminal tail, with non-canonical isoform 2 having an additional initial MAGHR sequence, which may perhaps prove to be of functional importance. The structure of Vps29 is a phosphoesterase fold and binds divalent metal cations, which led to speculation that it may have some phosphoesterase activity [147, 153]. There has been one report of some phosphoesterase activity of Vps29 against the retromer cargo CI-MPR [154], however no phosphoesterase activity has been detected in any other studies [133, 147].

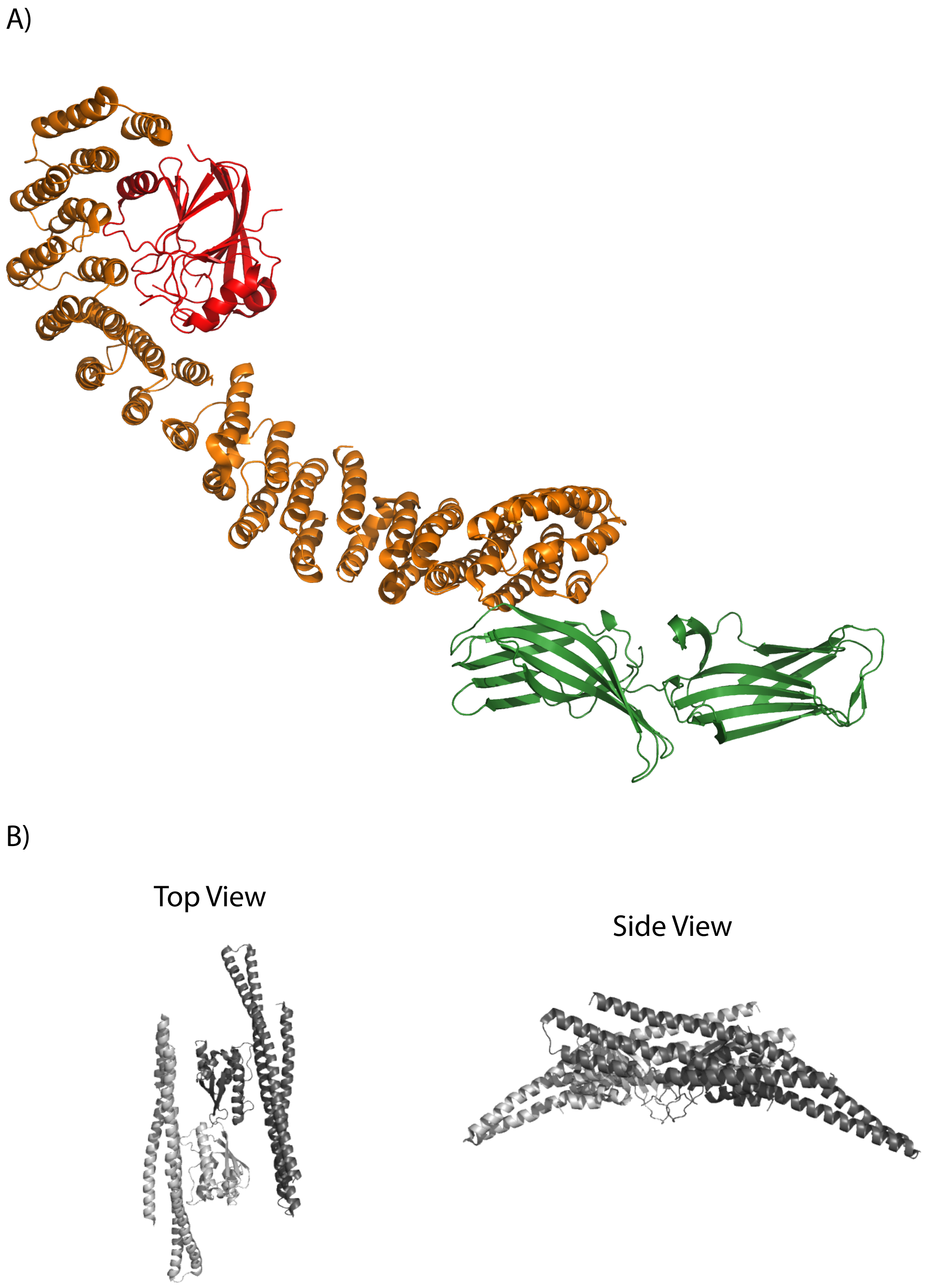


Figure 6: Cartoon depictions of the structure of various retromer proteins from *Chaetomium thermophilum.* A) Retromer core: Vps29 in red, Vps35 in orange, Vps26 in green. B) SNX-BAR dimer of Vps5. Both from pdb 6H7W, as published by Kovtun et al [149].

1.3.1.2 Interactions of the retromer core complex

Heteropentameric yeast retromer has been shown to mediate the retrieval of numerous cargos, including the TGN sorting receptor Vps10p [155], from a pre-vacuole compartment to the TGN. Despite this core being common to retromer of all species, as a result of the different architecture of mammalian retromer, the functions of the core complex (comprising Vps35, Vps26 and Vps29) seem to be somewhat dependent upon which accessory proteins the core interacts with. The core complex has long been speculated to directly bind various cargo proteins, as supported by some experimental studies. Key TGN sorting receptors must be recycled from the endosome back to the TGN and there is evidence that the retromer core complex subunits are directly involved in binding some of these receptors as cargo to be trafficked. For example, CI-MPR has a conserved [FW]L[MV] motif in the cytoplasmic tail to which Vps35 can bind, and it has been shown that Vps26 can bind to receptors SorLA, Sortillin and CI-MPR [156-158]. However, reports on the cargo binding ability of retromer core subunits do vary considerably. For instance, although Arighi et al didn't find an interaction between CI-MPR and Vps26 [42], Seaman et al detected an interaction between the two later [158]. Not only this, but two recent studies have shown that the cation-independent mannose-6-phosphate receptor (CI-MPR), a widely accepted cargo of retromer core specifically, is directly bound by SNX proteins SNX5/6, and it is SNX proteins that determine CI-MPR localisation, with no retromer contribution [159, 160]. Subsequently Cui et al conducted a study that suggests that CI-MPR may have multiple trafficking itineraries, with only part of the trafficking due to retromer core and the non-SNX-BAR SNX3, with the Golgi tether bound by the tubule also having an impact [161]. This raises questions about the true functions of both the retromer core subunits and the SNX-BAR proteins.

A body of evidence is now building to refute the original model of retromer cargo binding, and this has resulted in the current name 'retromer core complex' becoming accepted and replacing the historical, less accurate name of 'cargo binding complex'. As well as the finding of CI-MPR binding SNX5/6 by Simonetti et al [159], a recent publication by Kovtun et al detailing the organisation of the retromer complex on the membrane also raises questions about the method of cargo binding [149]. Figure 7 shows the Electron Tomography determined structure of retromer subunits from the single cell eukaryote *Chaetomium thermophilum* assembled on an artificial membrane. From Figure 7, it is clear that the retromer core complex projects away from the surface of the membrane, with Vps26 closest to the membrane bound to Vps5 [149]. Certainly, the arrangement of molecules in this structure suggest much steric hinderance to any cargo binding by the Vps35 subunits. However, how this arrangement might alter in the presence of cargo has yet to be investigated and some Vps5 subunits may be dispensed with in this case to allow cargo to bind. Conversely, it could be the case that SNX proteins provide more cargo binding capability than previously thought, which is discussed later. In the Vps26:SorLA binding study conducted by Fjorback et al, no SNX proteins were directly detected in a native pulldown study with subsequent mass spectrometry analysis, although it should be noted that only bands of interest were selected for MS analysis [157]. The same study utilised CI-MPR as a positive control for retromer binding and some SNX2 and SNX6 peptides were detected [157], which is consistent with the later finding by Simonetti et al [159]. However, the retromer core component Vps35 was also detected binding to CI-MPR [157], as reported by Seaman [158]. As a result of these seemingly conflicting results, understanding the specific function of the retromer core complex is a matter of ongoing investigation.

So far, the Vps29 subunit of the retromer core complex has not been shown to directly bind any cargo, however it does bind to regulatory/accessory proteins such as the putative Rab7 activating protein TBC1D5 and Varp; it is also the target of the *L.Pnuemophila* infectivity factor RidL [7, 162, 163]. Although Vps29 has a high degree of conservation across all eukaryotes, TBC1D5 is not uniformly conserved. Interestingly, yeast Vps29p appears to contain a loop of amino acids that is not present in mammalian retromer, and there is a good correlation between the absence of this loop and the presence of TBC1D5 in the genome [164]. TBC1D5 is widely thought to promote the GTPase activity of Rab7, with the inactivation of TBC1D5 recently being shown to increase the intracellular levels of GTP-Rab7a [165]. This is of particular interest, as the

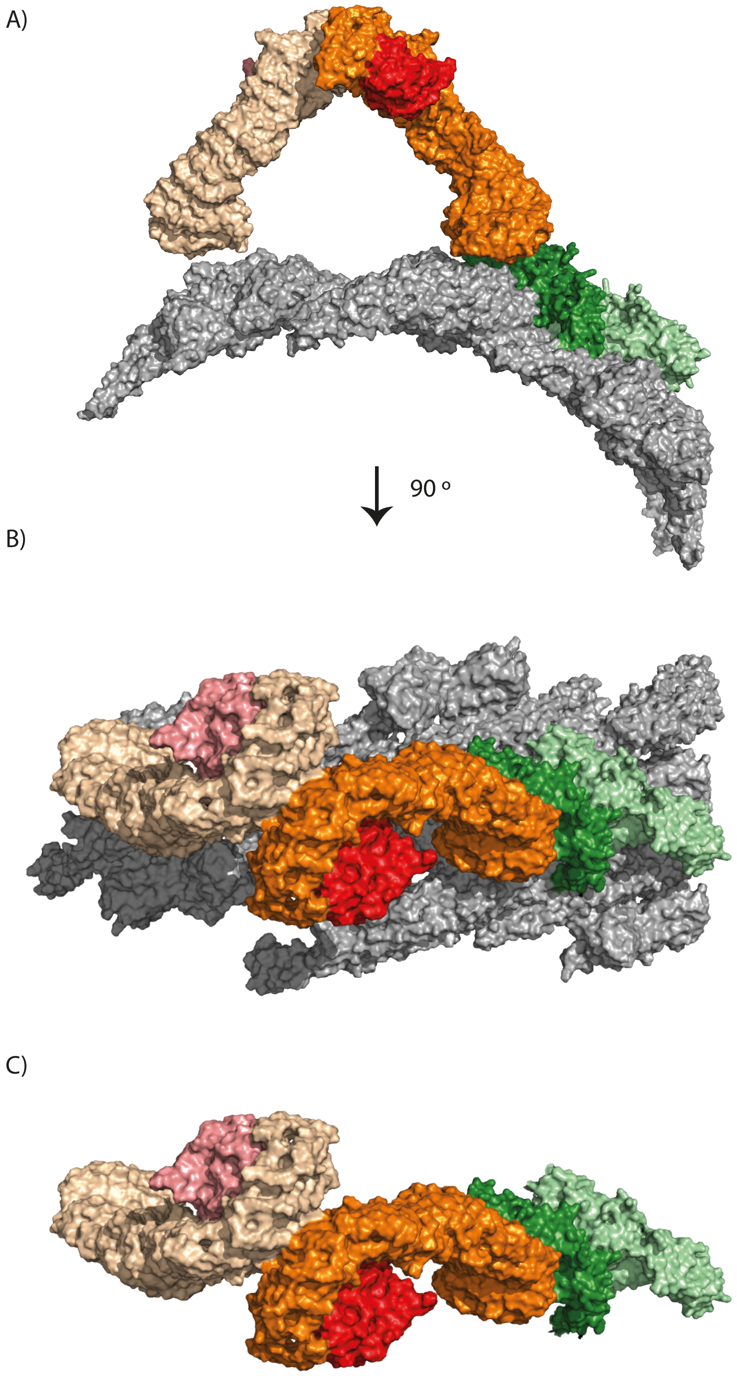


Figure 7: A diagram depicting the organization of retromer core complex, with SNX on the membrane. All from pdb 6H7W, as published by Kovtun et al [149]. A) Vps5 molecules coloured gray, Vps26 in green, Vps35 in orange and Vps29 in red. Vps5 is bound to a PI3P containing membrane (not shown). B) A view 90° of the complex in A). The dimerization of Vps35 is evident in this view. C) A view of B), with the Vps5 component removed.

membrane association of the retromer core complex seems to be somewhat dependent on Rab7 [162, 164, 166]. It has been demonstrated Rab7 interacts with a small area of Vps35, with the same finding for Ypt7 in yeast [162, 167, 168]. However, the Ypt7:Vps35p interaction in yeast is not required for membrane localisation of retromer, as in the case of Ypt7 depletion the heteropentameric yeast retromer complex remains associated with endosomal membranes, probably as a result of the yeast complex containing a tightly-bound SNX-BAR dimer [167, 169]. Also, Kovtun, Leneva et al demonstrated that Rab7/Ypt7 is not needed [149], nor has any effect upon (personal communication) the recruitment of retromer to liposomes when forming tubules *in vitro*. It has also been shown that a Rab5 effector, phosphatidylinositol-3-kinase (PI3K), can promote retromer recruitment onto membranes, presumably by establishing phosphatidylinositol-3-phosphate (PI-3-P) enriched domains at the endosome membrane, which SNX proteins bind to [166].

In addition, the mammalian retromer complex has been shown to associate with the WASH complex, which can regulate actin dynamics by activating the Arp2/3 complex that mediates actin branching [170]. The WASH complex is formed of strumpellin, Fam21, SWIP, ccdc53 and WASH [171], with the WASH complex being recruited to endosomal membranes in part by Vps35 binding to Fam21 [172]. There has also been one report that suggests Esp15 homology domain-containing protein 1 (EHD1) interacts with retromer to stabilise SNX1 tubules [173], although this interaction has yet to have been shown conclusively. EHD family proteins have a dynamin fold, and also bind nucleotide tri-phosphates and therefore it is hypothesised that EHD proteins may be involved in vesicle fission. However, the entire EHD family is missing in yeast.

1.3.2 Retromer-associated Sorting Nexins

Sorting nexin proteins or SNX proteins are characterised by the Phox homology (PX) domain that binds to phosphatidylinositols in membranes, most commonly phosphatidylinositol 3-phosphate (PI-3-P). Most SNX proteins also contain another domain and this allows classification into 5 subfamilies: SNX-PX, SNX-BAR, SNX-FERM, SNX-PXA-RGS-PXC and SNX-MIT proteins. In the case of retromer, the SNX-BAR, SNX-PX and SNX-FERM subfamilies are relevant.

1.3.2.1 SNX-BAR retromer proteins

As previously discussed, SNX 1/2 and SNX5/6 form obligate SNX-BAR dimers that can bind to the retromer core complex. SNX1, 2, 5, 6 can all bind to the retromer core complex [132, 174]. The SNX-BAR proteins contain a banana-shaped Bin-Amphiphysin-Rvs (BAR) domain that is thought to associate with the membrane using basic residues that form electrostatic interactions with the membrane lipids, in addition to the PX domain [175, 176]. Another SNX-BAR protein, SNX32, has high homology to SNX6 and is thought to be a tissue restricted retromer protein, although no studies have been undertaken to confirm this [136]. A review of the SNX-BAR protein subfamily role in endosomal sorting appears in [177].

The retromer SNX-BAR proteins have functions independent of the retromer core complex. Indeed, in mammals it has been demonstrated that SNX1 membrane recruitment is unaffected in Rab7 siRNA knock out cells, in contrast to the retromer core complex components [162]. It is also evident that SNX1 can bind cargo independently of the core complex. GST Pulldown, surface plasmon resonance (SPR) and mutational studies were all employed to demonstrate that SNX1 is capable of binding 10 GPCRs *in vitro* [141]. SNX1 also seems to have a role in binding proteins with a function in trafficking, including RME-8 [178]. RME-8 uncoats clathrin coated vesicles [179] and also binds the WASH complex component Fam21 [178]. It has also been suggested that SNX2 plays a critical role in embryonic development, more so than SNX1 [180].

Yeast 2 Hybrid studies by Wassmer et al. have shown that SNX5 or SNX6 can interact with SNX1 or SNX2 to form a SNX-BAR dimer, but not with each other [181]. Additionally, Wassmer et al. found that SNX5 and 6 can bind directly to the p150$^{\text{glued}}$ component of dynactin, a complex that activates the minus-end-directed microtubule motor dynein [181]. As previously mentioned, research by Simonetti et al [159] has demonstrated that CI-MPR trafficking is dependent not on the retromer core components as previously thought, but instead due to sequence specific interaction with SNX5/6, and this is likely to lead to a re-examination of the role the retromer core plays in cargo binding. Clearly, the SNX-BAR proteins are fundamentally important within the cell.

1.3.2.2 SNX-PX and SNX-FERM retromer proteins

The retromer core complex in mammalian cells binds not only to Vps5p and Vps17p SNX-BAR homologs and orthologs, but also other types of SNX proteins. The trafficking of Wntless is independent of SNX1 and requires the SNX-PX SNX3 instead [143]. In fact, the recruitment of retromer core complex to early, PI-3-P-containing endosomes has been demonstrated to be dependent on Vps35 binding to SNX3 [143, 182]. The structure of the Vps35:Vps26:SNX3 complex shown in Figure 8 elucidates how SNX3 is bound to the retromer core complex. When considered with the retromer core:SNX-BAR dimer structural information, it is evident that the binding of SNX3 and SNX-BAR proteins to the retromer core complex are mutually exclusive. More recently, Cui et al have shown that SNX3 is required for CI-MPR trafficking in GCC88 tethered endosomal carriers [161].

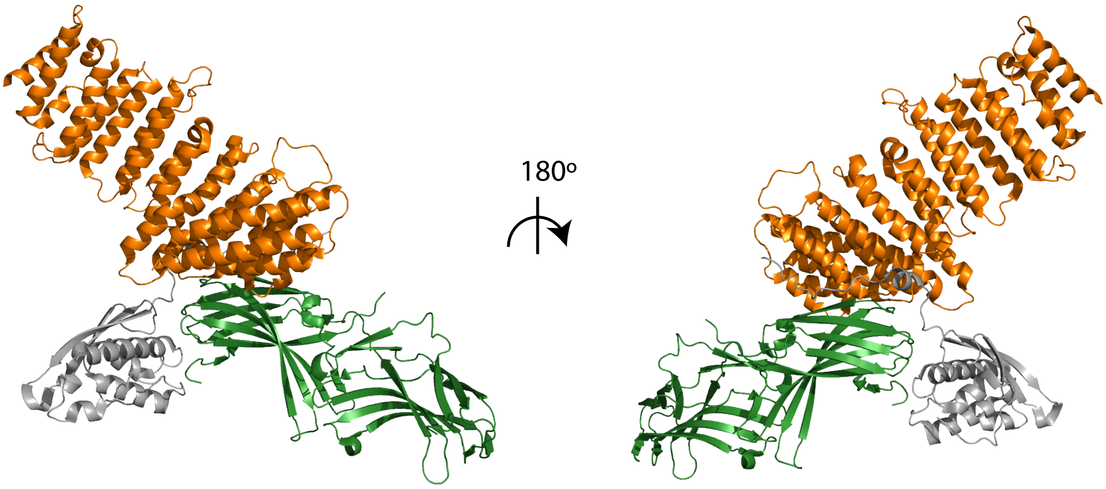


Figure 8: Vps26 and Vps35 in complex with SNX3, from pdb 5FOJ. Vps35 shown in orange, Vps26 shown in green, SNX3 shown in gray.

There is also a role for the SNX-FERM protein SNX27 in retromer trafficking. SNX27 is a SNX-FERM with a PDZ domain that is found only in higher eukaryotes. PDZ domains generally interact with proteins that contain a PDZ-binding motif. However, the PDZ domain of SNX27 is required to bind to Vps26, which doesn't contain the PDZ domain-binding motif [183]. SNX27 can bind PDZ-binding motif containing proteins and Vps26 at the same time, as Vps26 binds to a different part of the PDZ domain than the canonical PDZ domain-binding motif [184]. Additionally, the FERM domain of SNX27 has been shown to interact with SNX1 and also Fam21 of the WASH complex [183, 185, 186]. The structure of SNX27 PDZ domain bound to Vps26, in Figure 9, shows a distinctly different binding patch than for SNX3 which binds via the PX domain, as shown in Figure 8. A model of full length SNX27 bound to retromer derived from small angle X-ray scattering (SAXS) data was published in [184] . The role of SNX27 bound retromer core complex in cargo selection and trafficking, including GLUT1, is now beginning to be determined and it appears that SNX27 seems to particularly mediate endosome to plasma membrane traffic [183, 184, 187].

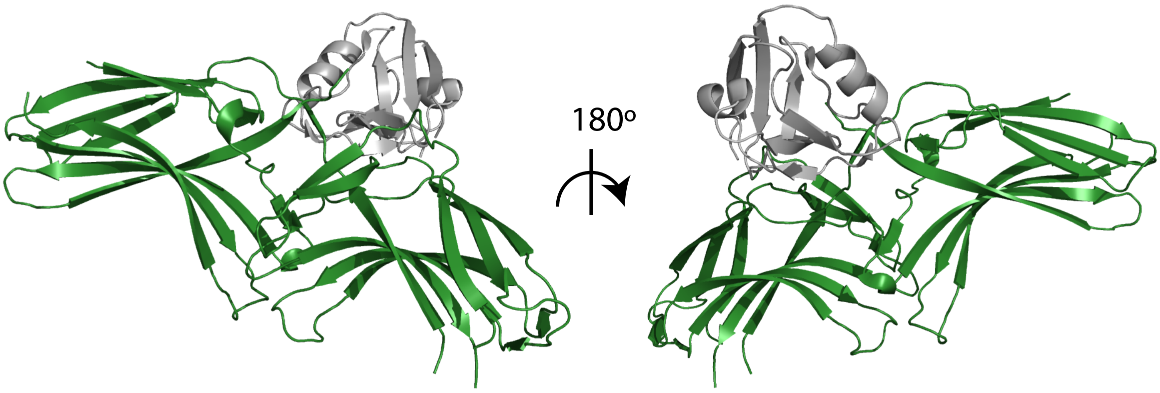


Figure 9: Vps26 (green) in complex with the SNX27 PDZ motif (gray), from pdb 4P2A.

It is clear that the retromer core complex and the associated retromer sorting nexins bind a complement of proteins, not only cargo proteins but also many proteins involved in both the regulation and undertaking of intracellular trafficking. Recent studies have shifted the original concepts of retromer mediated trafficking by redefining the roles of both the core trimer and the SNX proteins that interact with it. The retromer core is no longer thought to be the main 'cargo binding' complex, but instead is likely to have structural roles whilst associated cargo selection is mediated to a significant degree by SNX proteins.

1.3.3 Retromer Cell Biology

In vivo studies provide an opportunity to view the location of proteins throughout the cell in their native environment at a steady state and using live cell imaging provides the ability to dynamically perturb the network. This is of particular use in monitoring cargo transport through the endosomal network. The integral membrane protein glucose transporter 1 (GLUT1) can usually be detected at the plasma membrane (as demonstrated in Figure 10 reproduced from [7]) and once there is the main mechanism for transporting extracellular glucose into the cell. The steady state localisation of GLUT1 is at the plasma membrane, with a small proportion in endo/lysosomal compartments, as GLUT1 cycles from the plasma membrane to endosomes, or is directed to the lysosome for degradation, when appropriate. This may well be dependent on cell type, the state of nutrition or other factors.

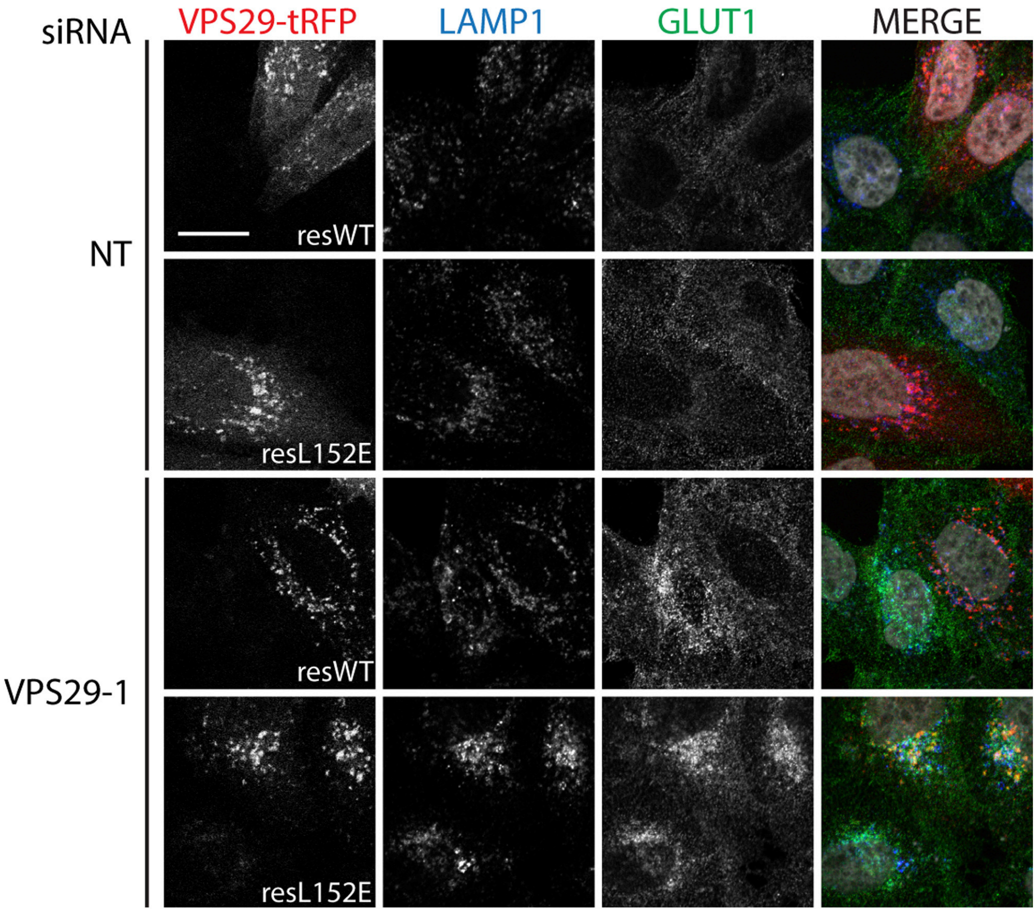


Figure 10: Images and experimental information (below) reproduced from [7] under a Creative Commons License (CC BY 3.0). HeLa cells were knocked down with single siRNA oligonucleotides at 100nM, with either nontargeting control (NT) or Vps29-1 oligonucleotides. Twenty-four hours prior to fixation, cells were transfected with siRNA-resistant Vps29-tRFP WT (resWT) or L152E mutant (resL152E). Cells were imaged by IF confocal microscopy. tRFP, red; LAMP1, blue; GLUT1, green; nuclei, white; merged panels (MERGE). Scale bars, 20μM.

Immunofluorescence confocal microscopy studies by Hesketh et al showed that whilst WT Vps29 could rescue a Vps29 siRNA knockdown GLUT1 trafficking relocalisation phenotype in HeLa cells, the L152E mutant Vps29 rescue experiments could not, resulting in GLUT1 co-localising with lysosomal marker LAMP1 [7], as reproduced here in Figure 10. Clearly, the recycling of GLUT1 from endosomes to the plasma membrane is Vps29, and therefore retromer core complex, dependent. If this recycling mechanism is perturbed, then GLUT1 degradation at the lysosome is increased, presumably as endosomes mature into lysosomes, without normal GLUT1 retrieval. A very similar GLUT1 mistrafficking phenotype is seen upon the siRNA mediated knockdown of SNX27, with GLUT1/LAMP1 colocalization increasing as a result [186], shown in Figure 11. This suggests that GLUT1 is trafficked to the plasma membrane by SNX27-bound retromer, and not SNX-BAR dimer bound retromer.

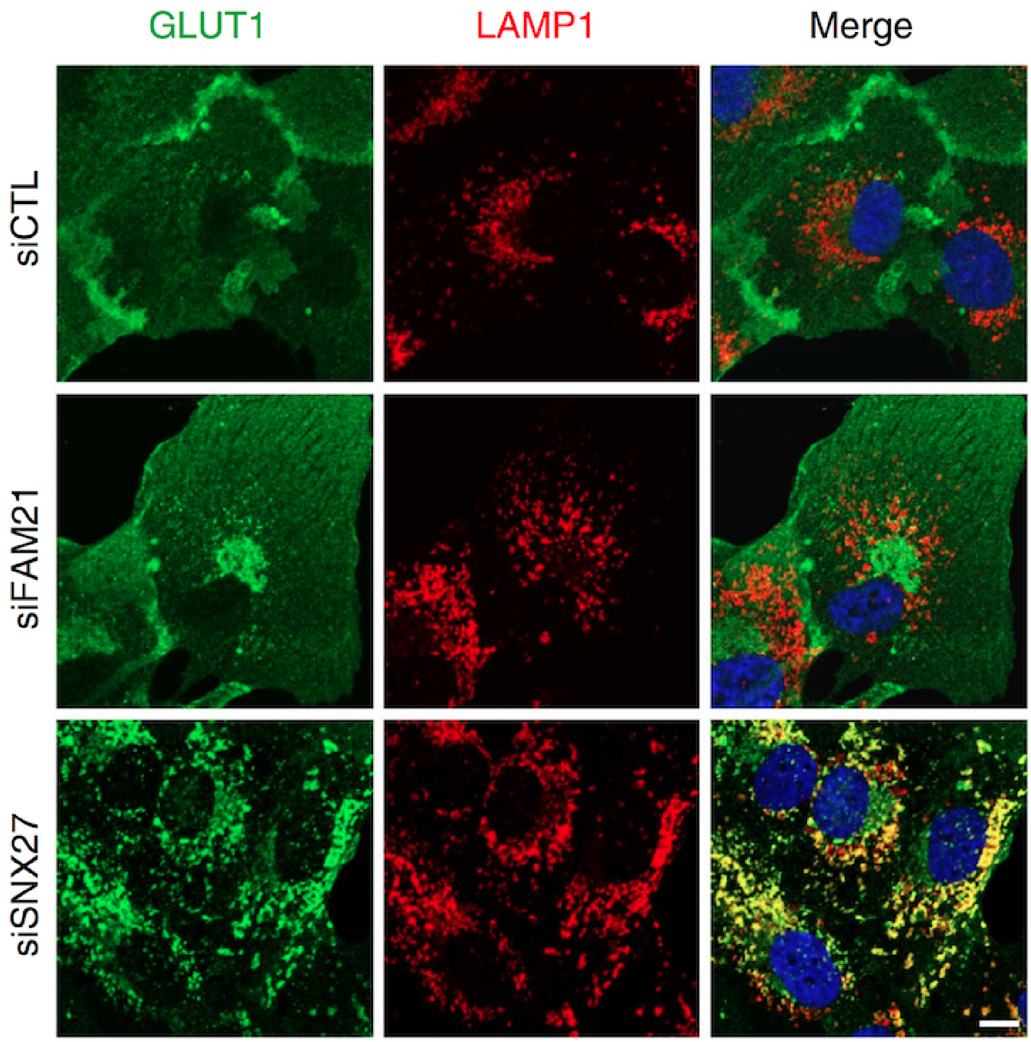


Figure 11: Image and experimental information (below) reproduced from [186] under a Creative Commons license (CC BY-NC-SA 4.0). hTERT-RPE1 cells transfected with the indicated specific or control (siCTL) siRNAs were immunostained for GLUT1 (green) along with LAMP1 (red). Merged images with 4',6-diamidino-2-phenylindole staining (blue) are to the right. Scale bars, 10μM.

Curiously, the depletion of Fam21 (one subunit of the large retromer-interacting WASH complex) resulted in GLUT1 being somewhat more dispersed throughout the cell, as reproduced in Figure 11 from [186]. Lee et al [186], showed that GLUT1 dispersal after Fam21 depletion by siRNA corresponded with an increase in co-localisation of GLUT1 and GM130, a *cis*-Golgi matrix protein. In the same study, Lee et al also demonstrated that Fam21 is needed to recruit SNX27 to membranes [186], so if Fam21 recruits SNX27, and it is a SNX27 retromer pathway that recycles GLUT1 back to the plasma membrane, then one might expect the Fam21 and SNX27 knockdowns to produce similar phenotypes. However this is not the case, a SNX27 knockdown results in GLUT1 co-localising with LAMP1 (Figure 11), whilst a Fam21 knockdown results in GLUT1 co-localising with GM130. One possible model is that Fam21 is part of a decision process in which cargo is prevented from going to the Golgi and subsequently SNX27 selects the cargo destined for the plasma membrane to prevent lysosomal degradation. However, more experimental verification is required to confirm this model, as it is not clear why GLUT1 would be aberrantly trafficked to the *cis*-Golgi, as opposed to the *trans*-Golgi.

The integral membrane protein Cation Independent Mannose-6-Phosphate Receptor (CI-MPR) acts to bind newly synthesised lysosomal hydrolases in the TGN and deliver them to endo/endolysosomal compartments. Once liberated of the cargo, CI-MPR must then be recycled back to the TGN, via endosomal compartments, to continue the cycle. Retromer-dependent CI-MPR recycling activity was long attributed to the retromer core complex, as it was proposed that Vps35 binds directly to CI-MPR. However, recent evidence suggests a direct role for SNX5/6 in this process, independent of Vps35 [159]. Whilst Simonetti et al found that a small proportion of CI-MPR is found in Vps35 positive, SNX2 negative membrane domains, CI-MPR can also be found in SNX2 positive, Vps35 negative membrane domains, as reproduced here in Figure 12. This indicates that Vps35 is not an absolute requirement for CI-MPR localisation. However, it should be noted that the majority of CI-MPR is found in both Vps35, SNX2 positive membrane domains, as shown in Figure 12. The immunofluorescence studies depicted in Figure 12 also show that CI-MPR can also be found on SNX1 positive, Vps35 negative tubules and also SNX6 positive, Vps35 negative tubules that seem to emanate from Vps35 positive endosomal structures.

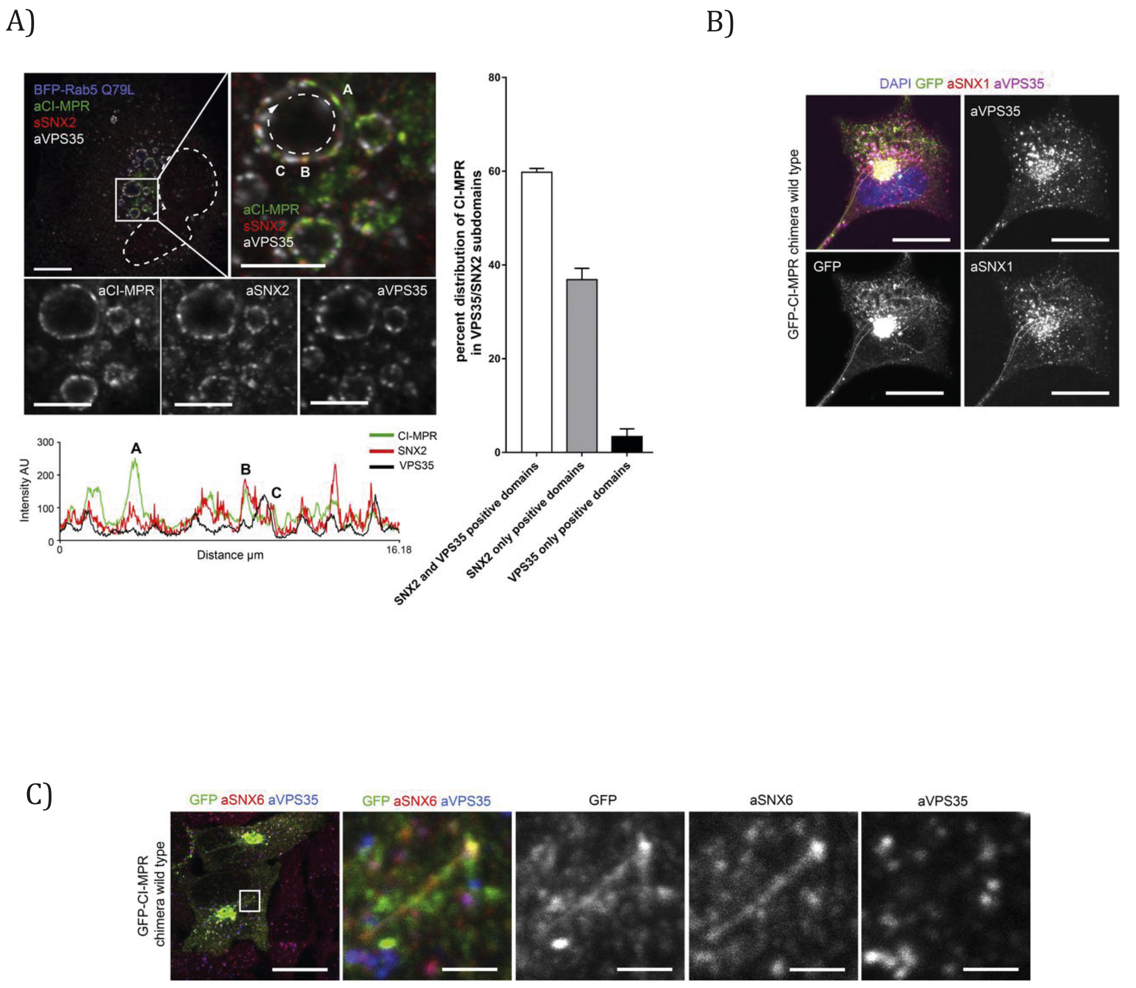


Figure 12: Images and experimental information (below) reproduced from [159] under Creative Commons License (CC BY-NC-SA 4.0). A) CI-MPR colocalizes with SNX2-positive subdomains on enlarged endosomes. HeLa cells were transfected with BFP-Rab5Q79L and immunostained for endogenous CI-MPR, SNX2, and Vps35 after 48h. The dashed line in the top left image indicates the contour of the nucleus, and the dashed line on the top right image refers to the enlarged endosome from which the intensity line scan was measured. Bars (main images) 10μM; (zooms) 5μM. Line scan of signal intensity across the circumference of enlarged endosome. Distribution of CI-MPR in retromer subdomains; n = 3 independent experiments. CI-MPR signal was quantified in 36 enlarged endosomes (means +/- SEM). B) A subpopulation of GFP-CIMPR chimera tubules are decorated with endogenous SNX1 but not endogenous Vps35. HeLa cells were transfected with GFP-CI-MPR chimera WT and immunostained for SNX1 and Vps35 after 48h. Bars, 20μM. C) Some SNX1/2-SNX5/6-decorated CI-MPR-containing tubules were observed to emanate from Vps35-positive endosomes. HeLa cells were transfected with GFP-CI-MPR chimera WT and immunostained for SNX6 and Vps35 after 48h. The box in the leftmost panel indicates the area depicted in the other four panels. Bars: (main images) 20μM; (zooms) 2μM.

The involvement of Vps35 in this recycling process is thus somewhat unclear. A recent hypothesis suggests that Vps35 concentrates CI-MPR on endosomal membranes, after which SNX-BAR proteins bring CI-MPR into tubules for recycling to the TGN, which is a tempting model that brings together the known experimental data [188].

The situation is further complicated when one considers other evidence on the trafficking itinerary of CI-MPR. It is known that AP-1 is responsible for vesicle transport out of the TGN and that clathrin in complex with AP-1 participates in the formation of MPR-containing transport vesicles at the TGN [189, 190]. Indeed, AP-1 and GGAs (Golgi-localised, γ-ear-containing, ADP-ribosylation factor binding proteins) have been shown to play important roles in mediating the transport of CI-MPR from the TGN to endosomes [191-194]. CI-MPR has both the Yxxϕ motif recognised by the μ subunit of adaptor protein complexes, and also the dileucine motif (DXXLL) which was also initially thought to be recognised by AP-1, as it is similar to a known conserved AP γσ heterodimer binding motif [D/E]XXXL[L/I], but which was later shown to bind to GGAs instead [191] (reviewed in [190]).

Curiously, CI-MPR was seen to accumulate in early endosomes in cells lacking AP-1 [195], which suggests a role for AP-1 in retrograde transport, and also suggests that GGAs can compensate for the lack of AP-1 in the anterograde direction. Additionally, it was demonstrated that the AP-1 binding protein epsinR is needed for retrograde transport of MPRs [196]. The role of AP-1 in retrograde endosome to TGN transport is now firmly established. Clearly, there are multiple trafficking pathways of MPRs (including CI-MPR) from the endosome to the TGN, and it has been suggested that the multiple trafficking mechanisms may reflect the different stages of endosome maturation within the cell.

As mentioned previously, Vps29 has not been shown to be involved in cargo binding, however Vps29 is known to bind the RabGAP protein TBC1D5. TBC1D5 is thought to be a Rab7 GAP, meaning that it activates Rab7 GTPase activity, thus promoting the conversion of Rab7GTP to Rab7GDP and allowing Rab7GDP clearance from the membrane by GDI. TBC1D5 remains membrane associated when siRNA knockdown of Vps29 is rescued by WT Vps29, however not when L152E mutant Vps29 is used as a rescue, as shown by Hesketh et al [7], and reproduced here in Figure 13A. TBC1D5 has been shown to bind to a patch on Vps29 containing the L152 residue [197]. When TBC1D5 expression is silenced, retromer core complex levels at the membrane seem to be elevated, and it has been speculated that this is due to artificial longevity of Rab7GTP at the membrane in the absence of TBC1D5 [165]. Presumably, the usual role of the putative Rab7 GAP TBC1D5 is to inactivate Rab7GTP to promote the removal of the late endosomal membrane identity of the emanating cargo laden tubules and if this does not occur, the formation of retromer tubules is stalled.

In the same 2014 study, Hesketh et al demonstrate that in untreated, cytosol extracted, Vps29-tRFP transfected cells, Varp and Vps29 co-localise very well in punctate structures [7], which is reproduced here in Figure 13B . Upon siRNA knockdown of Vps29, it seems that there is much less Varp on the membrane and even when rescued by WT Vps29-tRFP there is less Varp on endosomal membranes, but the Varp that is on the membrane is co-localised with Vps29.

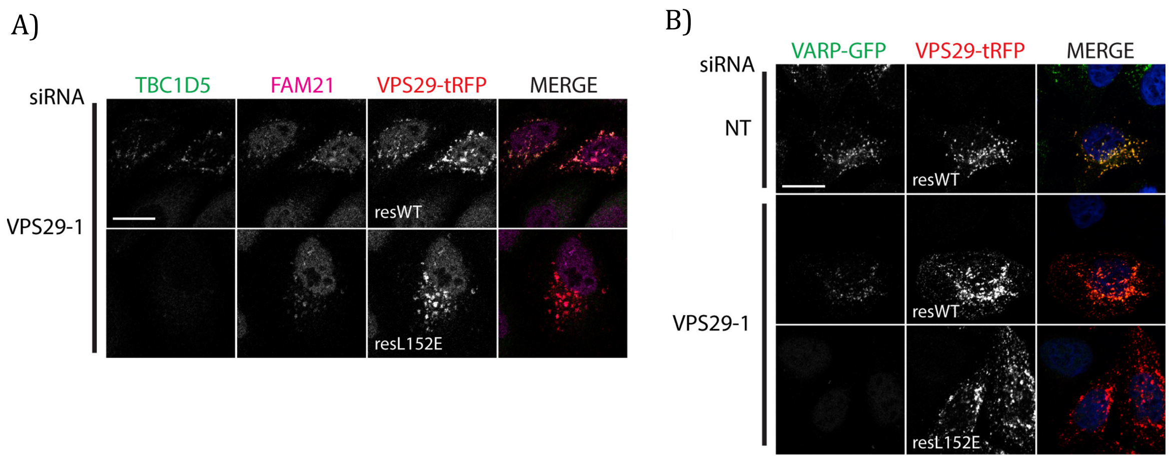


Figure 13: Images and experimental information (below) reproduced from [7] under a Creative Commons License (CC BY 3.0). A) IF confocal microscopy of cytosol-extracted HeLa cells following knockdown with 100nM Vps29-1 siRNA oligonucleotides at 100nM (NT, Non-targeting control, Vps29-1). Twenty-four hours prior to fixation, cells were transiently transfected with Vps29-tRFP WT or L152E mutant (TBC1D5 [green], FAM21 [magenta], Vps29-tRFP imaged by native fluorescence, and nuclei [blue, merged panels]). Scale bars, 20μM). B) IF confocal microscopy of cytosol-extracted Varp-GFP cells knocked down using single siRNA oligonucleotides at 100nM (NT, non-targeting control, Vps29-1). Twenty-four hours prior to fixation, cells were transiently transfected with Vps29-tRFP WT or L152E mutant. MERGE, merged panels. Scale bars 20μM.

Rescue with an L152E mutant of Vps29 completely abolished Varp recruitment to the endosomal membrane, as shown in Figure 13B, similar to the case of TBC1D5.

It should be noted that Vps29 has also recently been found in a non-retromer complex that also acts to recycle cargo from endosomes to the plasma membrane. The retriever complex described by McNally et al, shows that Vps29 can form a complex with DSCR3 and C16orf62, with the overall structure of the complex thought to be similar to that of retromer [198]. The retriever complex utilises the SNX-FERM protein SNX17 to recycle retromer-independent cargo such as α5β1 integrins back to the plasma membrane [198, 199]. In line with the known role of the WASH complex in non-retromer cargo sorting [199], the knockout of Fam21 results in the loss of retriever on the membrane, and subsequent mis-sorting of α5β1 integrins into lysosomes [198]. It has been reported that WASH complex localisation to endosomes was as a product of the interaction between Fam21 and Vps35 of the retromer core complex [172, 200]. However, McNally et al found that whilst Fam21 did become more cytosolic on knockout of Vps35, some Fam21 remained co-localised with SNX1 endosomes and so the membrane localisation of the WASH complex does not appear to be entirely dependent on retromer [198]. An additional complex, the CCC complex, formed of CCDC22 and CCDC93 and various COMMD proteins was also shown to be targeted to membranes by Fam21; and both the CCC complex and WASH complex are required for SNX17 mediated retriever recycling of α5β1 integrins to the plasma membrane, with no retromer requirement [198]. However, in their analysis McNally et al did discover that some SNX17-dependent cargoes levels in the plasma membrane are depleted under SNX27 or Vps35 suppression conditions [198].

Clearly then, there may be some overlap between the retriever and retromer pathways, and this can be demonstrated in the case of amyloid precursor protein (APP). Amyloid precursor protein is processed to the neurotoxic β-amyloid peptide (Aβ) as part of the pathogenic progression of Alzheimer's disease (AD). APP is known to be directly bound by SNX17, and siRNA knockdown of SNX17 decreases cellular APP levels, and also causes an increase in Aβ generation, as determined by Western Blotting and ELISA techniques [201]. Additionally, a large body of evidence exists to demonstrate the role of retromer in APP trafficking. The role of retromer in AD was first described in

2005 by Small et al., where analysis of post-mortem AD patient brain samples revealed that Vps35 and Vps26 levels were reduced in regions specifically vulnerable to AD [202]. Subsequent works have described that many genes involved in the APP recycling process may have AD pathology-related mutations, including SNX1, SNX3 and Rab7a [203], whilst the Vps10-domain containing family receptors (SorL1/SorLA, SorCS1, SorCS2, SorCS3) have variants that are associated with AD [204], two of which SorL1/SorLA and SorCS1 have been shown to bind to retromer [157, 205]. The role of retromer in pathological processes is still the subject of much investigation, with many detailed reviews describing how retromer deficiencies may be responsible for not only disordered APP processing, but errors in both developmental and neurologic systems, including mutations of Vps35 in the development of Parkinson's Disease [14, 206-209].

One hypothesis on how retromer deficiencies result in AD is that as a result of impaired retromer function, SorL1/SorLA trafficked APP [210] is held at the endosome for a longer amount of time than usual, which increases the probability of APP processing to pathogenic Aβ, as initially conducted by β-site APP cleaving enzyme (BACE) in endosomes (Reviewed in [14]). Additionally, there is a possibility that AD-promoting effects arising from mis-trafficking of CI-MPR in retromer deficient cells. CI-MPR binds and transports protease Cathepsin D to pre-lysosomal compartments, including endosomes. In retromer-deficient cells, it is thought that CI-MPR is retained in endosomes as a result of the impaired endosome to TGN retrograde trafficking, this seems to suggest that AP-1-mediated CI-MPR retrograde trafficking cannot completely compensate for the loss of retromer-mediated trafficking in this case. Since CI-MPR is not recycled efficiently, it fails to pick up Cathepsin D at the TGN and thus Cathepsin D levels in the endosome fall which subsequently results in improper processing of Tau, a microtubule associated protein that causes neurofibrillary tangles, discussed in a recent review [14]. The evidence for this is still inconclusive however, and so the recent finding by Simonetti et al that CI-MPR does not require the retromer core complex for recycling [159] will surely lead to further investigation into the precise role of the retromer core complex and various associated SNX proteins in AD. Not only this, but the discovery of the retriever complex will likely add complexity to this discussion, asking if SNX17 acts independently of the retriever complex in APP trafficking, or if the retriever complex is as important in APP trafficking as retromer is proving to be.

1.4 Varp

The finding by Hesketh et al, showing that the retromer core recruits the protein Varp to endosomal membranes provided the foundation for the work in this thesis [7]. Although the nature of this interaction was speculated upon, it was felt that structural characterization of the interaction would greatly improve the understanding of this interaction.

First identified by Zhang et al. in a Yeast 2 Hybrid screen for Rab21 effector proteins [2], Varp is a Vps9 domain containing protein consisting of 1050 residues and multiple functional domains, as shown in Figure 14. Subsequent studies have identified that Varp has other binding partners, including other Rab GTPases: Rab32/ Rab38, and Rab40c, and also the R-SNARE, VAMP7. As a result of this interaction network, Varp likely has a role in late endocytic trafficking, not only of receptors to the plasma membrane but the trafficking of melanogenic enzymes to melanosomes as well. Additionally, with some highly conserved regions in Varp yet to be studied, there is a possibility that the interaction network of Varp is much larger than currently understood.

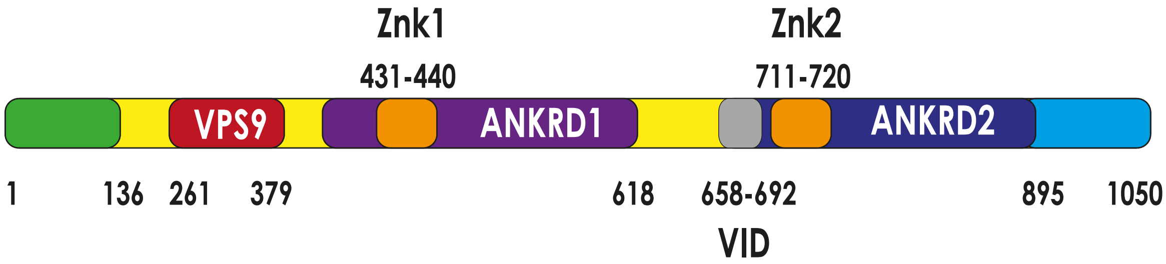


Figure 14: A schematic of the domain architecture of Varp.

1.4.1 Domain organization and known binding partners of Varp

The Rab21 guanine nucleotide exchange factor (GEF) activity of Varp is conferred by the Vps9 domain, and this binding activity is needed for neurite outgrowth [211]. Whilst there are multiple Vps9 domain-containing proteins in humans, only Rabex-5 and Varp have Rab21-GDP binding activity [212, 213]. Although the localisation of Rab21 is still yet to be conclusively determined, Rab21 has been detected at the plasma membrane with a small proportion located on recycling endosomes [214], (personal communication, David Owen). It has been demonstrated that Rab21 binds to α integrins to direct the internalised integrins to endosomes and a constitutively active (GTP-bound) Rab21 mutant (Rab21 Q76L) showed an increased ability to bind integrins compared to WT Rab21 [215]. As Varp acts to convert Rab21GDP to Rab21GTP, there is a strong possibility that integrin recycling will involve Varp. Not only this, but as the retriever-interacting SNX17 has been demonstrated to act in the recycling of integrins back to the plasma membrane [199], it is probable that Vps29, a component of both the retromer and retriever complexes and a Varp binding partner, is involved as well.

Since the initial discovery of Varp, it has been demonstrated that Varp interacts with a multitude of different proteins in the cell. For instance, the ANKRD1 domain of Varp is required for binding Rab32/38, however, unlike for Rab21, Varp has a role as a Rab effector protein instead of a GEF [3, 4, 216]. The structure of the Rab32:Varp ANKRD1 complex is shown in Figure 15B. Rab32 and Rab38 influence TYRP1 trafficking from endosomes to melanosomes in melanocytes, and the Rab32/38 effector activity of Varp is required for this [4]. Mutations in Rab38 cause a phenotype of chocolate coat colour in mice, which is thought to be due to the defective trafficking of TYRP1 in these cells [217]. An interaction between Varp and VAMP7 has also been shown to be critical in the Rab32/38 mediated trafficking of TYRP1 to melanosomes [216].

The interaction of Varp with the R-SNARE VAMP7 is mediated by the first ankyrin repeat of the Varp ANKRD2 stack (~658-692 of Varp) packing on top of the longin domain and trapping the portion of the VAMP7 SNARE motif between them. The structure of the VAMP7:Varp complex is shown in Figure 15A. The SNARE motif is required in order to make fusion-competent four helix bundle SNARE complexes with 3 Q SNARE motifs, and hence Varp binding to VAMP7 can prevent VAMP7 from forming SNARE complexes [5, 129, 216]. In non-specialised cell types, VAMP7 normally cycles from endosomes/endolysosomes to the plasma membrane and back, whilst a Varp-binding deficient mutant of VAMP7 mislocalises to the TGN [7, 30]. More recently, Dennis et al have demonstrated that VAMP7 mediates fusion of tubular carriers that carry the melanogenic enzyme TYRP1 with melanosomes [218]. Not only this, but they showed that the recycling of post-fusion VAMP7 away from melanosomes after cargo delivery uses tubular carriers that are distinct from the incoming, TYRP1 positive, cargo delivery tubules and are also positive for both Varp, and Rab38 [218]. Interestingly, mutation of the retromer binding site in Varp (double HL mutant: H432S, L434A, H712S, L714A) did not prevent Varp localising to melanosomes, and in fact VAMP7 and Rab32/38 alone are required to recruit Varp to melanosomes [218].

The ANKRD2 domain, although very similar to ANKRD1, doesn't show any binding activity to Rab32/38 but instead binds to Rab40c [9]. Rab40c promotes the proteasomal degradation of proteins, including Varp. Not only this but RACK-1 (receptor of activated protein kinase C 1) has been shown to bind to ANKRD2 of Varp and RACK-1 binding seems to compete with Rab40C, and thus can protect Varp from degradation [10]. Both KIF5A and golginA4 have also been reported to bind to Varp, although the binding sites of these have yet to be confirmed [6]. Preliminary data from the Luzio (CIMR) laboratory suggests that it may be a plasma membrane coiled coil tether and not golginA4 that is a true Varp binding partner, although this remains to be conclusively proven (personal communication).

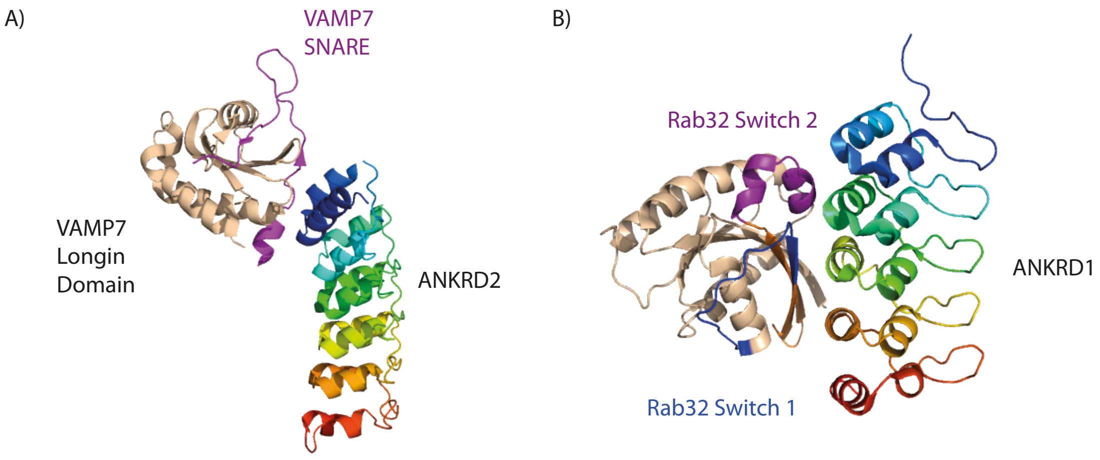


Figure 15: Structure diagrams of Varp with binding partners. A) The longin domain of VAMP7 with the Varp VID (VAMP7 Interaction Domain), as stabilized by ANKRD2.

Image from pdb 4b93. B) Rab32 in complex with the Varp ANKRD2 domain, from pdb 4cym.

Varp also binds to the Vps29 subunit of the retromer core complex, via the two ‘zinc knuckle’ domains [7], depicted as ‘Znk1’ and ‘Znk2’ in Figure 14 and the highly conserved sequences of which are shown in Figure 17. Co-localisation studies demonstrate that Varp and Vps26 have high co-localisation on both punctate and tubular structures within the cell, [7] as reproduced here in Figure 16. Indeed it has been shown that retromer, not VAMP7 or Rab32, is responsible for recruiting Varp to endosomal membranes in HeLa cells, and that this localisation can be disrupted by knockdown of either Vps29, or Vps35 [7], as reproduced here in Figure 16. The images in Figure 16 also demonstrate that Varp co-localises well with Fam21, and also that whilst Varp is present on EEA1 containing structures, Varp and EEA1 seem to exist in separate microdomains. A recent review of the roles of Varp in endosomal trafficking is also available in [212].

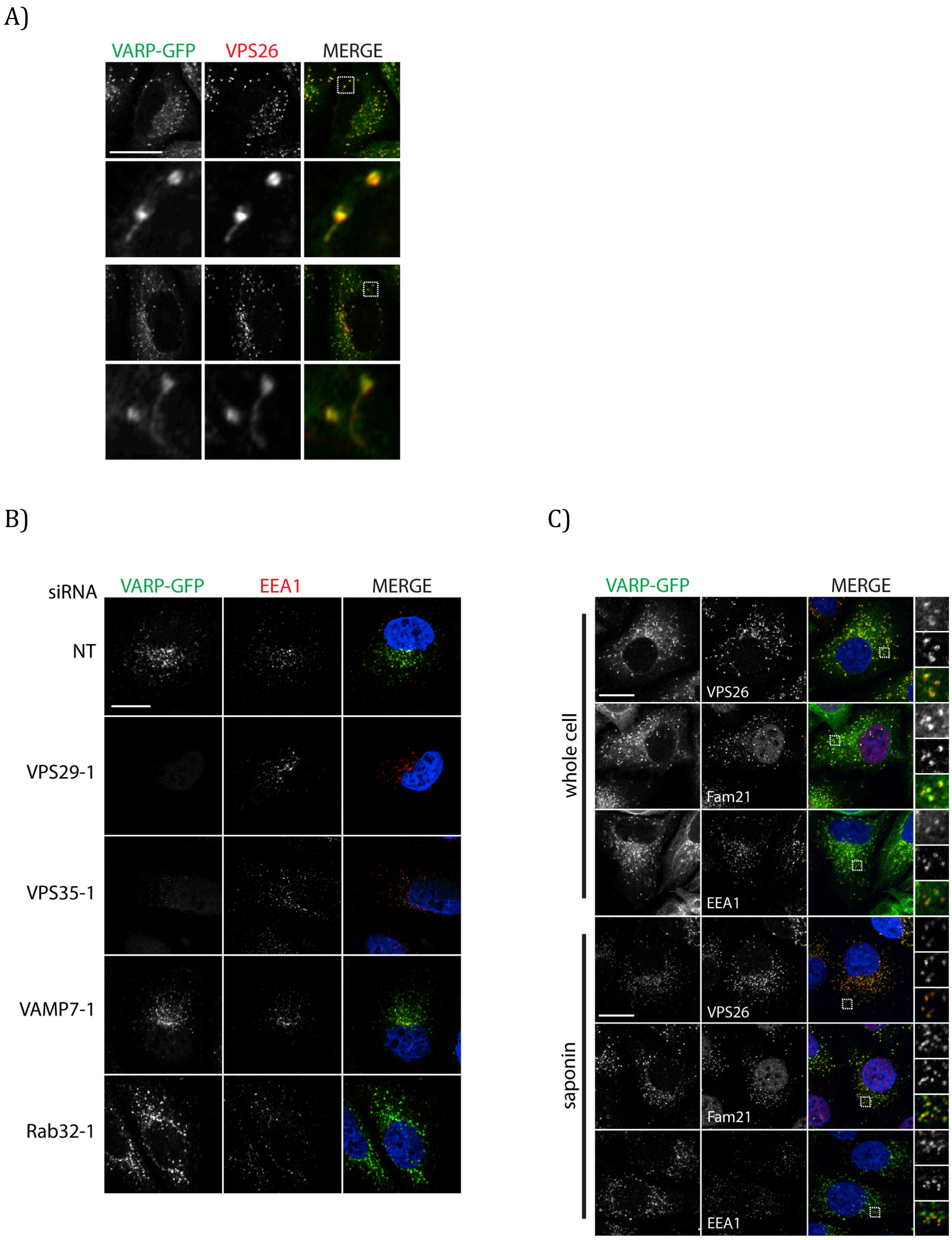


Figure 16: Images and Experimental information adapted from [7], under a Creative Commons License (CC BY 3.0). A) IF confocal microscopy image of VARP-GFP cells. GFP, green; Vps26, red; the boxed regions (shown in the MERGE column) are enlarged below each image, demonstrating colocalization of VARP-GFP and Vps26 on both vacuolar and tubular domains of endosomes. Scale bars, 20μM. B) IF confocal microscopy of cytosol-extracted HeLa cells stably expressing Varp-GFP, which were knocked down using single siRNA oligonucleotides at 100nM (NT, non-targeting control, Vps29-1, Vps35-1, VAMP7-1, Rab32-1). GFP, green; EEA1, red; nuclei, blue, merged panel (MERGE). Scale

bars, 20μM. C) IF confocal microscopy of Varp-GFP cells. GFP, green; and either Vps26, Fam21, or EEA1, red; nuclei, blue, merged panels. Cells were either fixed intact (whole cell) or after cytosol extraction (saponin). Boxed regions in the merged panels are shown as separate green (top), red (middle), and merged (bottom) channels on the right. Colocalization coefficients for Vps26, Fam21, and EEA1 versus GFP were as follows, respectively (saponin only): Vps26, 0.56; Fam21, 0.16; EEA1, 0.67. Scale bars, 20μM.

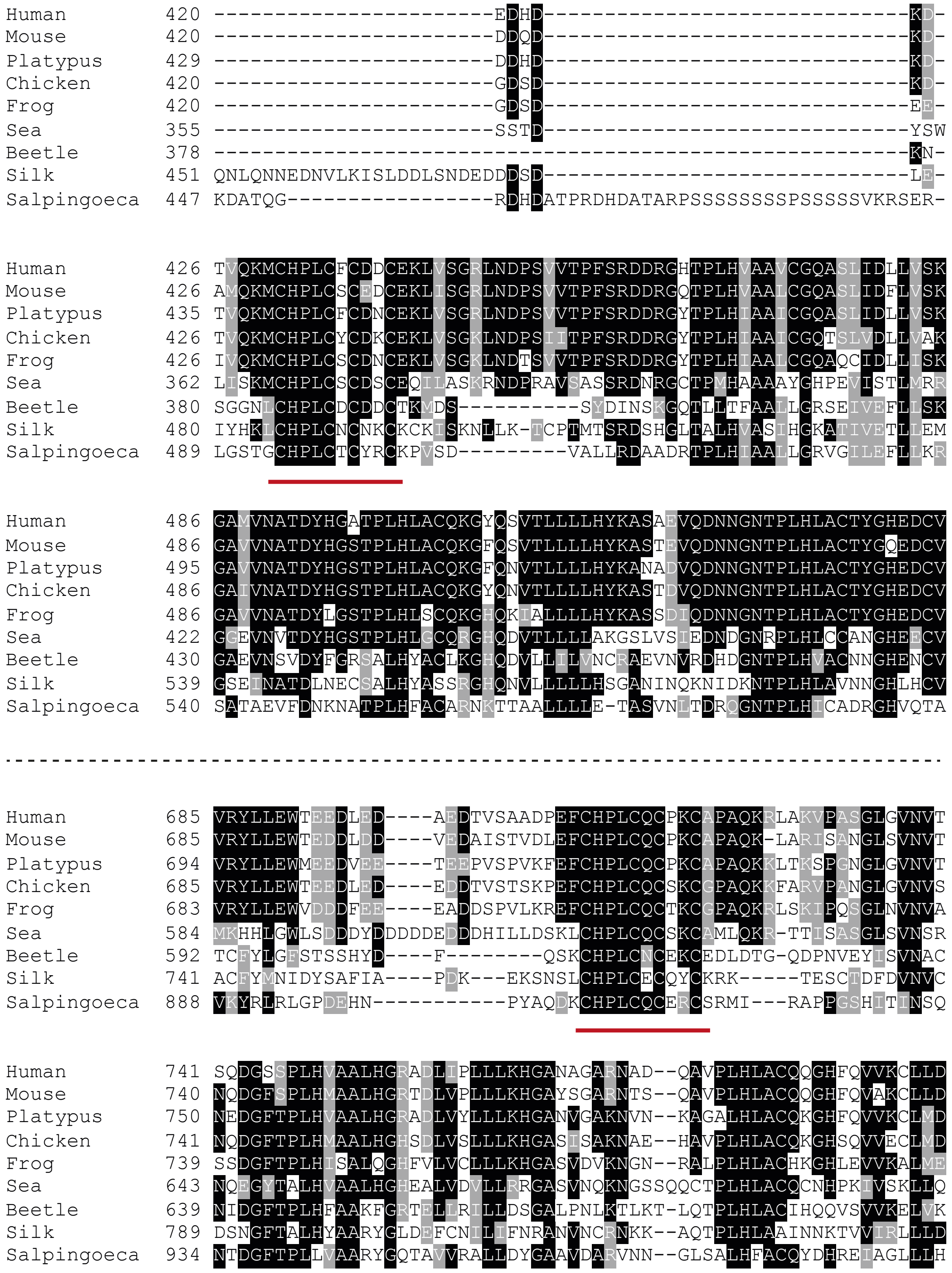


Figure 17: Sequence Alignment of Varp, with zinc binding sequences highlighted by red bars. Aligned using TCoffee, and formatted using BoxShade. Black shading represents a widely conserved amino acid, whilst gray shading indicates amino acid similarity within a defined group. See Section 2.1.1 for details.

Interestingly, it is only recently that a Vps9-containing Varp-like protein has been described in yeast, named Varp-like Protein 1 (VRL1). This is due to a frameshift mutation in all laboratory strains of yeast which means that a premature stop codon prevents the expression of VRL1 [219]. Analysis suggests that VRL1 has no Rab 32/38 binding site, no VAMP7 binding site and no Cys-Rich region but it probably contains a number of ankyrin repeats and the N-terminal 136 residues are very conserved in higher species [219]. Despite the lack of the Cys-rich region, Vrl1 has been genetically linked to retromer trafficking [219].

1.4.2 Zinc binding domains in biology

Zinc binding domains are an abundant feature of eukaryotic proteins, with much research being conducted into both their structures and functions. They have been classified in terms of their structural similarity into different fold groups, as summarised in reviews by Krishna et al. [220], Ravasi et al. [221] and more recently, Andreini et al [222] . In the case of Varp, the zinc knuckle Znk1 and Znk2 domains bear the most similarity to the 'zinc binding loop' category, described by Krishna as follows: '*at least three zinc ligands are very close to each other in sequence and are not incorporated into regular secondary structural elements*' [220]. In the Znk1 and Znk2 domains of Varp, the 4 zinc-binding ligands are all close to each other in sequence, as highlighted in Figure 18, invariantly spaced throughout evolution (shown in sequence alignment in Figure 17), similar to the loop in 45kDa polypeptide Rpb3 of DNA directed RNA polymerase (PDB: 1I3Q-C).

```
Znk1    428  QKMCHPLCFCDDCEKL 443
Znk2    708  PEFCHPLCQCPKCAPA 723
```

Figure 18: The sequences of the two zinc knuckles in hVarp.

Whilst other classification schemes exist, grouping zinc-binding domains by ligand geometry or type, classification based on structure may help to understand or even predict the function of a particular zinc-binding domain. A variety of functions can be associated with zinc binding domains, due to the diverse geometry, ligand coordination and structure. For example, zinc fingers have long been known to be present in DNA and RNA binding proteins and to be involved in homo-multimerization.

This thesis presents the NMR-derived structure of the Vps29:Varp Zinc Knuckle 2 complex, as well as further biochemical and biophysical characterisation of complex formation.

1.5 Nuclear Magnetic Resonance Spectroscopy

Liquid state Nuclear Magnetic Resonance has been used in the study of biological macromolecules for many years, and with modern advances in technology, NMR remains the leading technique for studying small protein complexes at the atomic level, in the solution state. The use of NMR was considered in this project because attempts to crystallise the Vps29:Varp Zinc Knuckle 2 complex had previously been unsuccessful, although the structure of Vps29 has previously been determined by X-Ray crystallography in both the apo state and in complex with other protein ligands. Whilst Vps29 is a well-folded and highly structured protein, the protein sequence either side of both of the CHPLCxCxxC motifs in Varp are thought to be very flexible and unstructured. It is also worth noting that the interaction between the two proteins is reported to be relatively weak, in the micromolar range [7], hence making a homogenous sample as would be needed for crystallisation would be very difficult.

When using NMR techniques, the protein of interest is directly observed in a state more representative of physiological conditions than in X-Ray crystallography, and allows for the dynamic nature of proteins to be considered. However, the larger a protein is, the more difficult it becomes to make residue and atom-specific assignments for the individual signals from the protein, as is required for the interpretation of essentially all NMR data.

1.5.1 Structure determination

1.5.1.1 Resonance assignment

Determining which resonances arise from which atoms in a protein can be an involved process, particularly if the NMR spectra contain many overlapped signals, as can happen for instance when there are many residues of the same type. In the first stage of the process, a variety of experiments based on J-couplings are used. These link atoms that are connected through one, or (in the case of smaller couplings) a few, chemical bonds, and they are used to find atom-to-atom connections to establish spin systems. Each spin-system encompasses the atoms within an amino acid residue. In this case, ^{15}N-HSQC spectra were recorded to correlate N and NH resonances for each residue. Using this information as a starting point, 3D experiments such as CBCACONH and CBCANH, were then used to link the amide proton to other atoms in the peptide backbone, namely Cα, Hα, Cβ, and Hβ of both the self-residue spin-system (i) and the preceding residue spin-system (i-1). The Cβ chemical shifts of some amino acid types are distinct enough to be able to determine the residue type of a spin system uniquely at this stage. For example, Glycine, Serine and Threonine all have diagnostic Cβ shifts that usually allow them to be uniquely identified. Finding connections between sequential spin-systems (sequential assignment) further aids the process, as each connected string of spin systems can then be placed in the context of the entire protein sequence to assist in residue specific assignment.

Further extending the resonance assignment to encompass the sidechain atoms involves using HCCH-COSY and HCCH-TOCSY type experiments. Both HCCH-COSY and HCCH-TOCSY experiments correlate Carbon and Hydrogen atoms of the residue sidechain through the chemical bonds, however the HCCH-COSY detects only those atoms that are covalently connected through pathways of J-couplings involving only a single ^{13}C-^{13}C coupling, whereas the HCCH-TOCSY detects connections over pathways involving more than one ^{13}C-^{13}C coupling. The quantity, and chemical shift, of resonances within a residue are distinctive of sidechain type, thus further residue type assignments can be made and previous assignments confirmed at this stage.

Whilst these experiments allow the identification of a high proportion of the resonances within a protein, there are additional experiments that are needed to assign the sidechain NHs of Asn, Gln, Arg and Trp and also the aromatic rings. NOESY-type experiments (see below) in conjunction with the ^{15}N-HSQC can be used to assign sidechain NH groups of Asn, Gln, Arg and Trp whilst suitably optimised ^{13}C experiments (including HSQC and NOESY) can be used to determine the aromatic residue sidechain assignments.

The assignment process as used in this project is described in greater detail in Chapter 4, particularly in Section 4.2.

1.5.1.2 Structural constraints

Structural constraint information can be gathered by the use of Nuclear Overhauser Effect (NOE) experiments. NOE cross peaks arise as a consequence of dipolar cross-relaxation, which occurs when two NMR-active nuclei, normally protons, are close in space (typically < approx. 5Å apart). Commonly used NOE-based experiments include 2D (^{1}H,^{1}H)-NOSEY (which identifies pairs of protons that are close together in space) and 3D HSQC-NOESY (which further correlates one of the NOE-linked protons to its corresponding directly bonded ^{13}C or ^{15}N atom, thereby dispersing the signals into a third frequency dimension and so helping to relieve signal overlap). In addition, the sizes of J-couplings that span three chemical bonds can sometimes be analysed to give angular torsion constraints.

1.5.1.3 Structure calculation

Many programmes, such as XPLOR-NIH [223], are available to determine protein structures using mainly NOE-derived structural constraints together with other constraints such as those from J-couplings. In this work, we adapted a hybrid NMR/X-ray crystallographic approach recently published by Eustermann et al. [224], whereby pre-existing X-ray co-ordinates are used to constrain the conformation of parts of a structure, whilst other parts are allowed to evolve during the calculation so as to satisfy NMR-derived structural constraint information [224]. In the original work, this approach was used to calculate the structure of a complex of a two-domain fragment of the protein PARP-1 that binds with one domain each side of a DNA single-strand break; the conformation of each half of the complex (i.e. one protein domain and the region of DNA double-helix to which it binds) was constrained to resemble crystal structures of the corresponding domain bound to a DNA double-strand break, while the relative arrangement of the two halves of the complex was constrained by the NMR data, which in that case comprised a combination of RDCs (residual dipolar couplings) and NOEs. In the present work, we used a pre-existing crystal structure of Vps29 (pdb 2R17) to specify the conformation of the Vps29 component in the complex, while allowing the conformation of the Varp fragment to evolve under the NMR-derived constraints, which here comprised a mixture of intra- and inter-molecular NOEs and some J-coupling-derived angle constraints for the Varp fragment. A more detailed description of the method appears in Chapter 2, Section 2.3.3.

1.5.2 Chemical shift perturbation

Chemical shift perturbations (CSPs) are determined by measuring the chemical shift differences between free and ligand-bound states of a protein. Whilst CSPs are most commonly used to identify a region likely to be a ligand-binding site, quantitative analysis of CSPs during a titration with ligand can also provide information about binding affinities. The most widely used resonances for CSP studies are the ^{1}H and ^{15}N signals of amide groups in isotopically labelled proteins; these are particularly sensitive as they respond to small structural changes that can affect the hydrogen bonding network within a protein [225]. However, if a protein changes its conformation as a result of ligand binding, this can cause structural movements that may sometimes be distant from the ligand binding site, and recognising such secondary effects within CSP data can be difficult.

1.6 Thesis Goals and Implications

In very general terms, the purpose of the work described in this thesis is to deepen the understanding of the complex machinery involved in cellular trafficking processes, more specifically to explore the role of Varp by way of its interactions with proteins within the endocytic trafficking system and larger protein complexes, especially retromer.

The structure of the Varp-Vps29 complex is important for understanding the way in which Varp acts in concert with other Vps29 binding proteins to orchestrate the movement of cellular material. To date, some Varp domains have been studied at a molecular level, which has led to mechanistic understanding of the function of Varp within the cell. However, there are still regions of Varp that have yet to be characterised so that their roles can be defined, and this applies particularly to the 'zinc knuckles' that are the main focus of this work. Another question is, what is the function of the N-terminal region of Varp? There are sections of the N-terminus that seem to be at least as well conserved as the zinc knuckles of Varp which is very suggestive of important intracellular function. There are also protein-protein interactions reported in the literature that have yet to be mapped to binding sites on Varp, and it may be possible that the N-terminal region contains these binding sites.

How do the small zinc knuckles of Varp enable binding to Vps29, what is the role of zinc, and does binding of Varp affect the ability of Vps29 to form complexes with other proteins? Vps29 is known to bind not only other subunits of retromer core, but also the potential RabGAP TBC1D5, as well as being a target for the subversive role of RidL in *Legionella Pneumophila* [163]. Solving the structure of the Varp:Vps29 complex should determine if binding to any of these interacting proteins is mutually exclusive, or if several interactions can occur at the same time, thus allowing a deeper understanding of retromer core complex function within the cell. Moreover, since the architecture of retromer core and SNX-BAR has been determined, information about Varp binding can be considered in the framework of the entire retromer coat.

This thesis presents the structure determination of the Vps29:Varp Znk2 complex, determined by NMR spectroscopy and utilizing information from a previously determined X-Ray crystal structure of Vps29, pdb 2R17. This is supplemented by biochemical and biophysical techniques to further characterize the nature of this protein-protein interaction. This allows structure directed mutants to be designed to probe the *in vivo* function of in collaboration with DJO and JPL labs (Cambridge Institute of Medical Research). In addition, studies undertaken with the N-terminal region of Varp reported in this thesis help to further expand understanding of the role of Varp within the cell.

Chapter Two

2 Materials and Methods

2.1 Molecular Biology

2.1.1 Sequence Alignments

Sequence alignments were performed using TCoffee accessed at https://www.ebi.ac.uk/Tools/msa/tcoffee/. Results were then formatted using the BoxShade server accessed at https://embnet.vital-it.ch/software/BOX_form.html. BoxShade colours conserved amino acids black, whilst amino acids that are conserved within one of the following groups are coloured grey: aromatic/cyclic and hydrophobic (FYW), hydrophobic (ILVM), positively charged (RK), negatively charged (DE), small (GA), polar (TS), or polar and amidic (NQ).

2.1.2 Cloning

All oligonucleotides for PCR were obtained from Sigma Aldrich (Haverhill, UK). PCR products were purified by Agarose gel, extracted using a Gel Extraction kit from QIAprep (Qiagen, Crawley, UK), the PCR product was then cut with restriction enzymes from New England Biosciences (Hitchin, UK) and ligated into the appropriate vector. All of the vectors used in this study were shown to be correct by sequence analysis by Beckman Coulter Genomics (Bishops Stortford, UK) using either standard primers, or sequence specific primers obtained from Sigma Aldrich (Haverhill, UK).

2.1.3 Polymerase Chain Reaction

DNA required for insertion into vectors was produced using the following standard techniques, including KOD PCR and BioTaq PCR. The solutions and cycle settings for each were as follows.

Standard KOD PCR solution:

5μl KOD Buffer

5μl Deoxynucelotidetriphosphates (dNTP) mix (2mM each)

2μL 10pmol/μL forward oligonucleotide

2μL 10pmol/μL reverse oligonucleotide

3μL $MgSO_4$ (25mM)

1μL template DNA (5ng/μL)

1μL KOD enzyme

31μL sterile H_2O

To screen ligation colonies and Yeast-2-Hybrid colonies for vectors with expected inserts, the following solutions, and colony screening protocol was used.

Colony screening PCR solution:

2μL BioTAQ buffer

0.5μL $MgCl_2$ (50mM)

0.3μL BioTAQ polymerase

2μL dNTPs (2mM each)

2μl forward primer (10μM)

2μl reverse primer (10μM)

Sterile H_2O to 20μL

Template DNA – touch colonies with toothpick, dip into aliquot, mix and remove toothpick.

Each solution was vortexed thoroughly and placed in a heated lid thermocycler using the below settings.

Standard KOD PCR:

1. 95°C for 2 minutes

2. 95°C for 20 seconds

3. 55°C for 10 seconds

4. 70°C for 15 seconds (10-25s (target dependent) per kb of DNA)

Repeat steps 2-4 x 30 cycles.

BioTAQ Colony PCR:

1. 95°C activation for 3 minutes,
2. 95°C denaturation for 30 seconds,
3. Annealing temperature 5°C lower than primer melting temperature (Tm),
4. 70°C extension (30-45s per kb of DNA),

Repeat 2-4 for 20 cycles.

2.1.4 Agarose Gel Electrophoresis

PCR products were run on a TBE agarose gel (90mM Tris Borate, 2mM EDTA) with 1/10,000 (v/v) SYBR Safe DNA Gel Stain (ThermoFischer Scientific, Loughborough, UK) for visualization. The percentage of agarose (Sigma Aldrich, Haverhill, UK) was altered to suit the size of the DNA product. The relative size of the PCR product was determined by comparison with DNA Hyperladder ™ 1Kb (Bioline, London, UK).

2.1.5 Purification

PCR products were purified by excising agarose gel fragments and isolating the DNA using the QIAquick Gel Extraction kit (Qiagen, Crawley, UK).

2.1.6 Restriction digest

PCR products and vectors were digested using the necessary restriction enzymes (New England BioLabs, Hitchin, UK) for 1 hour at 37°C in the appropriate buffer (as recommended by the manufacturer).

2.1.7 Ligation*s*

Using no more than a total of 200ng DNA per reaction, Roche Rapid DNA ligation kit (Sigma Aldrich, Haverhill, UK) was used to ligate insert and vector at a 1:3 ratio of vector to insert.

Purified insert DNA (e.g. 50ng of 500bp) was added to vector (e.g. 50ng of 1500bp) and made up to 10µL with DNA dilution buffer (supplied with Roche Rapid DNA Ligation Kit).

10µL Ligation buffer

1µL Ligase

Left for 5 minutes at room temperature and used for transformation immediately thereafter.

2.1.8 Chemically competent cells

Chemically competent DH5α *E.coli* cells were created to amplify plasmid DNA. A batch of DH5α cells were plated onto a 9cm LB plate and grown overnight at 37°C. A single colony was then used to inoculate 2mL LB medium, which was then incubated at 37°C with 210rpm shaking overnight. 1mL of the overnight culture was used to inoculate 100mL LB medium. The culture was then grown on at 37°C with 210rpm shaking until it reached an OD_{600} of 0.25-0.3. The culture was then chilled on ice for 15 minutes, following which the cells were centrifuged for 10 minutes at 3300x g at 4°C. The resulting pellet was then resuspended in 30-40mL of 0.1M $CaCl_2$ at 4°C. These cells were then kept on ice for 30 minutes after which they were centrifuged as before. The pellet was resuspended in 5mL in 0.1M $CaCl_2$ plus 15% glycerol at 4°C. The cells were then frozen on dry ice and subsequently stored at -80°C.

2.1.9 Chemical Transformations

For cloning purposes

2µl of ligation mixture was added to 50µL of a suspension of chemically competent DH5α *E.coli* cells on ice and incubated for 15 minutes. This mixture was then heat shocked at 42°C for 1 minute and recovered on ice for at least 1 minute before the addition of 250µL SOC (Super Optimal broth for Catabolite repression) media. These cultures were incubated for 1 hour at 37°C with 210rpm shaking. The cultures were pelleted by centrifugation (5000 rpm, 10 minutes). 150µL of supernatant was removed, and the pellet resuspended in the remaining 150µL of liquid. The resuspension was then plated on a pre-warmed (37°C) TYE agar plate containing the antibiotics appropriate for the plasmid and incubated overnight.

For protein expression purposes

Chemically competent BL21(DE3) PLysS *E.coli* cells for protein expression were produced as for DH5α. 1µL plasmid DNA was incubated with 50µL of a suspension of BL21(DE3) PLysS *E.coli* cells for 15 minutes on ice. The mixture was then heat shocked at 42°C for 1 minute and recovered on ice for at least 1 minute before the addition of 250µL SOC media. This culture was then incubated, pelleted and resuspended in the same way as for transformed DH5α cells. The 150µL resuspension was then plated onto prewarmed (37°C) 9cm TYE agar plates containing antibiotics appropriate for the plasmid, and 34µg Chloramphenicol and incubated overnight at 37°C.

2.1.10 Plasmid preparation

Plasmids from DH5α cells were isolated and purified using a QIAprep Spin Miniprep kit (Qiagen, Crawley, UK).

2.2 Protein Chemistry

2.2.1 SDS Polyacrylamide Gel Electrophoresis (SDS PAGE)

Protein samples were boiled in SDS (sodium dodecyl sulphate) sample buffer (125mM Tris-HCl pH6.8, 4% glycerol, 0.002% bromophenol blue, 8% w/v SDS, 5% β-mercaptoethanol) for 5 minutes. The samples were spun at 10,000 rpm for 1 minute to collect all the liquid, and run on the appropriate gel at a constant current of 45mA. Either 4-12% or 12% NuPAGE™ Bis-Tris Protein Gels were used to resolve proteins using MES buffer (2.5mM MES, 2.5mM Tris Base, 0.005% SDS, 0.05mM EDTA, pH7.3) and the XCell SureLock™ Mini-Cell electrophoresis system (ThermoFischer Scientific), unless otherwise specified. Protein bands were visualised using either InstantBlue™

(Expedeon Protein Solutions, CA, USA) or Coomassie Blue stain (1g Coomassie blue, 250mL methanol, 50mL acetic acid, 200mL H_2O). A solution of 10% methanol, 10% acetic acid was used where destaining was required.

2.2.2 Protein Expression

2.2.2.1 Expression of human Varp Znk2 in *E.coli*

Residues 692-746 of human Varp (one of two Zinc-binding motifs – referred to as Znk2) had previously been cloned by Ingmar Schafer [7] into a pGEX6P1 vector containing an N-terminal Glutathione-S-Transferase tag with a Prescission protease cleavage site (LQVLPE/GPLGS) between. This vector was used to transform BL21 PLysS *Escherichia coli* cells for protein expression. Single colonies resistant to both Ampicillin and Chloramphenicol were used to inoculate 10mLs of 2xTY medium with 100µg/mL Ampicillin and 34µg/mL Chloramphenicol.

Non-isotopically labelled protein was then prepared as follows. The 10mL starter cultures were used to inoculate 1L of 2xTY medium, grown to an OD_{600} between 0.8-1.2 and induced with 0.4mM IPTG, with the addition of 0.5mM $ZnCl_2$. The induced cultures were grown for 16 hours at 22°C with 210rpm shaking, before pelleting by centrifugation at 6000 rpm for 30 minutes at 4°C. The pellets were resuspended and stored in 20mM Tris, 200mM NaCl, pH7.4, 0.2mM βME, 0.2mM AEBSF or PMSF, 10 units DNase (Sigma Aldrich, Haverhill, UK) at -20°C.

For expression of isotopically labelled protein, after 5 hours growth, the 10mL starter cultures (in 2xTY) were gently pelleted at 4000rpm for 20 minutes. These pellets were gently resuspended in 10mL M9 minimal media (enriched only with $^{15}NH_4Cl$ (0.5g/L) or with both $^{15}NH_4Cl$ (0.5g/L) and $^{13}C_6$-glucose (2g/L)) and then added to a total volume of 1L of the same isotopically enriched M9 minimal media (recipe in Appendix 1). These 1L cultures were then grown at 37°C with 210rpm shaking to an OD_{600} of between 0.8-1.2, where the cultures were induced with 0.4mM IPTG. Upon induction, 0.5mM $ZnCl_2$ was added to the media. The induced cultures were grown for 18-20 hours at 28°C, before pelleting by centrifugation at 6000 rpm for 30mins at 4°C. The pellets were resuspended and stored in 20mM Tris, 200mM NaCl, pH7.4, 0.2mM βME, 0.2mM AEBSF or PMSF, 10 units DNAse at -20°C.

For both isotopically labelled and non-isotopically labelled protein purification the following applies.

Upon thawing, the cells were lysed by disruption (CF2 cabinet, Constant Systems, Daventry, UK) using a pressure of 30kPSI and then centrifuged at 30,000 rpm for 1 hour. The supernatant was then incubated with GST-Sepharose resin (GE Healthcare, Amersham, UK) for 1 hour at 4°C with gentle stirring, before being washed with 500mL of 20mM Tris, 100mM NaCl, pH7.4. The resin was then incubated overnight at 4°C with 0.5mg Prescission Protease. The cleaved protein was eluted from the beads under gravity flow in 5mL batches using the previous wash buffer, and after elution 0.5μM $ZnSO_4$ and 0.2mM βME were added to the elution fractions. The protein-containing fractions, as determined by SDS-PAGE electrophoresis, were pooled and loaded onto either a 16/26 or 10/300 Hiload S75 Superdex column (GE Healthcare, Amersham, UK) equilibrated with 20mM Tris, 200mM NaCl, pH7.4 with an appropriate flow rate for the column. The protein containing fractions were pooled and used for subsequent experiments. When required, the protein solution was concentrated using VivaSpin 3000mwco spin filters (Sartorius, Epsom, UK) that had been washed with 3-5x the concentrator volume of 20mM Tris, 200mM NaCl, pH7.4.

To exchange the buffer solution for other applications, a VivaSpin 3000mwco spin filter (Sartorius, Epsom, UK) was washed with 3-5x the concentrator volume of buffer. The concentrated protein solution was then diluted by $^{1}/_{3}$ with the desired end buffer and reduced in volume by $^{2}/_{3}$. This process was repeated 5 times so as to reduce the level of the original buffer to less than 0.5%.

2.2.2.2 Expression of human Varp Znk1 in *E.coli*

Residues 369-460 of human Varp (one of two Zinc-binding motifs – referred to as Znk1) had previously been cloned by Ingmar Schafer [7] into a pGEX6P1 vector containing an N-terminal Glutathione-S-Transferase tag with a Prescission protease cleavage site (LQVLPE/GPLGS) between. Recombinant Znk1 was prepared in an identical fashion to that described above for Znk2.

2.2.2.3 Expression of murine Vps29 in *E.coli*

Full length murine Vps29 was previously cloned by Matthew Seaman (CIMR, Cambridge, UK) into a pGEX4T2 vector containing an N-terminal Glutathione-S-Transferase tag with a Thrombin cleavage site (LVPR/GS) between. Murine Vps29 differs from human Vps29 by only 1 residue, with the final C-terminal residue being Serine in mVps29 and Proline in hVps29. This vector was used to transform BL21(DE3) PLysS *Escherichia coli* cells for protein expression. Single Ampicillin and Chloramphenicol resistant colonies were used to inoculate 10mL of 2xTY medium with 100μg/mL Ampicillin and 34μg/mL Chloramphenicol.

Non-isotopically labelled protein was produced as follows. The 10mL starter cultures were used to inoculate 1L of 2xTY which were then grown at 37°C, 210rpm shaking, to an OD_{600} of 0.8-1.2, where the cultures were induced with 0.4mM IPTG. These cultures were then grown at 18°C for 16 hours, before pelleting by centrifugation at 6000rpm for 30mins at 4°C. The pellets were resuspended and stored in 20mM Tris, 100mM NaCl, pH7.4, 0.2mM βME, 0.2mM AEBSF/PMSF, 10 units DNase at -20°C.

Isotopically labelled protein was produced as follows. The 10mL starter cultures (in 2xTY) were grown for 5 hours before being spun at 4500 rpm for 20 minutes. The resulting pellets were then resuspended in 10mLs M9 minimal media (enriched only with $^{15}NH_4Cl$ (0.5g/L) or with both $^{15}N_4Cl$ (0.5g/L) and $^{13}C_6$-glucose (2g/L)) and then added to a total volume of 1L of the same isotopically enriched minimal medium (recipe in Appendix 1). The 1L cultures were grown at 37°C, 210rpm shaking, to an OD_{600} of between 0.8-1.2, when the cultures were induced with 0.4mM IPTG. The induced cultures were grown at 25°C for 16 hours, before pelleting by centrifugation at 6000rpm for 30mins at 4°C and stored in 20mM Tris, 100mM NaCl, pH7.4, 0.2mM βME, 0.2mM AEBSF/PMSF, 10 units DNase at -20°C.

For both isotopically labelled and non-isotopically labelled protein purification the following applies.

Upon thawing, the cells were lysed by disruption (CF2 cabinet, Constant Systems) using a pressure of 30kPSI and then centrifuged at 30,000rpm for 1 hour at 4°C. The supernatant was then incubated with GST-Sepharose resin (GE Healthcare, Amersham, UK) for 1 hour at 4°C with gentle stirring, before being washed with 500mLs of 20mM

Tris, 100mM NaCl, pH7.4. The resin was then incubated overnight at room temperature with Thrombin (50U/mL resin). The protein was eluted in 10mL batches using the previous wash buffer with additional 2mM DTT. The protein-containing fractions, as determined by SDS-PAGE electrophoresis, were pooled and loaded onto a 16/60 or 26/60 Hiload S200 Superdex column equilibrated with 20mM Tris, 200mM NaCl, pH7.4. The protein containing fractions were pooled and used for subsequent experiments. When needed, the protein fractions were concentrated using VivaSpin 10000mwco (Sartorius, Epsom, UK) spin filters, which had previously been washed 3-5X concentrator volume with 20mM Tris, 200mM NaCl, pH7.4.

To exchange the buffer solution for other applications, a VivaSpin 10000mwco spin filter (Sartorius, Epsom, UK) was washed with 3-5x the concentrator volume of buffer. The concentrated protein solution was then diluted by $^1/_3$ with the desired end buffer and reduced in volume by $^2/_3$. This process was repeated 5 times so as to reduce the level of the original buffer to less than 0.5%.

2.3 NMR Spectroscopy

All data were acquired on DRX 500MHz, DRX 600MHz, DMX 600MHz, Bruker Avance 2 700MHz, or Bruker Avance I 800 MHz spectrometers, each equipped with a triple resonance (^{1}H, ^{15}N, ^{13}C) 5mm cryoprobe at 298K, unless otherwise stated. ^{1}H chemical shifts were calibrated using sodium-3,3,3-trimethylsilylpropionate (TSP) as an external reference. ^{15}N and ^{13}C chemical shifts were indirectly referenced against ^{1}H, using the ratio of gyromagnetic ratios [226].

Spectra were processed using TOPSPIN 3.2 or 3.5 (Bruker GmbH, Karlsruhe) and analysed in CCPN Analysis 2.4.2 [227]. For a list of all NMR experiments used in this thesis, see Appendix Two – NMR .

2.3.1 NMR Sample Preparation

All NMR samples were prepared in either 20mM Tris, 200mM NaCl, pH7.0, 1mM DTT in H_2O or 20mM deuterated Tris, 200mM NaCl, pH7.0, 1mM deuterated DTT in either D_2O or H_2O unless otherwise stated. Samples were sealed with a septum in a standard NMR tube, drawing out dissolved air using a vacuum pump and replacing with Argon/Nitrogen in several rounds.

2.3.2 Chemical Shift Perturbations

Differences were calculated for the backbone amide ^{15}N, ^{1}H chemical shifts of residues in ^{15}N labelled Vps29 alone or in complex with unlabelled Varp Znk2; these chemical shifts were monitored using ^{15}N HSQC experiments. This was repeated for labelled Znk2 in complex with natural abundance Vps29. For each residue, the individual shift differences for ^{15}N and ^{1}H were combined into a single CSP using the formula below, with an α value of 0.2:

$$CSP = \sqrt{((\Delta\delta\,^{1}H)^{2})} + ((\alpha \cdot \Delta\delta\,^{15}N)^{2})$$

2.3.3 Structure Determination

To determine the structure of Vps29 binding Varp Znk2, a hybrid X-ray crystallographic/NMR approach was used, similar to that described by Eustermann et al [224]. The conformation of the majority of the Vps29 component was restrained to a template conformation adapted from a previously published X-ray structure (see below), while the conformations of the Varp peptide and interfacial residues of Vps29 (selected on the basis of preliminary structures) were allowed to evolve under a combination of intermolecular and intra-peptide NOE-derived distance restraints, as well as limited NOE-derived restraints for the interfacial region of Vps29 and J-coupling-derived χ^{1} dihedral angle restraints for the interfacial region of the Varp peptide. Details of the NMR constraints used and their analysis (including distance calibration in the case of NOE-based constraints) are discussed in Sections 5.1 and 5.3. All of the NOESY datasets used for structure calculations were acquired using pulse sequences modified to ensure equal RF heating in each case, e.g. for ^{13}C experiments a period of ^{15}N decoupling equal in length to the acquisition period was applied at the beginning of the inter-scan delay, and for ^{15}N experiments an equivalent period of ^{13}C decoupling was similarly applied.

A previously published X-ray crystallographic structure of Vps29 (pdb 2R17) was used to create a template for Vps29 in the Vps29:Varp Znk2 structure [137]. Residues 1-181 of Vps29 of chain A from pdb 2R17 were used, with all seleno-Met residues changed to Met and 11-12, 40-41 and 91-92 peptide bonds set to a *cis* conformation. Details of the rationale for selecting 2R17 as a template and for selecting the cis conformation of these peptide bonds are discussed in Section 5.2.

The protein co-ordinates in the template were adapted as follows: Hydrogen atoms were added according to standard geometries, then all atoms except for the guanidinium protons of all arginines and all atoms of residues 40 and 41 were fixed and the structures then subjected to Powell energy minimisation (1,000 steps), Langevin dynamics at 1,000K (20,000 steps), increase of the Van der Waals force constant in 2,000 step cycles, cooling to 300K in 2,000 step cycles and final Powell minimisation (1,000 steps); this stage of the calculation allowed adaptation to the *cis* conformation of the 40-41 peptide bond as well as resolving atom definition issues for the arginines. All atoms except backbone amide N, C' and O were then released and the structures again subjected to Powell minimisation (100 steps). Only very small movements of the protein backbone occurred during this minimisation; for residues 1-181 the backbone co-ordinate shift (rmsd for N, Cα, C') was 0.194Å.

To calculate an ensemble of models for the complex, 50 starting structures were first created by randomising the backbone ϕ and ψ angles of Varp 687-746 and placing the peptide at a random distance (Gaussian distribution around 150Å) and orientation relative to a copy of the template structure of Vps29. These starting structures were then subjected to a two-stage simulated annealing protocol to generate an ensemble of conformers consistent with NOE and dihedral angle constraints, all whilst applying the Ramachandran database potential of mean force [228] with a force constant of 1 kcal.mol^{-1}. Since the XPLOR-NIH calculations utilized r^{-6} summation for all groups of equivalent protons and non-stereospecifically assigned pro-chiral groups, and since no stereoassignments were made, (and the assignment-swapping protocol within XPLOR-NIH for deriving stereoassignments indirectly during the structure calculation itself was not applied), all constraints involving protons within such groups were converted to group constraints (by using wildcards such as HB*). All lower bounds were set to zero

[229]. Stage 1 of the two-stage protocol comprised Powell energy minimisation (500 steps), Langevin dynamics at 1,000K (5,000 steps), increase of the van der Waals force constant and tilting of the NOE potential function asymptote in 1,000 step cycles, switching to a square well function then cooling to 300K in 500 step cycles and final Powell minimisation (500 steps). The force constants used for both the distance and the dihedral angle restraints were 50kcalmol^{-1}. In stage 2 zinc ions were first defined and placed at the geometric average of the co-ordinates of the four zinc-binding cysteinyl sulphur atoms, and all necessary bond and angle terms added to the force field, following which structures were subjected to Powell energy minimisation (500 steps), Langevin dynamics at 1,000K while progressively increasing the torsion angle force constant in 1,000 and 2,000 step cycles, switching to a square-well NOE function then cooling to 300K in 250 step cycles and final Powell energy minimisation (1000 steps).

Throughout these calculations, strong non-crystallographic symmetry (NCS) constraints were used to maintain the internal structure of the Vps29 component. To achieve this the Vps29 structure was duplicated, the co-ordinates of one copy shifted by 500Å then rigidly fixed and unfixed copies so as to maintain the Vps29 structure in the evolving co-ordinates of the unfixed copy. Once preliminary rounds of calculation had established likely regions where contacts between Vps29 and Varp might occur, the force constant for NCS terms applied to backbone atoms in just these regions of Vps29 were much reduced and those applied to sidechain atoms in the same regions reduced further still; the NCS constraints used in the final rounds of calculations are shown below:

Atoms constrained	Force constant (kcal.mol^{-1})
Residues 7-19 and 60-148, all atoms	100.0
Residues 1-6, 20-59 and 149-181, N, Cα and C' atoms; All remaining atoms of residues 1, 3-6, 20-24, 26, 28-29, 31-59, 149-151, 153, 155-160, 162, 164, 166-171, 173 and 175-181	2.0
All remaining atoms of residues 2, 25, 27, 30, 152, 154, 161, 163, 165, 172 and 174	0.1

Finally, the remaining atoms of the Vps29 and N- and C- terminal tails (residues -9-0 and Ser 182) were added in a separate annealing protocol. Initially all atoms of the full structure were placed at fully randomised positions within a 200Å cube, then for those residues included in the previous stages of the calculations (i.e. residues 1-181 of Vps29 and all residues of Varp Znk2) the randomised co-ordinates were replaced by the previously calculated values and rigidly fixed in place. The structures were then subjected to 1,000 cycles of Powell minimisation, followed by 5,000 steps of Langevin dynamics at 500K, increase of the van der Waals force constant in 1,000 step cycles, cooling to 300K in 1,000 step cycles and final Powell minimisation (1,000 steps). This protocol for adding the tails and linker was repeated independently 5 times for each input structure using a different randomisation seed each time, and the structure with the lowest value of Etotal retained. The 25 models with lowest total energy were accepted to form the final ensemble. The program CLUSTERPOSE [230] was used to calculate the mean rmsd of ensembles to their mean structures, and structures were visualised using the program PYMOL ([231]). Ensembles were superposed using the co-ordinates of their respective average structures; the average structures themselves are not shown.

2.4 Amino Acid Analysis

A sample of Varp Znk2 was sent to the Protein and Nucleic Acid Chemistry facility (PNAC facility (Dept. of Chemistry, Cambridge, UK)) for amino acid analysis. This involves hydrolyzing the protein and using an ion exchange analyser to detect individual amino acids to obtain accurate protein concentration data. The sample as sent had a concentration of 389µM (45.9µg in 20µL, or 2295µg/mL) as determined by spectroscopy at A_{280} using an extinction coefficient calculated using the ProtParam program (ExPASY).

2.5 SEC-MALS Analysis

SEC-MALS (Size Exclusion – Multiangle Light Scattering) analysis was performed using an Agilent HPLC system (Agilent, California, USA) coupled to a Wyatt Helios II 18 angle light scattering instrument, a Wyatt Optilab rEX online refractive index detector and a Wyatt QELS module (Wyatt, Haverhill, UK).

100µL of each sample was applied consecutively to a S75 10/30 column (GE Healthcare, Amersham, UK), which was set to run at 0.5ml/min via the Agilent HPLC system (Agilent, California, USA). The eluent of this run was then passed through the light scattering and refractive index detectors. The resulting data was compared to a BSA standard.

2.6 GST Pulldown Technique

To capture GST-Znk2, 40µL of Glutathione (GSH) Sepharose Bead solution (GE Healthcare, Amersham, UK) in 20mM Tris, 500mM NaCl, 0.5µM $ZnCl_2$, 0.1mM TCEP was incubated with either 0µg, 20µg, 50µg or 100µg GST-Znk2 in the presence or absence of 200µg Vps29 for 30 minutes at 4°C. These samples were spun for 10 minutes at 13,200rpm, washed with 1mL buffer and spun again (repeated 3x). Then, 40µL SDS gel loading buffer (see SDS PAGE protocol) was added and the samples boiled for 5 minutes. These samples were then run on a 4-20% Tris-Glycine gel (180V for 45mins) in Tris-Glycine running buffer (25mM Tris base, 192mM glycine, 0.1% w/v SDS, pH 8.3).

Western blotting was then used to detect Vps29 in the samples. The SDS PAGE gel was placed in a blotting tank with transfer buffer (25mM Tris, 1.5mM glycine, 20% methanol and 0.1% SDS) and run at 100V, 400mA for 1.5 hours to transfer the protein onto nitrocellulose membrane. To check for protein transfer, the membrane was stained with Ponceau S (2% w/v Ponceau S, 30% w/v trichloroacetic acid) and destained with water.

The nitrocellulose membrane was then incubated in blocking buffer (100mM Tris, 15mM NaCl, 0.1% Tween, 5% milk powder) for 30 minutes. A 1/1000 dilution of Goat pAb Anti-Vps29 was prepared in blocking buffer and incubated with the nitrocellulose membrane for 1 hour at room temperature with rocking incubation. The membrane was then washed in wash buffer (100mM Tris, 15mM NaCl, 0.1% Tween) for 5 minutes, repeated 3x with fresh buffer. A dilution of 1/10,000 Horse Anti-Goat Ab conjugated to HorseRadish Peroxidase (HRP) was prepared in blocking buffer and incubated as before, followed by washing as before. 5mL of each peroxide solution and luminol enhancer reagent was used to develop the blot with a 1 minute incubation period, after which the excess reagents were wicked away. The blot was then developed with X-ray film after 10s, 30s and 60s exposures.

2.7 Surface Plasmon Resonance

All SPR experiments were performed on a Biacore T200 instrument with a CM5 Sensor Chip and the 'GST Capture Kit' (GE Healthcare, Amersham, UK) unless otherwise stated.

To prepare the chip flow cell surface for ligand immobilization via amine coupling [232], a mixture of EDC (N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride) and NHS (N-hydroxysuccinimide) was injected over the cell at a rate of 10µL/min for 7 minutes. The ligand in this case, Anti-GST Ab (GE Healthcare), was thus immobilized onto the previously prepared cell by flowing the antibody solution in 10mM sodium acetate, pH5.0, at a concentration of 30µg/mL with a flow of 10µL/min for 5 minutes, at 25°C, across the reactive surface. Any remaining reactive esters were then deactivated using ethanolamine injected over the surface for 7 minutes at 10µl/min. This process resulted in an antibody immobilization of 1000-1500 response units (RU).

Response units are a measure of the change in refractive index of the sensor surface over time. The RU measurement does have a positive correlation with the amount of protein attached to the sensor chip, however, it cannot be used as an accurate quantification of the absolute amount of protein attached to the chip. Older user manuals describe a value of 1000RU being equal to 1ng of protein per mm^2, however this advice has since been withdrawn because individual proteins have different effects on the refractive index value and therefore the amount cannot be accurately stated for all proteins.

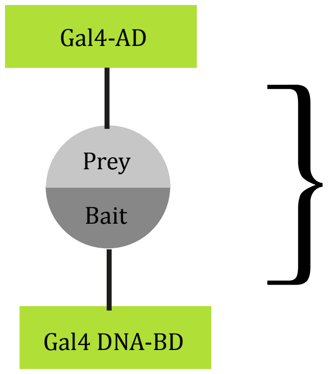
In order to ensure that GST binding was fully reversible during subsequent measurements, high affinity (non-reversible) GST binding sites present on the immobilized antibody were first blocked by injecting 5μg/mL recombinant GST over the surface at 10μL/min for 3 minutes. Subsequently, 10mM glycine-HCl at pH2.2 was injected at 10μL/min for 3 minutes to dissociate GST from the lower affinity (reversible) GST binding sites on the antibody. Figures to illustrate this process appear in Chapter 3, Section 3.3. This provides a surface on which to reversibly bind GST-fusion proteins to the immobilized antibody surface.

Once the surface has been prepared as described, either GST alone (to provide a reference cell) or GST-tagged fusion proteins were flowed over the surface. To ensure reproducibility between consecutive runs, as well as a stable point from which to do kinetic analysis of data, the response of the two cells was carefully monitored, with any runs that resulted in a response substantially higher or lower than average being discounted. Either GST-Znk2 or GST-Znk1 was immobilized by flowing a 10μg/mL solution of GST-tagged protein over the immobilised antibody for 3 minutes. With the surface prepared in this way, varying concentrations of either wtVps29 or mutant Vps29 analyte were then injected over the sensor surface at a flow rate of 10μL/min. All analyte-ligand interactions were measured using 20mM Tris, 200mM NaCl, 0.1mM TCEP, 0.5μM $ZnCl_2$, pH 7.4 as a buffer unless otherwise stated.

To regenerate the GST binding surface after each binding cycle, the GST tagged protein bound to its ligand was removed by injecting 10mM glycine-HCl, at pH2.2, for 3 minutes at a flow rate of 10μL/min.

2.8 Yeast-2-Hybrid Screen

The Mate and Plate™ Yeast-2-Hybrid (Y2H) system [233] employs a strategy whereby the transcription of four reporter genes, under the control of Gal4 transcription activator responsive promoters, can be used to detect a protein:protein interaction. In this case, to identify novel interactions, the bait protein is expressed as a fusion protein with the Gal4 DNA-binding domain (DNA-BD), and a variety of prey proteins are expressed as fusion proteins with the Gal4 Activation domain (AD). Upon bait:prey interaction, the proximity of the two Gal4 domains is sufficient to promote the transcription of the reporter genes, and allow yeast colony growth on highly selective media.



Reporter Genes:

HIS3 - Confers ability to biosynthesise Histidine

ADE2 - Confers ability to biosynthesise Adenine

MEL1 - Confers ability to express and secrete α-galactosidase

AURI-C - Confers resistance to usually toxic drug Aureobasidin A

Figure 19: Four reporter genes are activated upon bait:prey protein interaction. This is due to the Gal4 DNA-binding domain (Gal4-DNA-BD) and Gal4 Activation domain (Gal4-AD) coming into close enough proximity to promote the transcription of these four genes.

Reporter genes with a Gal4-responsive promoter are present in the Y2HGold *S.cerevisiae* strain into which the bait plasmid is transformed: **HIS3**, **ADE2**, **AURI-C** and **MEL1**.

HIS3 confers the ability to biosynthesise histidine,

ADE2 confers the ability to biosynthesise adenine,

MEL1 confers the ability to produce and secrete α-galactosidase, which in the presence of the substrate X-α-galactosidase will produce a blue colony colour,

AURI-C confers resistance to the drug Aureobasidin A by producing a mutant form of inositiol phosphoryl ceramide synthetase that is not susceptible to adverse effects of the usually toxic chemical.

The pGBKT7 plasmid into which the bait DNA is inserted contains the DNA-BD of Gal4, a Kanamycin resistance gene, and the **TRP1** gene which confers the ability to biosynthesize tryptophan and can therefore be selected for by growth on Single DropOut Tryptophan media (SDO (-Trp)). The Y187 *S.cerevisiae* strain into which the library of prey DNA is transformed contains only the **MEL1** reporter gene. The pGADT7 plasmid into which the library of prey DNA is inserted contains the Gal4-AD, an Ampicillin resistance gene, and the **LEU2** gene which confers the ability to biosynthesize Leucine and can therefore be selected for by growth on Single DropOut Leucine media (SDO (-Leu)). The multitude of genes allows a range of selection media to ensure that the detection of positive interactions is robust, see Table 2 for a list of media and selection specificity.

Yeast 2 Hybrid Selective Media	Selection Criteria
SDO (-Trp)	Selective for Bait plasmid pGBKT7
SDO (-Leu)	Selective for Prey plasmid pGADT7
DDO (-Leu, -Trp)	Selective for pGBKT7/pGADT7 positive diploid
QDO (-Leu, -Trp, -His, -Ade)	Selective for pGBKT7/pGADT7 positive diploids activating transcription of HIS3 and ADE2 reporter genes
QDO/X/A (-Leu, -Trp, -His, -Ade, +X-α-galactosidase, +Aureobasidin A)	Selective for pGBKT7/pGADT7 positive diploids activating transcription of HIS3, ADE2, and AURI-C reporter genes. Colonies are coloured blue if activating MEL1 reporter gene.

Table 2: Selective media regime for Yeast 2 Hybrid screening.

The dropout media were prepared as below.

Base dropout media:

3.45g Yeast nitrogen base without amino acids (Formedium™, Norfolk, UK)
15g Agar,
25mL 40% w/v Glucose,
475mL H_2O.

Add Kaiser formulation amino acid dropout supplement (Formedium™, Norfolk, UK) as appropriate (per 500mLs):

811mg -Leu
963mg -Trp
773mg -Leu -Trp
726mg -Ade -His -Leu -Trp

If required, with the addition of:

40μg/mL X-α-Gal (Clontech)
200ng/mL Aureobasidin A (Clontech).

2.8.1 Cloning

Residues 1-136 of human Varp were cloned into a pGBKT7 yeast vector in frame with the Gal4 DNA-binding domain, using Nde1/Sal restriction sites. The plasmid was cloned as described in Section 2.1, with the Standard KOD PCR solution altered with the addition of 2μL DMSO, and a reduction of H_2O volume to 29μL.

2.8.2 Yeast Transformation protocol

The desired yeast strain was streaked onto 9cm YPD (Yeast Petone Digest) plates and grown at 30°C until colonies reached 2-3mm in diameter. A single colony was used to inoculate 10mLs of YPD media, this culture was then vortexed briefly to distribute the cells in the solution, before being incubated at 30°C with 230rpm shaking for 16 hours.

The resulting culture was diluted to achieve an OD_{600} of 0.2-0.3, following which 60mLs was placed into a 250mL flask and grown at 30°C with 230rpm shaking for 3 hours. Once this culture reached a cell density of OD_{600} 0.4-0.6, 4mLs was spun at 700xg for 5 minutes at room temperature. The resulting pellet was resuspended in 20mL sterile H_2O before being spun again as above. The pellet was then resuspended in 1.5mL 1.1x TE/LiAc mix (11mM Tris HCl, 1.1mM EDTA, 110mM lithium acetate), before being spun at 13,000xg for 15s. From this, the pellet was resuspended in 0.5mL 1.1x TE/LiAc mix before being held on ice for 10 minutes.

Carrier DNA (Merck KGaA, Darmstadt, Germany) at 10mg/mL was heated at 100°C for 5 minutes, cooled to 4°C on ice, before being heated once again at 100°C for 5 minutes. Immediately following this, 50ng of the carrier DNA was added to the previously prepared cells and mixed gently before the addition of 0.5mL PEG/LiAc mix (40% w/v PEG, 10mM Tris HCl, 1mM EDTA, 10mM LiAC).

100ng of the plasmid DNA of interest was added to 5µL sterile H_2O. To this the cell suspension was added and the whole was gently mixed by inversion before incubating at 30°C for 30 minutes, with gentle mixing every 10 minutes. After this, 20µL of DMSO was added and mixed in by inversion, before heat shock at 42°C for 15 minutes, with gentle mixing every 5 minutes.

The cells were then pelleted by centrifugation at 13,000xg for 15 seconds at room temperature, the pellet resuspended in 1mL YPD media and incubated for 1 hour at room temperature with rocking. The cells were then centrifuged again as before and then resuspended in 50µL sterile H_2O before being plated on plates with the appropriate selective media, e.g -Trp for pGBKT7 containing cells, and sealed with tape. These were grown for 3 days at 30°C.

2.8.3 Autoactivation/Toxicity Test

To test for bait (GAL4-BD - hVarp1-136) autoactivation of the reporter genes, 100µL of pGBKT7-hVarp1-136 was transformed into the Y2HGold strain of

Saccharomyces cerevisiae. The resulting pGBKT7-hVarp1-136 transformed Y2HGold cells were then plated on SDO (-Trp), SDO (-Trp)/X and SDO (-Trp)/X/A, whilst the positive control diploid (Y2HGold strain [pGBKT7-53] mated with Y187 strain [pGADT7-T]) was plated on DDO (-Leu, -Trp)/X/A.

2.8.4 Yeast-2-Hybrid Screen

2.8.4.1 Mating and growth of interaction positive colonies

The prey library, Universal Human (Normalised) Mate and Plate™ Library (Clontech (now Takara), France), was supplied cloned into pGADT7 vectors in the Y187 yeast strain.

To create mated diploid cells, the bait strain (Y2HGold [pGBKT7-hVarp1-136]) was plated onto a Yeast extract Peptone Digest (YPD) medium and single colonies were used to inoculate an overnight culture. At an OD_{600} of 0.8, the culture was pelleted and resuspended to a cell density of >1x10^8 cells/mL in 4.5mL SDO (-Trp). This was used to inoculate 45mL YPD, 50µg/mL Kanamycin with 1mL (5x10^7 cells) of Universal Human (Normalised) Mate and Plate™ Library in Y187. These cells were incubated at 30°C for 24 hours with 40 rpm shaking. Then, the culture was pelleted at 1000xg for 10 minutes and resuspended in 7.5mLs of 0.5xYPD, 50µg/mL Kanamycin. A small amount of the mated culture was plated on 9cm SDO (-Trp), SDO (-Leu) and DDO (-Leu, -Trp) media at 100µL dilutions of 1/10, 1/100, 1/1000 and 1/10,000. Colonies growing on the 9cm SDO (-Trp), SDO (-Leu) and DDO (-Leu, -Trp) media were used to calculate the mating efficiency of the process and also how many clones were screened during the process.

122 colonies on 1/1000 plate –DDO plate.

7.5mL x 10 x 122 x 1000 = 9.15x10^6 colonies screened.

425 colonies on 1/1000 SDO –Leu plate.

(diploids / limiting partner - bait) cfu/mL x 100 = mating efficiency.

(1.22x10^5 / 4.25x10^5) cfu/mL x 100 = 29% mating efficiency.

The remaining 5mLs of cell culture was then spread across 20 15cm QDO plates (-Leu, -Trp, -Histidine (His), -Adenine (Ade). Growth on these plates indicates a positive result of two of four reporter genes activated. Any colonies that grew were then re-streaked onto QDO/X/A plates as a further round of selection with all four reporter genes needing to be activated. Any blue colonies were selected for PCR amplification and sequencing, and re-streaked onto QDO/X/A plate media to continue colony growth.

2.8.5 Colony screening

A small part of a positive single colony from the QDO/X/A screening plate was mixed into 50µL sterile H_2O, 2µL of this solution was dropped onto a QDO/X/A plate to ensure continuity of the colony and 2µL was used for the PCR screening.

BIOTAQ PCR Screen:

2µL target cells/H_2O mix,

2µL 10x NH_4 reaction buffer (Bioline, London, UK),

0.2µL (1 unit) BIOTAQ enzyme (Bioline, London, UK),

0.4µL 10mM Each dNTPs,

0.8µL 20mM $MgCl_2$,

1.0µL 10µM pGADT7 FOR primer,

1.0µL 10µM pGADT7 REV primer.

12.6µL H_2O.

PCR cycle settings:

96°C denaturing,

60°C annealing,

72°C extension 3.5 minutes,

35 cycles.

To determine if the PCR reaction amplified any bait DNA, the solution was run on a 2% agarose gel. The gel was made by dissolving 2g of agarose (Sigma Aldrich) in 100mL TBE buffer (90mM Tris base, 90mM Boric acid, 2mM EDTA, pH8.3), heating for 1 minute, then, when cool to the touch, 8µL of SYBR® Safe DNA Gel stain was mixed into the solution; this mixture was then poured into a DNA gel mould and left to cool. 15µL PCR samples were mixed with 6x DNA loading buffer (Invitrogen, UK) and loaded onto the gel with Hyperladder DNA marker (Bioline, London, UK) as a standard.

Any colonies that did not produce a band on an agarose gel could not be sequenced. If there was a band on the agarose gel, the band was excised and the DNA purified from the gel using the QiaQuick® Gel extraction kit (Qiagen, Germany) and sent for sequencing by primer extension (Genewiz, formerly Beckman Coulter, New Jersey, USA) with the pGADT7 FOR primer. The resulting sequences were BlastX (NCBI) searched using the 'translate protein from nucleotide' query at https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastx&PAGE_TYPE=BlastSearch [234, 235]. The matches were then recorded using the protein name, ID reference, the residues of the prey matched, the quality of the sequencing data and also any mutations between the library sequence and the database sequence were noted.

Chapter Three

3 Biochemical Studies on Varp and Vps29-Varp interaction

To characterise the interaction between Varp and Vps29, both a GST-pulldown assay and Surface Plasmon Resonance experiments were used. These studies resulted in the determination of the K_D for the interaction of either Znk1 or Znk2 of Varp with Vps29.

3.1 Pull-down assay demonstrates that GST-Znk2 binds to Vps29

Although the information it generates is qualitative information, a pull-down assay is a useful initial experiment to directly determine if two proteins interact with one another. To determine if Znk2 binds wtVps29, various concentrations of GST-Znk2 were bound to GST sepharose beads. These solutions were then incubated with 200μg wtVps29 to determine if there was a binding interaction, before running on an SDS PAGE gel and subsequent transfer to nitrocellulose paper for Western blotting.

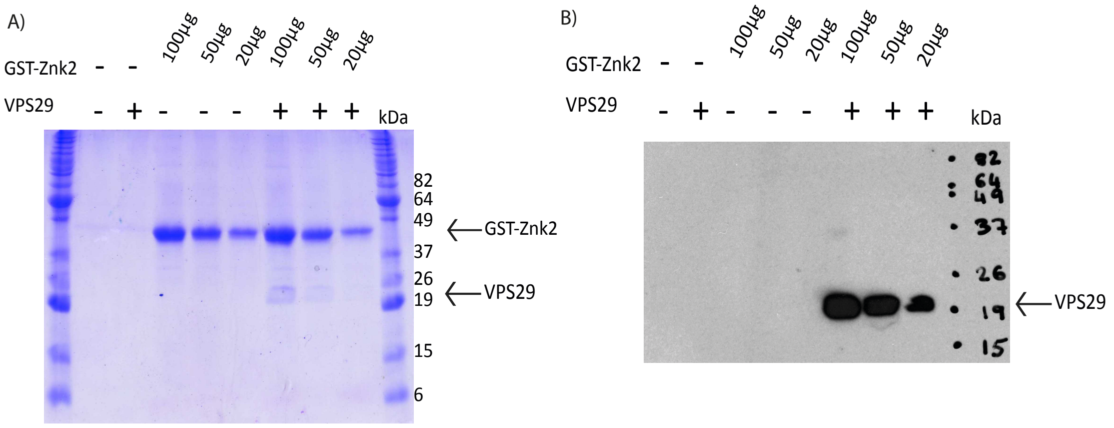


Figure 20: A GST-Pulldown experiment. A) 14-20% Tris-Glycine SDS-Page gel (run at 180V for 45mins) of GST sepharose beads with or without GST-Znk2 and with or without Vps29. Ladder in first and final lanes is Benchmark Prestained protein ladder. B) 60s exposure X-ray film of Western blot (using Goat Anti-Vps29 Ab, and Horse Anti-Goat Ab) developed from nitrocellulose transfer of the SDS-PAGE gel in A).

The Western blot in Figure 20B shows that wtVps29 does not bind to GST sepharose beads in the absence of GST-Znk2 (Lane 2), therefore the presence of wtVps29 in the other samples (Lanes 6,7,8) can be directly attributed to the presence of GST-Znk2 in those samples. This finding is further reinforced by the increased amount of Vps29 present when the concentration of GST-Znk2 is higher, 100μg vs 50μg vs 20μg (See Figure 20) showing a clear concentration dependence of the interaction.

3.2 Surface Plasmon Resonance shows that GST-Znk2 and GST-Znk1 bind Vps29 with a K_D in the micromolar range

As a more quantitative method than a pull-down, Surface Plasmon Resonance (SPR) is a valuable additional technique for looking at many aspects of binding events involving proteins. A ligand is bound to the suitably prepared surface of a sensor chip, and then the binding partner of interest (analyte) is flowed over the ligand bound surface at a range of different concentrations under a fixed temperature. The binding of the analyte to the ligand changes the refractive index of the sensor surface (measured in response units (RU)), which when plotted over time, produces a sensorgram. A variety of information can be deduced from the sensorgram, including both binding kinetics and the affinity of an interaction and by carrying out experiments at different temperatures, thermodynamic information can be obtained. In terms of kinetics, the Biacore T200 system can measure rates of association (ka) of slower than $3x10^9$ M^{-1} s^{-1} and rates of dissociation (kd) slower than $1x10^{-5}$ s^{-1}.

Amine coupling of Znk2 directly to a CM5 chip was attempted, to try to exclude any issues that may result from GST potentially binding Vps29. However, the amine coupled Znk2 showed no interaction with wtVps29. It was hypothesized that because Znk2 is small, and likely to be mostly unstructured, that the amine coupling destroyed any structure present in the molecule, thus preventing binding. It could also be the case that after amine coupling the Znk2 molecule was too close to the chip surface and hence this could obscure the Vps29 binding site.

As a result of failing to directly couple Znk2 to the sensor surface in a usable form, it was decided to use a GST Capture kit from GE Healthcare to capture GST, GST-Znk2 or GST-Znk1 onto a CM5 chip sensor surface using an anti-GST antibody which itself first needed to be amine coupled to the chip. The pulldown results described in Section 3.1 have already demonstrated that GST-Znk2 could bind wtVps29. In the SPR experiments, GST alone was used in a reference cell to provide control data that could be subtracted from the experimental GST-fusion protein results and so provide a true representation of fusion protein-Vps29 binding, if any.

3.2.1 CM5 Chip Surface Preparation

The CM5 chip surface is formed of a carboxymethylated dextran layer to which biomolecules can be covalently attached via a number of means, including amine coupling. Amine coupling involves activating the carboxyl groups on the surface of the chip with a mixture of 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC) and N-hydroxysuccinimide (NHS) which results in the formation of reactive succinimide esters. The esters then react with primary amine groups in a molecule of interest to covalently couple the molecule to the chip.

A GST capture kit from GE Healthcare was used to immobilize the supplied Anti-GST antibody onto the CM5 sensor chip (GE Healthcare), see Figure 21 A/C. Because the polyclonal anti-GST antibody has a small fraction of high-affinity sites that can be difficult to regenerate, these high-affinity sites should be occupied prior to starting an experiment. This is achieved by slowly injecting recombinant GST over the prepared Anti-GST surface for 7 minutes (See Chapter 2, and Figure 21 B/D).

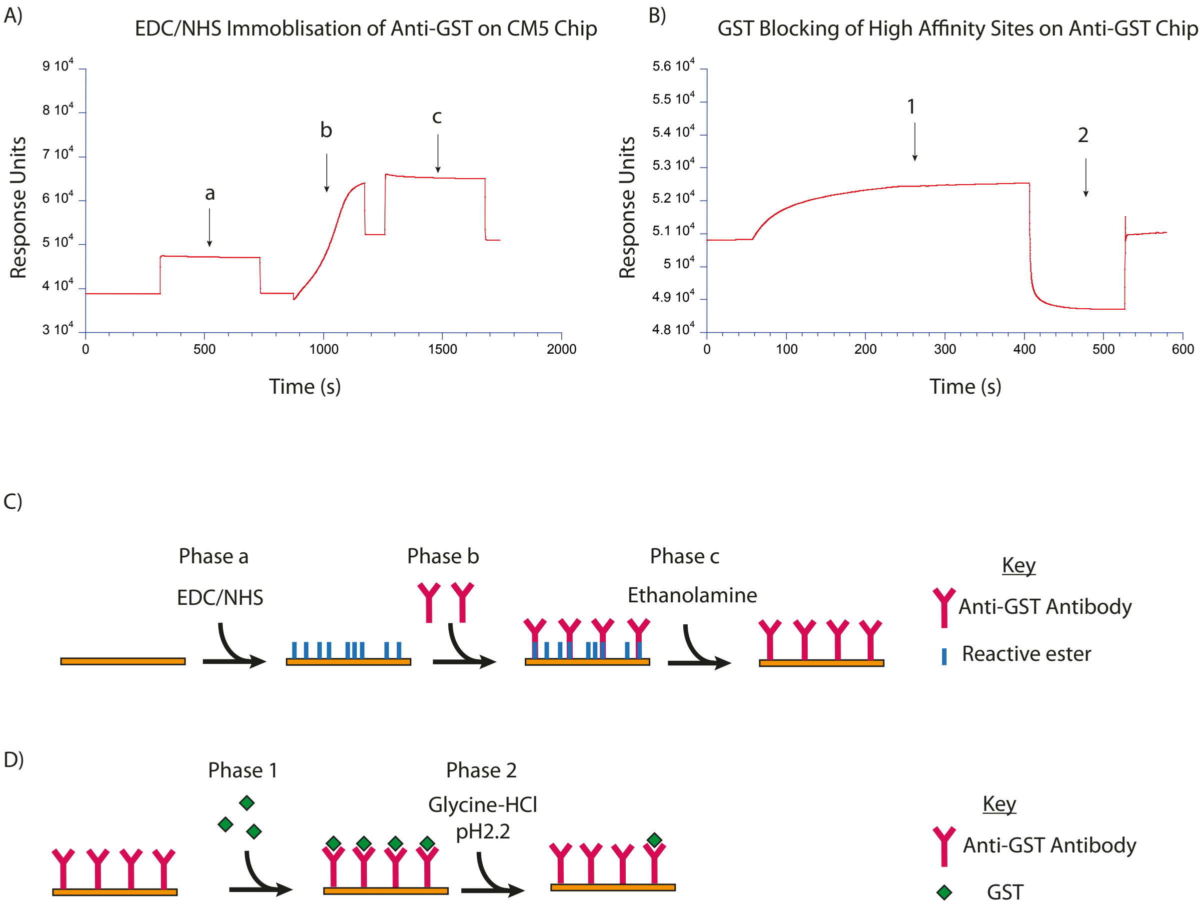


Figure 21: Generating a functional CM5 sensor chip for SPR experiments. A) A sensorgram depicting the stages of anti-GST antibody immobilization onto a CM5 sensor chip via amine coupling. Phase a represents injection of EDC/NHS to activate the sensor surface creating reactive esters, Phase b shows the gradual association of Anti-GST Ab binding to the activated surface, and Phase c shows the deactivation of the surface, removing reactive esters by ethanolamine treatment. The increase in response unit from T=0s to T=1800s represents the antibody remaining bound to the chip surface. B) A sensorgram depicting the blocking of high affinity GST sites on a previously prepared Anti-GST Ab surface. Phase 1 demonstrates the binding of GST to the antibody surface whilst Phase 2 represents the regeneration of the antibody surface using glycine-HCl, pH2.2 to remove any GST that is not occupying a high-affinity site. C) A diagram illustrating the various phases as demonstrated by the sensorgram in A). D) A diagram to visually represent the various phases of the sensorgram in B).

3.2.2 Experimental Assay Setup

Once both the reference and active sensor surfaces were prepared as described, GST, GST-Znk2 or GST-Znk1 were injected over the appropriate surface to bind to the immobilized antibody, in 20mM Tris, 200mM NaCl, 0.1mM TCEP, 0.5μM $ZnCl_2$, pH7.4 buffer.

To test the developed protocol for the assay, both GST and GST-Znk2 were injected as 10mg/ml solutions for 180s. The relative response unit (RU) measurement of the binding of these ligands to the antibody surface was recorded. As discussed in Chapter 2, any experimental runs in which the binding of the ligand deviated too far from the average were discounted. Figure 22B shows a plot on which the 'capture level' of GST-Znk2 is shown for each cycle, it is clear from this plot that the ligand capture level (as indicated by the response unit measurement) does not vary by more than 20RU throughout the experiment, so it can safely be assumed that the level of GST-Znk2 bound to the chip remains approximately constant. To prepare the 'analyte' solutions for injection after the GST-ligand has been captured, Vps29 was expressed and purified by GST-sepharose binding, thrombin cleavage (all as described in Chapter 2, Section 2.2.2.3) with subsequent gel filtration using a size exclusion column (S75, 26/60) into 20mM Tris, 200mM NaCl, 0.1mM TCEP, 0.5μM $ZnCl_2$, pH7.4 and concentrated to a stock concentration of 2mM prior to injection over the sensor surface.

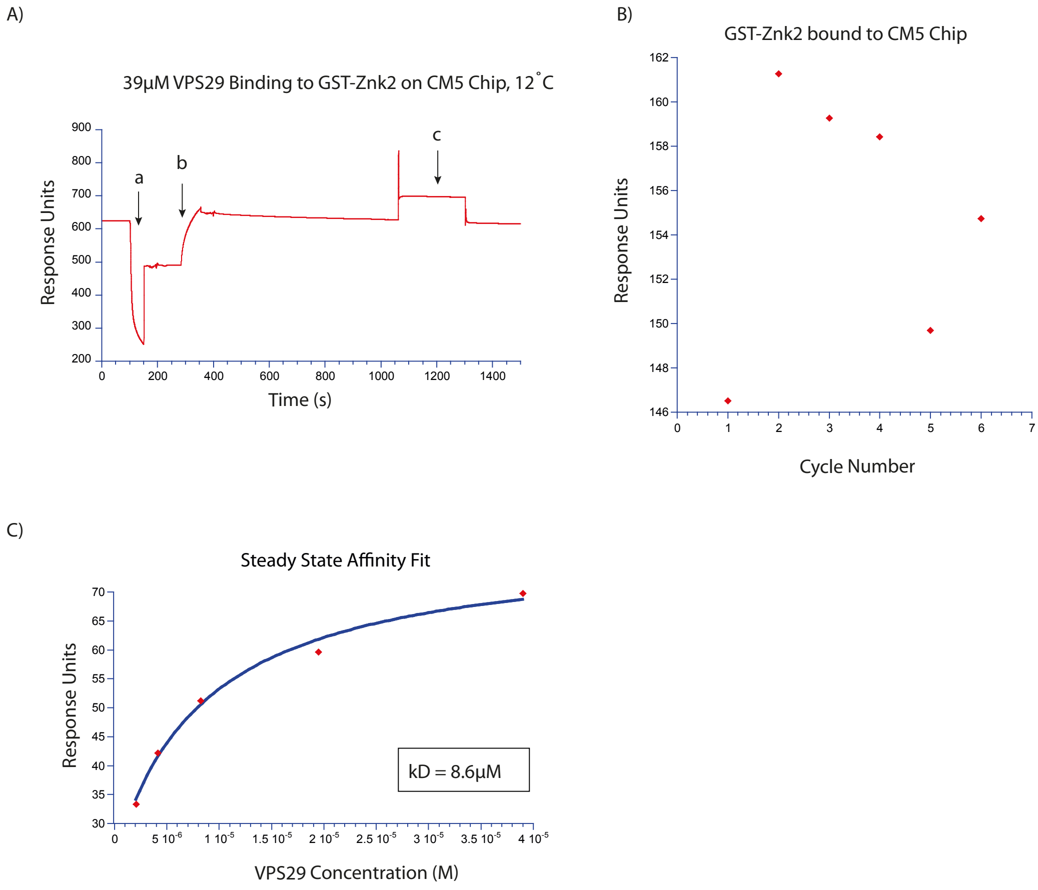


Figure 22: An SPR experiment showing Vps29 binding to GST-Znk2 at 12°C. A) A sensorgram of an experiment performed on a pre-prepared immobilized antibody surface at 12°C. This sensorgram results from the reference cell being subtracted from the active cell to correct for any non-specific binding events, and buffer effects. Phase x demonstrates the binding of GST to the reference cell whilst Phase y demonstrates the binding of GST-Znk2 to the active cell. Phase z is the injection of the analyte (in this case 39μM or 0.8mg/ml Vps29). The spike at the beginning of the injection is the result of the slight delay of the flow of analyte between the first cell (reference) and second cell (experimental). At the end of the analyte injection phase (Phase z), it is clear that the analyte immediately dissociates from the GST-Znk2, without a regeneration injection. B) A plot of GST-Znk2 capture versus the cycle number. It should be noted that in this and subsequent capture level plots the Y-axis range does not begin at zero, and is different in each case to correspond to the data range. C) A steady state affinity fit calculated from the response unit change upon injection of Vps29 at concentrations of 2, 4.1, 8.25, 19, 39μM. The affinity fit, and therefore the K_D value, was calculated using the Biacore T200 Evaluation software program, Version 2.0 (GE Healthcare, Amersham, UK).

The sensorgram in Figure 22A shows the phases of an individual experiment (one cycle), comprising GST/GST-fusion protein binding to the surface, analyte (in this case 39μM or 0.8mg/ml Vps29) injection into the flow cell and the dissociation of the analyte at the end of the injection. The regeneration injection subsequently removes not only the analyte from the chip, but also the GST-Znk2, leaving just the Anti-GST Ab on the chip so that the surface is ready for reuse in a subsequent 'cycle'; for a diagram depicting this see Figure 23.

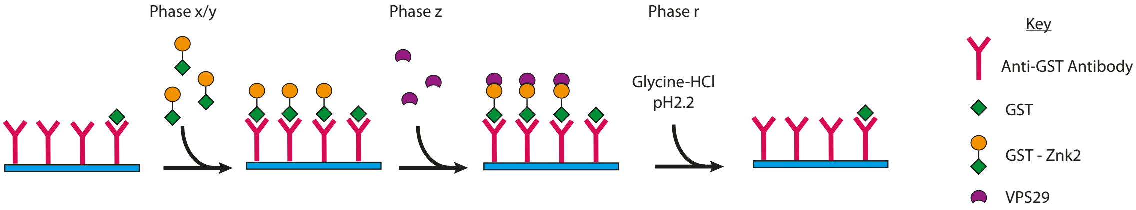


Figure 23: A diagram visualizing the various phases of the experiment shown by the sensorgram in Figure 22A. Phase x describes GST binding, and Phase y describes GST fusion protein binding to the antibody surface. Phase z demonstrates the injection of analyte protein over the surface and binding to GST/GST-fusion protein. Phase r (not shown in Figure 22) shows the regeneration injection that re-instates the original condition of the antibody surface.

Therefore, each cycle involves binding GST/GST-fusion protein back onto the chip. This is not ideal as the Anti-GST antibody will not continue to bind GST/GST-fusion protein to exactly the same extent (i.e to an identical RU value) indefinitely, thus introducing a potential error source. As mentioned previously, the level of GST-fusion protein binding to the Anti-GST antibody surface was recorded, and any experiment with large deviations in GST-fusion protein capture values were excluded from analysis. All analyte-ligand interactions were performed in 20mM Tris, 200mM NaCl, 0.1mM TCEP, 0.5μM $ZnCl_2$, pH 7.4 unless otherwise stated.

The analyte binding phase (phase z) in the sensorgram in Figure 22A depicts fast kinetics, with the binding reaching steady state rapidly (fast association), and the analyte dissociating immediately upon the cessation of the analyte injection. This, and the relatively small size of the proteins involved, makes it difficult to determine the binding kinetics of this interaction, and means that the affinity of the interaction can only be determined by steady state analysis, when the interaction is at equilibrium. To produce a steady state binding curve from the sensorgrams, a range of analyte concentrations was tested over the course of an experiment, with the flow rate and time being held constant for each. For the affinity fit in Figure 22C, 240s injections of Vps29 at concentrations of 2, 4.1, 8.25, 19, 39μM were used.

3.2.3 Results of Varp Znk2 and Vps29 binding Assay

The Biacore T200 Evaluation software, version 2.0 (GE Healthcare, Amersham, UK) was used to analyse the results using the steady state affinity fit calculation to determine the K_D of the interaction. When the data shown in Figure 22 were analysed in this way, the K_D value found for the binding of GST-Znk2 by Vps29 was 8.6μM, at 12°C. Biacore recommend that values ranging up to at least 2x K_D should be used in steady state affinity experiments to determine an accurate value, and in this case the analyte concentrations used to determine the K_D represent 0.233x, 0.478x, 0.962x, 2.216x, 4.549x the calculated K_D value.

Subsequently, the same experiment was repeated at 25°C using 8 Vps29 concentration points (0, 1, 2, 4, 8, 16, 24, 32, 48μM) to produce more data points for a more accurate affinity fit. Figure 24A shows a sensorgram displaying the overlaid, reference-corrected, analyte injection phase (phase z) for all the tested concentrations, including two blank cycles. The calculated K_D for the Vps29:Varp Znk2 interaction at 25°C, from the steady state affinity fit in Figure 24B, is 4.7μM, which is somewhat lower than the 8.6μM calculated at 12°C. This could be as a result of the increased experiment temperature, or the fact that GST-Znk2 binding to the chip decreased by more than 20RU over the experiment lifetime (as demonstrated in Figure 24C), thus meaning the response generated by injection of higher Vps29 concentrations is artificially lower.

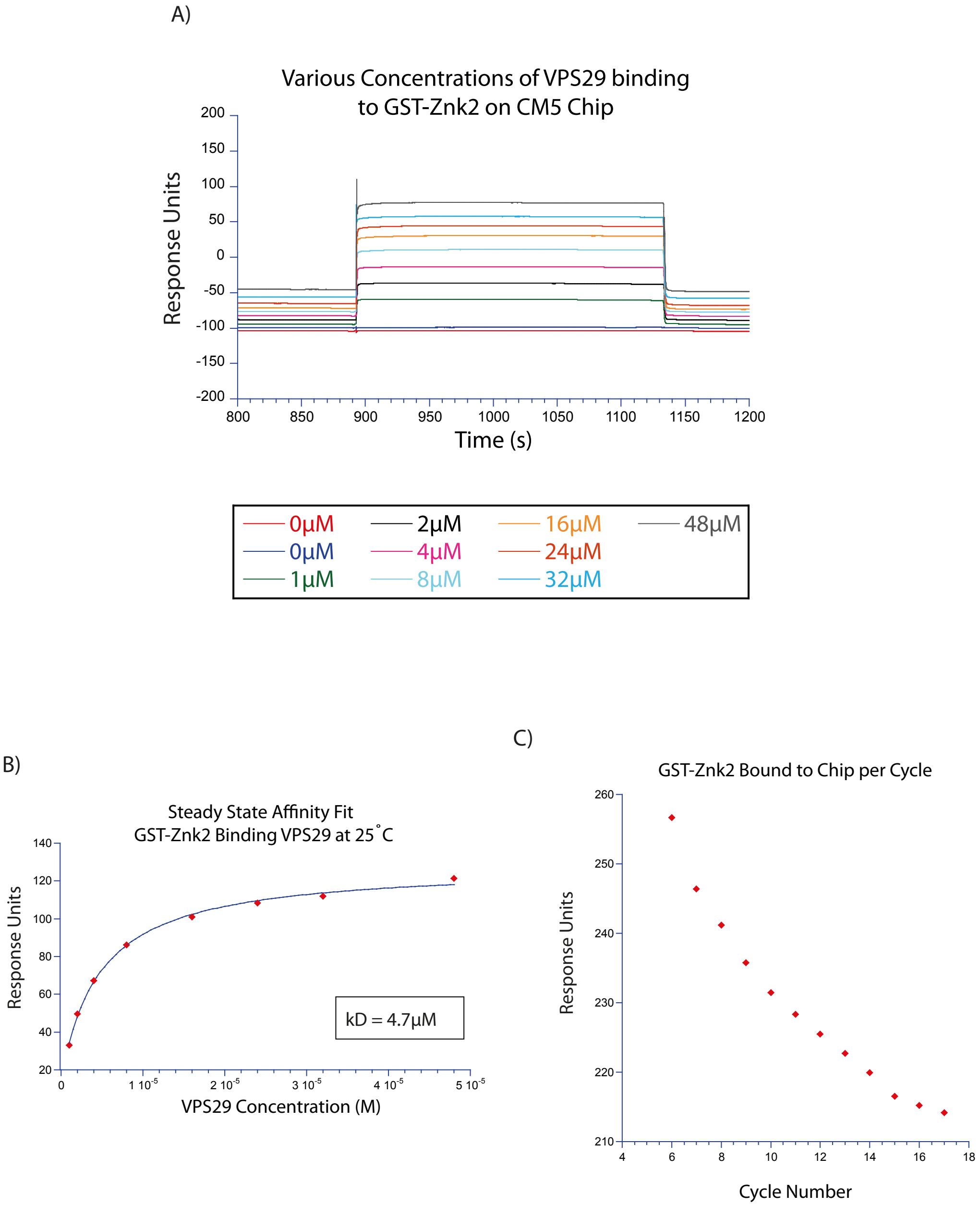


Figure 24: An SPR experiment showing Vps29 binding Znk2 at 25°C. An overlaid, reference-corrected, sensorgram displaying the analyte binding injection phase (phase z) of all cycles. Colour co-ordinated to Vps29 concentration. B) A steady state affinity fit plot of response unit versus Vps29 concentration. C) A capture level plot of GST-Znk2 bound to the active cell sensor surface.

In some cases, thermodynamic experiments can be useful to gather more detail about the kinetics of a binding interaction. Increasing or decreasing the temperature of the system can, in some instances, alter the rate of association and dissociation of the analyte to measurable levels. Previous experiments at 12°C and 25°C (sensorgrams in Figure 22 and Figure 24) both show association and dissociation rates that are too fast to measure by the instrument, as previously mentioned. In an attempt to determine kinetic parameters for the binding interaction, repeated experiments were carried out at temperatures of 12°C, 25°C and 37°C. In this case, 240s injections of Vps29 concentrations of 0, 1, 2, 4, 8, 16,24, 32, 48μM were flowed over a GST reference cell and an active, GST-Znk2 bound cell, prepared as before. However, as shown by the sensorgram in Figure 25A, the association and dissociation rate constants were consistently too fast to measure. The Biacore T200 system is not able to lower the temperature below 12°C, as the equipment can only hold at a minimum of -15°C from the external environmental temperature, and therefore kinetic analysis was not possible for the Vps29:Varp Znk2 interaction.

From Figure 25 it is apparent that there are slight fluctuations in the apparent K_D with temperature change, however these K_D values are within a small range. The K_D values calculated for both instances of Vps29:Varp Znk2 interaction at 12°C, 8.6μM for the initial experiment (Figure 22) vs 8.4μM for the thermodynamic experiment (Figure 25) agree very closely. This suggests there is good reproducibility in this case. However, the two instances of Vps29:Varp Znk2 interacting at 25°C result in more variation between the calculated K_D values. The initial experiment conducted at 25°C and displayed in Figure 24, resulted in a calculated K_D of 4.7μM, whereas the repeat 25°C experiment in Figure 25C resulted in a calculated K_D of 7.3μM. In the second instance, the variation in GST-Znk2 bound to the chip throughout the experiment was much less over the relevant cycles (<20RU deviation, see Figure 25E, cycle numbers 18-28) than for the previous experiment where the deviation was ~50RU. It is therefore likely that small changes in calculated K_D can be attributed to fluctuating GST-Znk2 levels bound to the sensor surface. However, both the values are within the same order of magnitude, and agree well with other affinities measured for the WT Vps29: WT Znk2 interaction.

Over the course of 30-40 cycles of ligand-analyte binding and surface regeneration during the experiment, the antibody surface appears to lose its capacity to bind GST. The capture level plot in Figure 25E clearly demonstrates that the level of GST-Znk2 attaching to the chip decreases over time. In this case, the problem only affects the cycles for the 37°C experiment, however the resultant K_D is still within an order of magnitude as the other results, and is therefore unlikely to be much affected. This would become more of an issue if the experiment on the chip required a greater number of cycles, making reproducibility an issue, as less protein on the chip would result in less binding of analyte and therefore would more significantly impact the K_D calculation. The Biacore Evaluation T200 software (version 2.0 (GE Healthcare, Amersham, UK)), does not allow any correction for the fluctuations of ligand concentrations and so any runs resulting in abnormal ligand capture that may drastically impact the K_D must be excluded, as previously discussed.

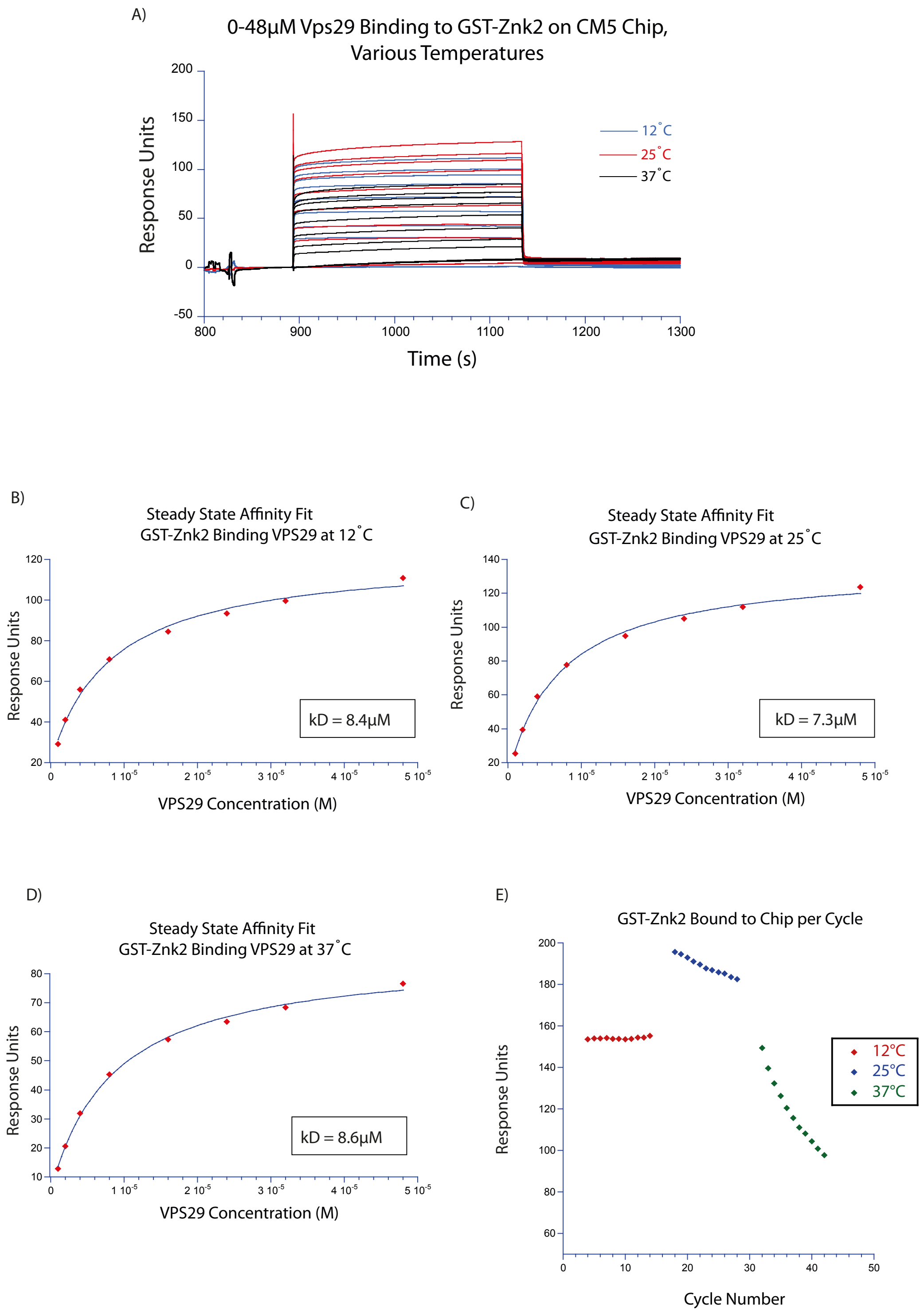


Figure 25: An SPR experiment showing Vps29 binding Znk2 at various temperatures. Overlaid reference–corrected sensorgrams, displaying only the analyte injection (phase z), of all cycles of 12°C, 25°C and 37°C experiments. B) Steady state affinity fit for 12°C experiment. C) Steady state affinity fit for 25°C

experiment. D) Steady state affinity fit for 37°C experiment. E) Capture level plot of GST-Znk2 binding over the cycles. Cycles 4-14 represent 12°C, Cycles 18 -28 represent 25°C, and Cycles 46-56 represent 37°C.

Following the 37°C experiment, a 12°C duplicate was run on the same chip to assess the impact of the falling GST-Znk2 levels. In the capture level plot in Figure 26 it is clear that by the time cycle 46 has been reached, the chip is no longer binding the same amount of GST-Znk2. The calculated K_D for the 12°C duplicate experiment was 16μM, compared to 8.4μM and 8.6μM on previous runs. Although this does not greatly affect the overall calculated K_D, as the K_D values are all within the in the same order of magnitude, the impact of fluctuating GST-Znk2 levels was mitigated in further experiments by monitoring the capture levels of the GST proteins throughout the experiment. This also required the preparation of a new Anti-GST Ab surface chip once the GST/ GST-fusion protein binding capacity dropped below 100 RU on any individual chip.

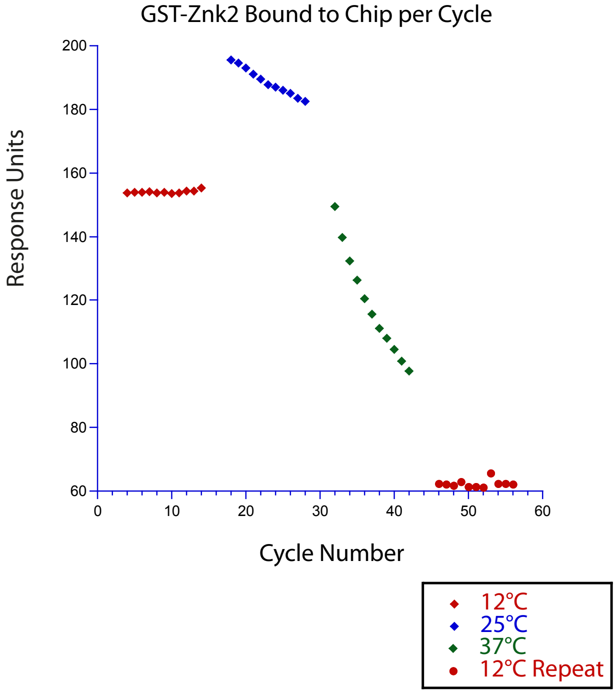


Figure 26: A capture level plot describing the relative amount of GST-Znk2 bound to the sensor chip per cycle.

3.2.4 Results of Varp Znk1 and Vps29 binding assay

The first zinc knuckle in Varp (residues 346-460) has been reported to bind Vps29 with a lower affinity than Znk2, with K_D values of 13µM and 5µM respectively [7]. To attempt to verify this, GST-Znk1 was attached to the chip via the anti-GST Ab as for GST-Znk2, and Vps29 binding experiments conducted as before. Each Vps29 concentration injection was repeated twice to control for the effect of fluctuating GST protein capture levels. As shown in Figure 27, GST-Znk1 binds Vps29 with an apparent K_D of 4.8µM. This suggests a tighter interaction than for Znk2, in contrast to the formerly reported result [7]. However, the difference between the calculated K_D values for Znk1 and Znk2 is very small and could simply be the result of experimental uncertainties with this technique. The GST-Znk1 capture level plot in Figure 27D shows that GST-Znk1 binding onto the chip was reproducible throughout the experiment with a deviation of <20RU. A table of all the calculated K_D values for the interaction of wtVps29 with both GST-Znk1 and GST-Znk2 (at all temperatures) is shown in Table 3.

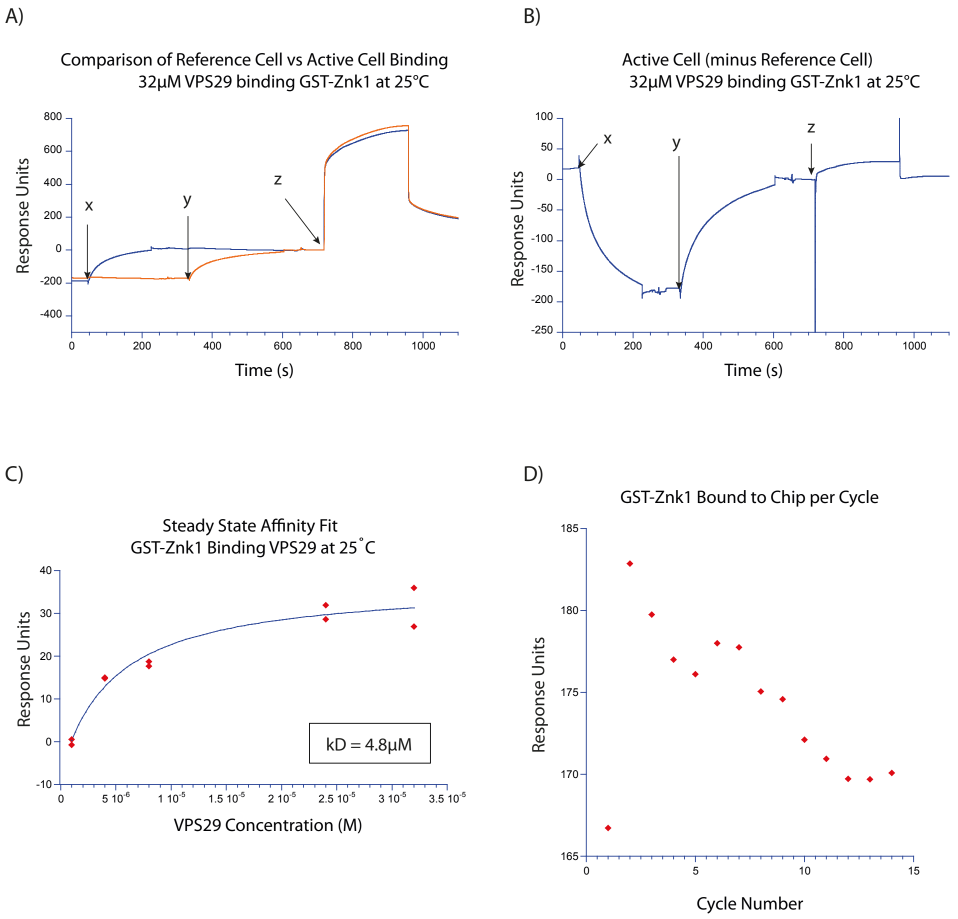


Figure 27: An SPR experiment showing Vps29 binding Znk1 at 25°C. A) Sensorgram showing both reference GST cell (blue) and experimental GST-Znk1 cell (red). Phase x describes the injection of GST, and GST binding onto the reference cell, Phase y shows the injection of GST-Znk1 into the active cell, and the binding of GST-Znk1 to the surface. B) Reference corrected sensorgram: the active cell minus the reference cell. The spike at the beginning of the analyte injection phase (phase z) is as a result of the time delay caused by the analyte injection reaching the reference cell slightly before the active cell. Likewise, the spike at the end of phase z is as a result of the time delay caused by the cessation of the analyte injection reaching the reference cell slightly before the active cell. C) Affinity fit, and K_D calculation for Vps29 binding to GST-Znk1. D) Capture level plot for all the cycles in the experiment.

Analyte	Ligand	Temperature	Average kD
wtVPS29	GST-Znk2	12°C	8.5μM (+/- 0.1)
wtVPS29	GST-Znk2	25°C	6.0μM (+/- 1.8)
wtVPS29	GST-Znk2	37°C	8.6μM
wtVPS29	GST-Znk1	25°C	4.8μM

Table 3: A table of the calculated K_D values for all SPR experiments with wtVps29.

3.3 EDTA treatment reduces the interaction affinity between Varp Znk2 and Vps29.

To determine if the binding interaction between Vps29 and Varp Znk2 was affected by the removal of the zinc from Varp Znk2, further SPR experiments were conducted in the presence of 10mM EDTA. A 10mM EDTA-containing buffer was used to remove the zinc from GST-Znk2, and wtVps29 was then injected at concentrations of 0, 1, 2, 4, 6, 8, 12, 16, 24, 32 and 48μM in duplicate to monitor the response level. The capture level of GST-Znk2 onto the chip fluctuated greatly at the beginning of the experiment (see Figure 28C), with cycles 4-9 being higher than the subsequent cycles. For this reason, those cycles (comprising one repeat each of 0, 1, 2, 4, and 8μM) were excluded from the affinity calculation in Figure 28B.

The calculated K_D of 86μM suggests that the binding affinity of Vps29 and GST-Znk2 is much reduced in the presence of EDTA, most likely because the zinc has been removed from GST-Znk2. As in the case of the mutant Vps29 L152E experiment, an analyte concentration of 2xK_D was not reached during this experiment, and therefore the actual affinity of this interaction is likely to be even lower than the calculated value under these conditions.

The intention of using SPR techniques was to provide a quantitive measure of the differences in binding between both the two zinc knuckles of Varp and between wtVps29 and the two mutants. However, issues with fluctuating GST capture levels, especially in prolonged experiments requiring multiple cycles, and reaching acceptable analyte concentrations for accurate determination of K_D (most particularly for low affinity interactions) has meant that the data is not sufficiently reproducible to allow robust quantitative conclusions to be drawn, and it will be necessary for some of these experiments to be repeated in the future to generate data of publication quality.

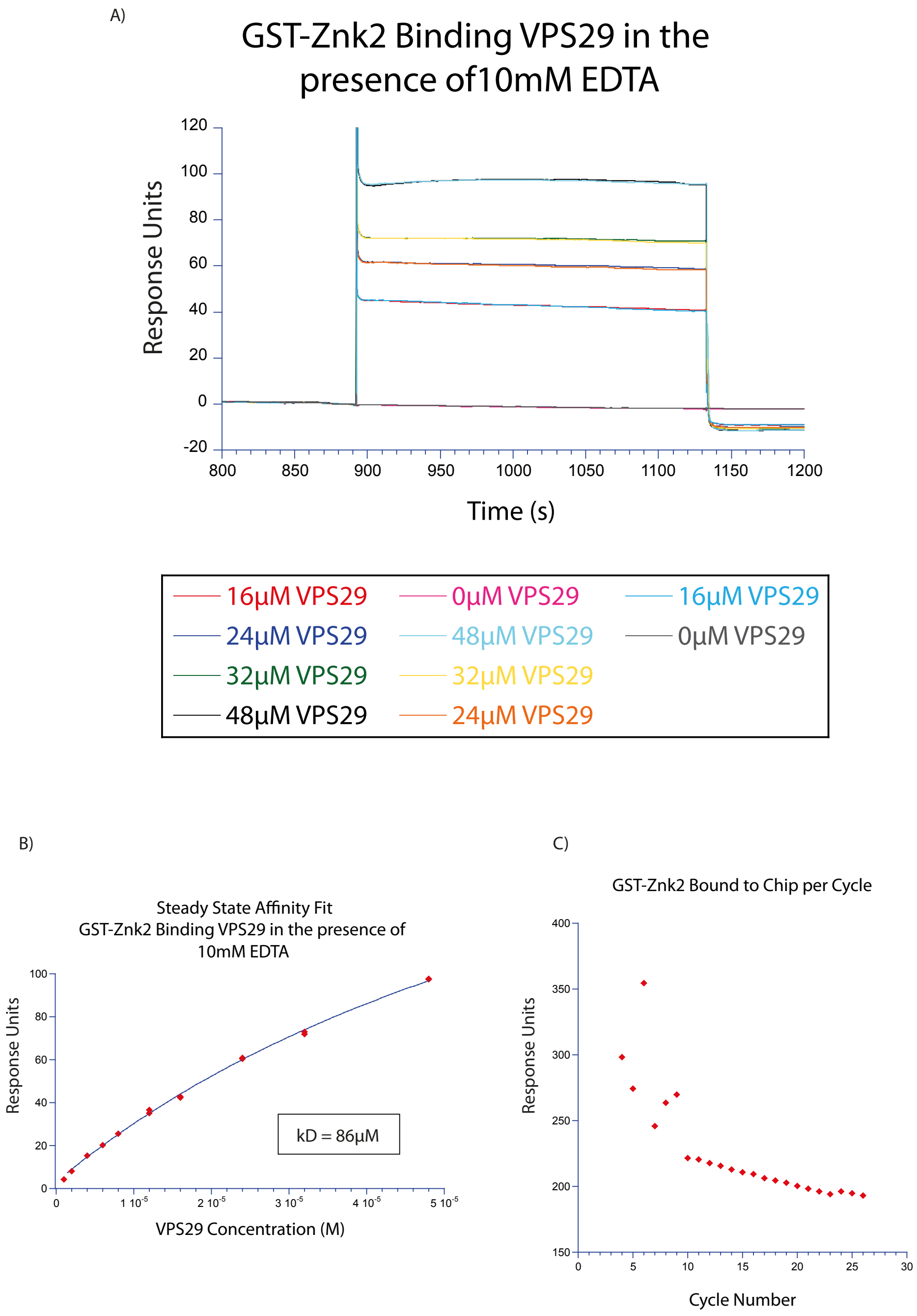


Figure 28: An SPR experiment showing Vps29 binding to Znk2 in the presence of 10mM EDTA. An overlay of reference subtracted sensorgrams depicting the binding of Vps29 to GST-Znk2 in the presence of 10mM EDTA. (Only cycles 12-19 + 26 shown as examples). B) A steady state affinity fit, using cycles 10-25. C) A capture level plot showing GST-Znk2 bound to the chip for all experimental cycles.

Chapter Four

4 NMR Studies on Varp and Vps29; sample preparation, signal assignment and chemical shift footprinting.

NMR is an extremely valuable tool for analysing the atomic structures of proteins. It allows an in-depth assessment of the structures and also the dynamic properties of proteins and protein-ligand complexes.

For the Vps29:Varp Znk2 complex, determining both the structure and the chemical shift perturbations (relative to the free components) firstly required assignment of all the resonances in each component in both the free state and following complex formation. The resonance assignment of residues in the knuckle region of Varp Znk2 in the complex was greatly hindered by significant line-broadening effects occurring mostly for residues at the interaction interface. This, and the fact that the short sequence of Znk2 contains 5 Prolines (which have no amide protons), made sequential assignment particularly difficult. In addition, the protein samples created for the experiments were prone to degradation and were therefore often quite short lived, despite having been sealed under Nitrogen or Argon. Due to this, multiple samples were needed to gather the required data. This meant that reproducibility was sometimes an issue, particularly because slight differences in the ratio of the two components between samples results in small differences in the chemical shifts and line widths between spectra, most noticeably in the interface region. The occurrence of such differences must therefore be taken into account during interpretation.

4.1 Characterisation of hVarp Znk2.

hVarp Znk2 (residues 692-746) was recombinantly expressed from a pGEX6P1 vector in *E.coli*., as described in Chapter 2, Section 2.2.2.1. The plasmid contains an N-terminal GST moiety with a Prescission Protease cleavage site (LEVLFQ/GP) preceding the recombinant Znk2. An additional C-terminal tryptophan was added with the aim of facilitating protein concentration determination by A280 during purification. This recombinant construct is shown in Figure 29.

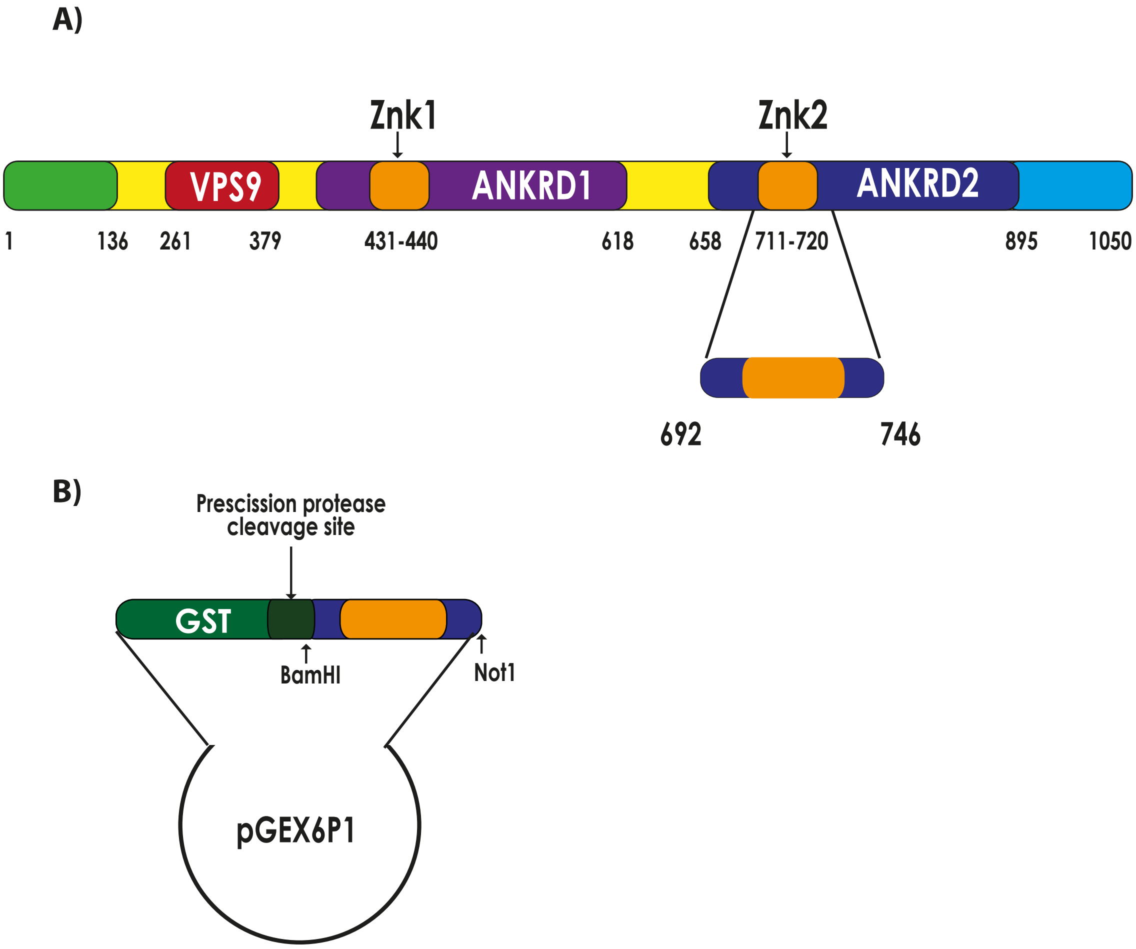


Figure 29: Two diagrams of recombinant Varp protein construct preparation. A) Schematic of hVarp depicting the different domains. The domains are colour coded as follows: green depicts the N-terminal region, red the Vps9 domain, purple and dark blue are the ankyrin repeat domains and the Zinc binding domains are highlighted in orange. Also depicted is the sequence of Varp used to create the Znk2 protein for experimentation in this thesis. B) A representation of the pGEX6P1 construct used to express GST-Znk2 in *E.coli*.

Following cleavage of GST, the resulting 58aa protein (GPLGS - hVarp Znk2 (692-746) - W) was predicted to have a molecular weight of 6048Da and a theoretical extinction coefficient of 5500 M^{-1} cm^{-1} (assuming all Cysteine residues are reduced) using the ProtParam program from the ExPASy server at http://www.expasy.org [236].

4.1.1 Amino acid analysis confirms theoretical hVarp Znk2 extinction coefficient for accurate quantification.

A sample of hVarp Znk2 was sent for Amino Acid analysis (University of Cambridge, PNAC facility, Cambridge UK) to determine the correct extinction coefficient for this protein to enable accurate protein concentration determination using the absorbance of light at 280nm. Using the ProtParam derived extinction coefficient of 5500 M^{-1} cm^{-1}, the sample was submitted at a calculated concentration of 389μM (45.9μg in 20μl, which is 2295μg/ml) for assessment. The average concentration determined for the sample by amino-acid analysis was 2378μg/ml +/- 62.29μg. This means that the Protparam derived extinction coefficient is sufficiently accurate to provide a reliable protein concentration reading for hVarp Znk2 using an A280 measurement.

It is important to note that there are a few amino acids not included in the analysis (see Table 4). This mostly relates to the limits of the technique, for example acid hydrolysis converts Asparagine to Aspartic acid and Glutamine to Glutamic acid, so values are combined into a result for Asx and Glx respectively. Also, Tryptophan cannot be accurately measured as the process of acid hydrolysis causes the amino acid to rapidly degrade. Another issue occurs due to Proline co-eluting with Cysteine during the chromatography step. This makes Proline difficult to quantitate, especially in the case of hVarp Znk2 as Cysteine and Proline residues comprise a large percentage of the total amino acid content (15.5%).

To determine the number of Cysteine residues, a separate Amino Acid analysis experiment was performed (Runs 3-5 in Table 4). Cysteine residues require oxidation prior to acid hydrolysis to be detected because, like Tryptophan, acid hydrolysis can degrade cysteine and affect the quantitation. Furthermore, unfortunately Glycine is a common contaminant and has had to be excluded. However, for this analysis, enough amino acids have been measured to ensure an accurate outcome (personal communication, Peter Sharratt, PNAC Facility, University of Cambridge).

Amino acid	Expected Value	Observed Value Run 1	Observed Value Run 2	Observed Value Run 3	Observed Value Run 4	Observed Value Run 5
Cys	4	excluded	excluded	3.52	3.60	3.87
Asx	6	5.88	5.82	6.06	6.12	6.07
Thr	3	2.79	2.75	2.80	2.80	2.77
Ser	6	5.20	5.19	5.23	5.27	5.21
Glx	8	7.72	7.72	7.75	7.74	7.68
Gly	0	excluded	excluded	excluded	excluded	excluded
Ala	7	6.66	6.68	6.64	6.61	6.59
Val	4	3.97	4.02	4.05	4.09	4.00
Met	0	0.00	0.00	0.00	0.00	0.00
Ile	0	0.00	0.00	0.00	0.00	0.00
Leu	5	4.91	4.94	4.88	4.90	4.91
Tyr	0	0.00	0.00	0.00	0.00	0.00
Phe	1	1.02	1.05	0.92	0.92	0.92
His	1	1.00	0.97	excluded	excluded	excluded
Lys	3	2.86	2.83	2.74	2.77	2.86
Arg	1	0.99	1.03	0.92	0.77	excluded
Pro	0	excluded	excluded	excluded	excluded	0.00
Trp	1	1.00	1.00	1.00	1.00	1.00

	Run 1	Run 2	Run 3	Run 4	Run 5	
Total sample	4.062 23.98	3.843 22.69	4.060 23.95	4.100 24.19	4.087 24.13	nmoles ug/ml
Concentration	406.16 2397.71	384.34 2268.89	405.78 2395.34	409.85 2419.33	408.74 2412.96	nmoles/ml ug/ml

Table 4: Table of results from Amino Acid Analysis performed by PNAC facility, University of Cambridge. Runs 1 and 2 exclude Cys and Runs 3-5 include Cys analysis.

For all residues but Serine, the expected results are very close to the observed results (within 10%), The mean observed value for Serine residues is 13% lower than the expected value, which may be due to the tendency of Serine to be slightly degraded by the process of acid hydrolysis.

As a result of this analysis, the calculated extinction coefficient of 5500 M^{-1} cm^{-1} (assuming all Cysteine residues are reduced) was used throughout to calculate protein concentration by A280.

4.1.2 hVarp Znk2 partially dimerizes leading to aggregates.

Throughout the purification of hVarp Znk2, it became clear that there were two prominent species of the protein present during gel filtration, when using 20mM Tris, 200mMNaCl, 1mM $ZnSO_4$, pH7.4. The two large peaks, the ratios of which varied over multiple purifications, were both identified as hVarp Znk2. When run on an SDS PAGE gel, shown in Figure 30, both peaks result in a band at 14kDa and, although hVarp Znk2 is only 6.048kDa, it is 'gel shifted' during SDS PAGE separation (as previously confirmed by Mass Spectrometry).

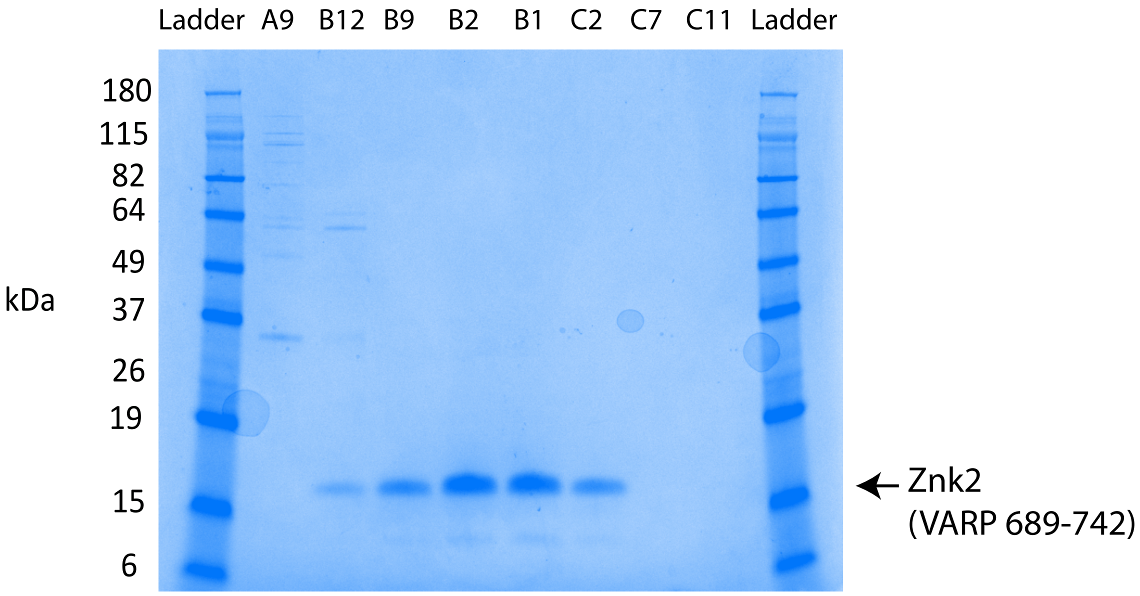


Figure 30: An SDS-PAGE gel of AKTA elution fractions.

Following gel filtration of hVarp Znk2 (see Figure 31 for elution profile), the two peaks of interest were collected separately and sent for analysis using Size Exclusion Chromatography with Multi Angle Light Scattering (SEC-MALS) to determine the oligomeric state of each of the samples. For SEC-MALS experiments, the samples were loaded separately onto a S75 10/300 column (GE Healthcare). The first peak P1 (37 minutes/ 11.1mls/0.46 column volumes) and the second peak P2 (42 minutes/12.6mls/0.53 column volumes) eluted at the same volume as previously in gel filtration. The calculated molar mass of the individual peaks indicates that P1 corresponds to a molar mass of approx. 12.5kDa whilst the second peak at P2 has a mass of approx. 6.53kDa (see the SEC-MALS profile in Figure 32). Additionally, the multi angle light scattering data in Figure 32 shows that the peaks themselves are monodisperse and do not contain mixtures of different oligomeric species. SEC-MALS is a useful technique to determine absolute molar mass as the results are not based on the apparent hydrodynamic radius of the protein. Because hVarp Znk2 is expected to be mostly unfolded, it will likely have a hydrodynamic radius that is large compared to folded proteins with a similar number of amino acids. This means that comparing the SEC data with folded protein standards is likely to be of little use when trying to determine the molecular weight of the peaks from SEC or indeed SDS-PAGE, as previously demonstrated.

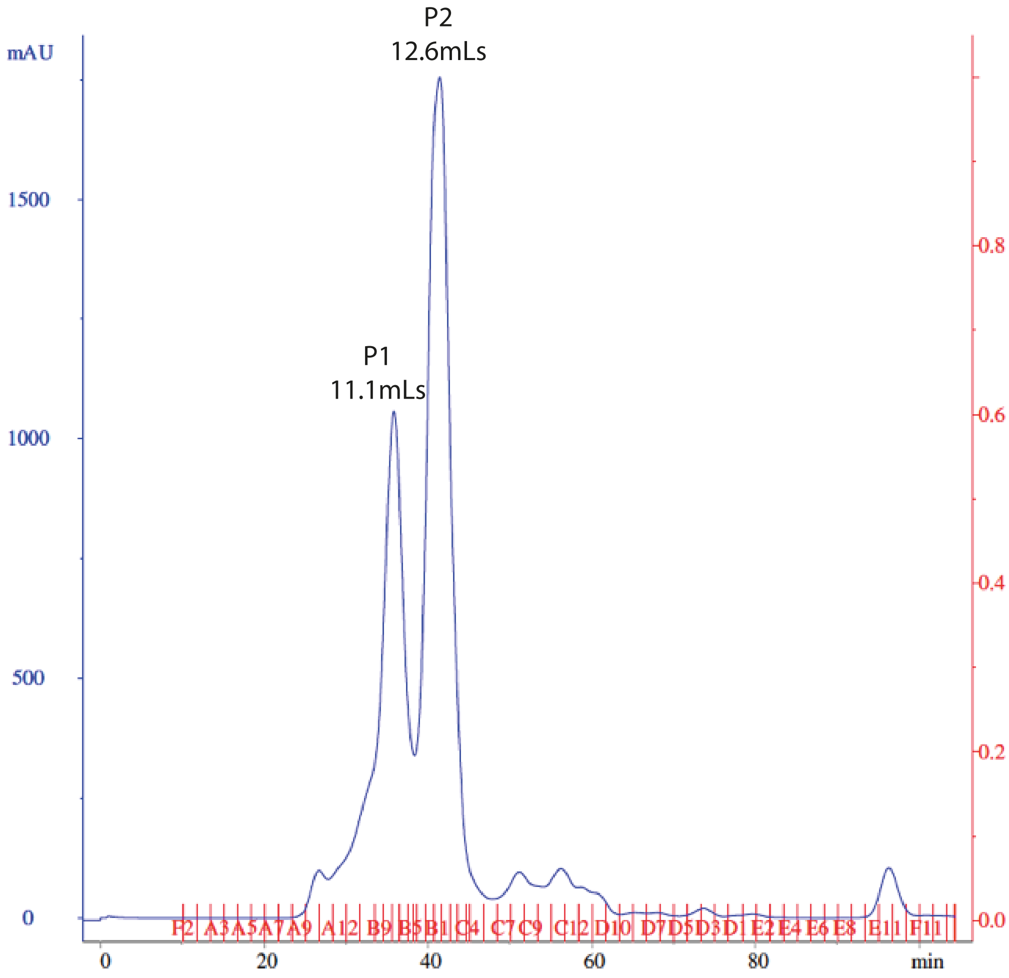


Figure 31: Elution profile of hVarp Znk2 on S75 10/300 column (GE Healthcare). The column was run in 20mM Tris, 200mM NaCl, 1mM $ZnSO_4$, pH7.4 at 0.3ml/min. The two peaks of interest are labeled, P1 at 37 minutes (11.1mLs) and P2 at 42 minutes (12.6mLs), as determined by SDS-PAGE gel electrophoresis.

The ProtParam calculated mass of the protein arising from the amino acid composition is 6.048kDa. It therefore follows that the peak arising at 6.53kDa is likely to represent a monomeric form of the protein, and the peak at 12.5kDa is most likely a dimer. The protein dimer sample was unstable and precipitated during any further purification and experimentation whereas the monomeric protein sample was more stable and was capable of being used for further work in this thesis.

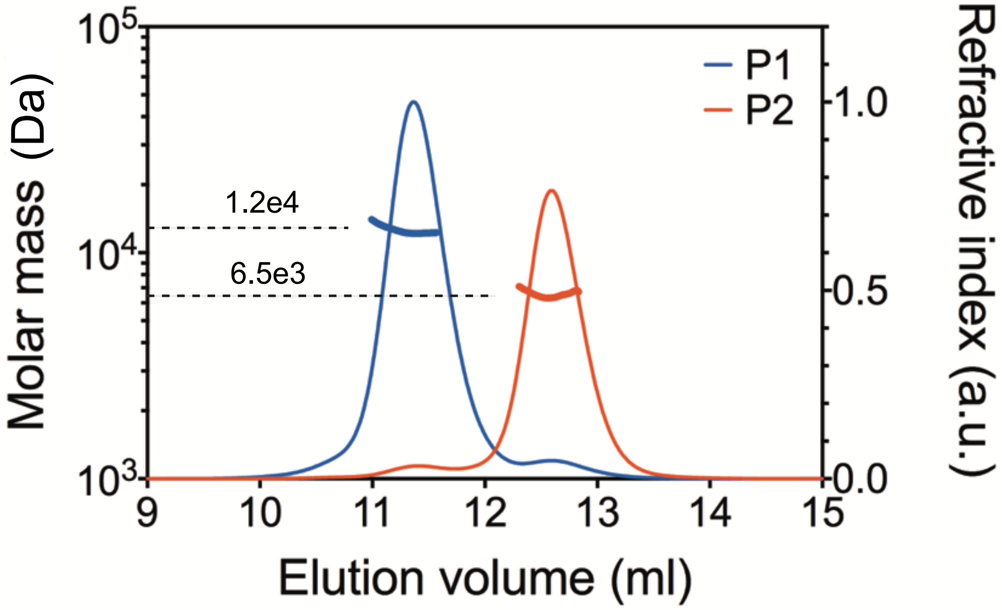


Figure 32: SEC-MALS profile of the two gel filtration peaks of Znk2. Performed using a S75 10/300 column (GE Healthcare) and run at 0.5ml/min.

The ratio of monomer to dimer varied depending mainly on the volume to which the gel filtration column input sample was concentrated, although the amount of time the protein sample was left between the GST cleavage purification step and the size exclusion gel filtration also had an effect. Using an S75 10/30 size exclusion gel filtration column required the sample volume to be between 25µL and 500µL, which was achieved using a VivaSpin 3000kDa molecular weight cut-off spin filter (Sartorius, Epsom, UK) with a Polyethersulfone (PES) filter. Reducing the sample to such a small volume meant that the concentration of contaminants in the sample was high, and also the increased surface area to volume ratio of the concentrated sample meant that a large percentage of the protein sample was in contact with the filter for a long period of time. This was not improved by the use of a cellulose membrane filter (Centriprep® filter units Millipore). This issue was reduced by the use of a preparative grade S75 26/60 column that can handle a sample volume of ≤13mL, although in this case the sample input volume was generally between 1-2mL, to retain good resolution between monomer, dimer, and contaminant peaks. By increasing the sample input volume in this way, the time to concentrate the protein sample prior to gel filtration was reduced. As a result, samples containing only monomeric protein were created for NMR experiments.

4.2 Resonance assignment of Varp Znk2

To achieve assignment of the backbone and sidechain atoms of free Varp Znk2, ^{13}C, ^{15}N – labelled Znk2 was prepared at 260μM, as described in the Materials and Methods, Section 2.3.1, using deuterated buffer components.

Initially, an ^{15}N-HSQC spectrum was acquired, which shows cross-peaks for amide resonances from both backbone and sidechain NH groups (Figure 33). Recombinant Znk2 is comprised of 61 amino acid residues, including 55 from the native protein, 5 N-terminal residues from the 5 non-native N-terminal residues resulting from the cloning and in addition a C-terminal tryptophan that was added to assist with protein concentration determination. The 5 proline residues in the protein produce no resonance in the ^{15}N-HSQC however, as proline lacks an amide proton. Conventionally, the ^{15}N-HSQC spectrum provides the initial starting point for resonance assignment. Using the CCPN Analysis software package [227], each peak in the ^{15}N-HSQC spectrum is given an arbitrary index number and is used as the basis for building up a spin-system; further resonances from other experiments (as described below) are subsequently added to form spin-systems, once it has been established they arise from the same residue. CBCANH and CBCACONH experiments are used to identify signals from carbon atoms that are in the same residue as a particular amide group, as well as those of backbone carbon atoms in the preceding residue. Overall, this process results in a set of N, NH, Cα and Cβ resonances that are linked into a spin system.

Additionally, the combination of CBCANH and CBCACONH experiments allows sequential spin systems to be linked to form a chain. The CBCANH experiment transfers magnetisation to a given amide not only from the Cα and Cβ atoms that are within the same residue, but also from the Cβ and Cα atoms in the preceding residue. In contrast, the CBCACONH experiment transfers magnetisation to the amide only from the Cβ and Cα in the preceding residue, (through the C' of the peptide bond and then on to the amide group for detection). As shown in the spectra in Figure 34 and the diagram in Figure 35, when the CBCANH and CBCACONH spectra are displayed overlaid on one another, this can reveal which peaks arise from the 'self' residue (i) (present only in CBCANH) and which arise from the preceding residue (i-1) (present in both spectra). In this way, it is possible to work back through connections identified in these spectra, linking together a series of resonances to form a chain.

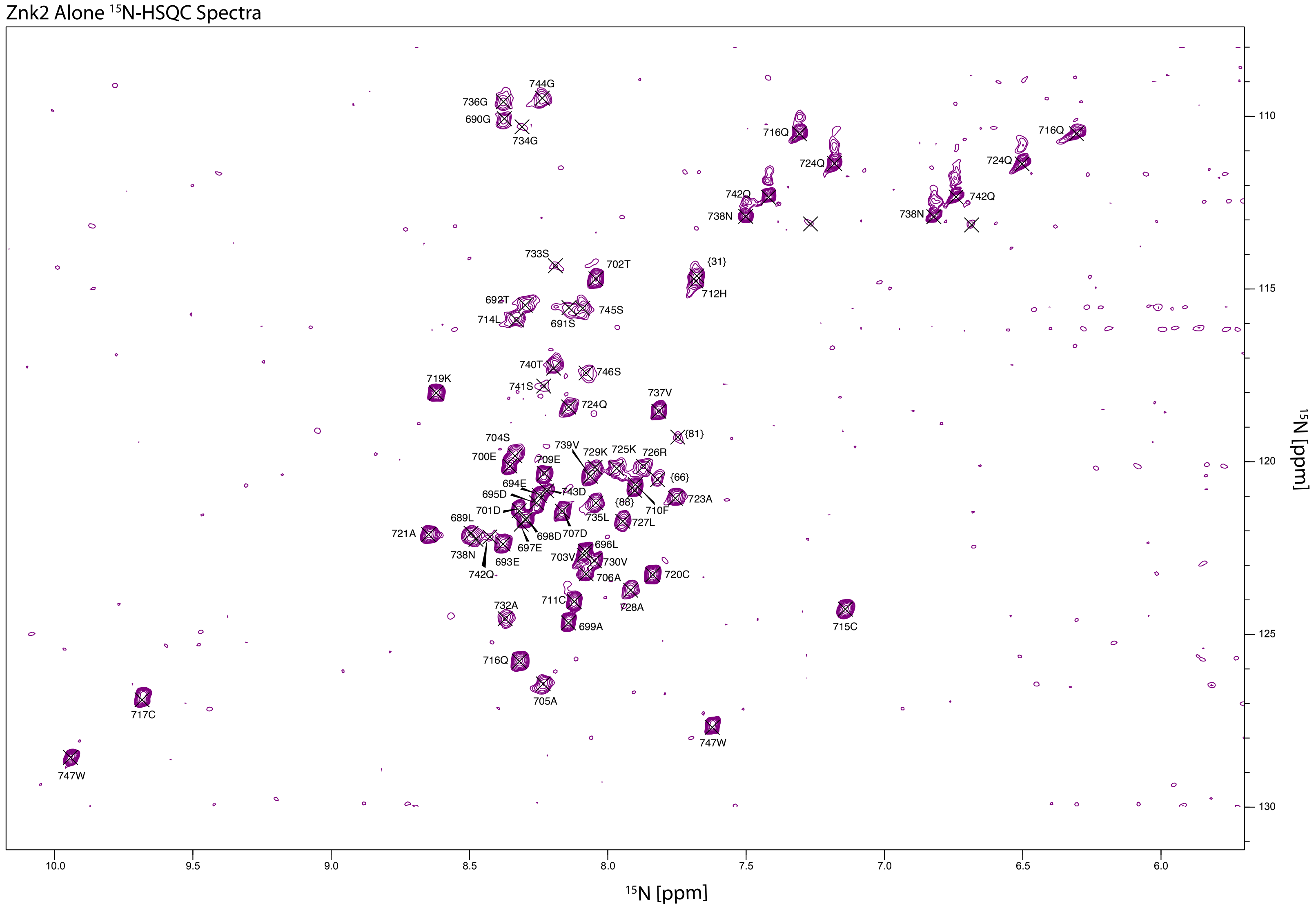

Znk2 Alone ^{15}N-HSQC Spectra

Figure 33: ^{15}N-HSQC of 260μM ^{15}N-labelled Znk2 in 20mM deuterated Tris, 20mM deuterated DTT, 200mM NaCl, pH7.0. Recorded using a Bruker Avance 800MHz spectrometer at 298K.

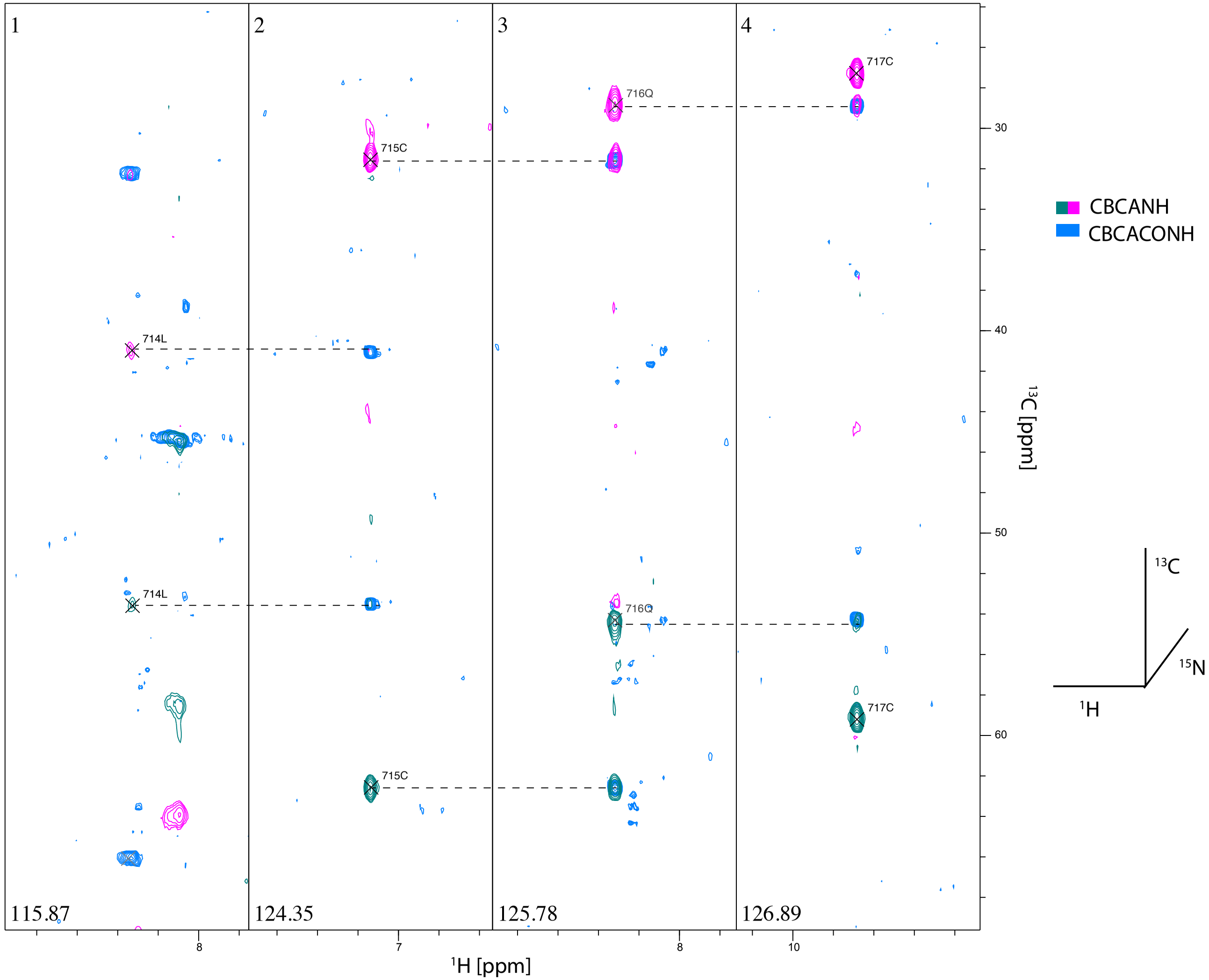


Figure 34: CBCACONH and CBCANH spectra of ^{15}N, ^{13}C labelled Znk2. These spectra have been overlaid and stripped in the ^{15}N dimension to show 2D slices of the spectra, illustrating an example of the connectivities used for making sequential assignments. The ^{15}N chemical shift (corresponding to the orthogonal dimension) is indicated at bottom left of each strip. The strips here show connectivites along the protein sequence from residue 714 to residue 717.

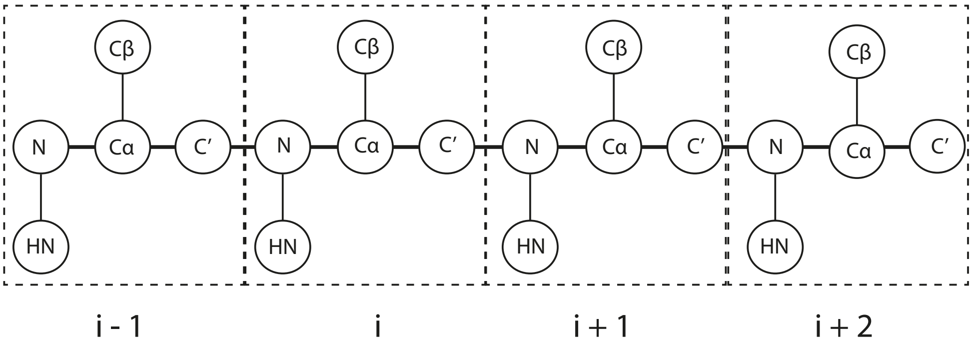


Figure 35: Diagram of atom connectivities in a protein sequence. In CBCACONH spectra, signals from Cα and Cβ of residue i-1 are visible, whereas in CBCANH spectra Cα and Cβ of both the i residue and the i-1 residue contribute (although the latter are weaker). Building a linked chain of amino acids relies on matching the intra residue resonances for Cα and Cβ in CBCANH of one NH to resonances in CBCACONH of another NH group.

Once a series of spin systems have been linked in this way, they must be individually assigned to specific amino acid positions in the known sequence of the protein. This is easier to achieve once the N, NH, Cα and Cβ of a given residue have been grouped together, as the distinctive Cβ and Cα shifts of some residues, coupled with the sequential information, facilitate identification of the amino acid type. In order to begin matching a linked chain of residues to specific locations in the known amino acid sequence, one must have a group of spin systems for which the linkage has been confidently and uniquely established, and which contain some distinctive amino acid types. Once one has identified such a group of perhaps 3 to 4 residues, which is often sufficient to correspond to a unique location in the sequence, this can be used as an anchor from which to continue the sequential assignment process. The reference chemical shifts of Cα and Cβ shown in Figure 36 demonstrate the most probable chemical shift values for Cα and Cβ for any amino acid type. To complete the assignment of the backbone atoms, HBHANH and HBHACONH experiments were used to find the shifts of hydrogen atoms attached to Cβ and Cα. Using this process, a nearly complete assignment of Znk2 backbone amide and carbon atoms in the free state was made.

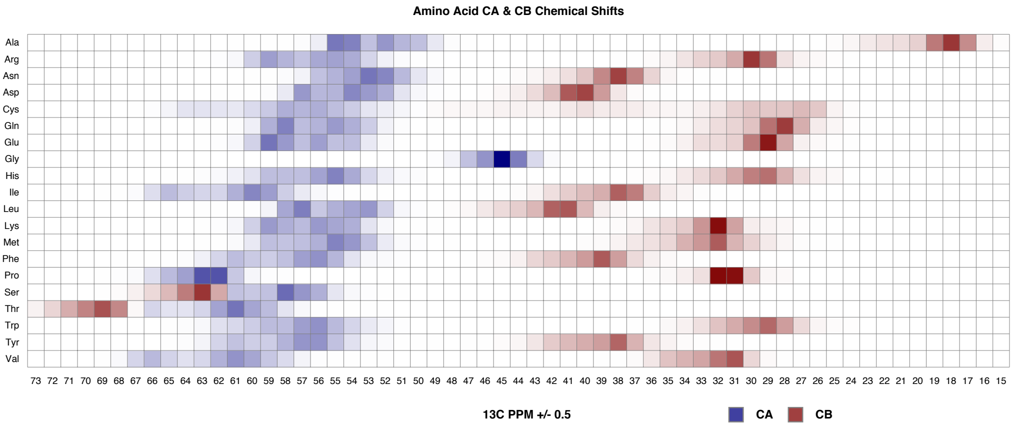


Figure 36: A schematic showing reference chemical shift ranges for Cα and Cβ from CCPN Analysis Software [227].These ranges are derived from statistics compiled by the Biological Magnetic Resonance Data Bank (BMRB) http://www.bmrb.wisc.edu.

Assignment of the sidechain atoms of the amino acids can often present more of a challenge. This is particularly the case for amino acids such as lysine and arginine that have long sidechains with multiple gamma, delta, epsilon or zeta groups, since these have the potential to contribute many overlapping resonances. To assign these side chain resonances unambiguously, HCCH-COSY and HCCH-TOCSY experiments were used in conjunction with ^{13}C-HSQC experiments, including some that were optimised for the detection of aromatic resonances. Using the previously assigned Cα and Cβ backbone shifts as a starting point, one can see the immediately adjoining carbon groups (i.e. through one bond only) in HCCH-COSY, while HCCH-TOCSY experiments show connections over multiple steps.

Similar to the Cα and Cβ shifts shown in Figure 36, reference chemical shifts are also available for the sidechain atoms of the various amino acid types. This information, coupled with the expected pattern of carbon resonances and number of carbons, called the ‘fingerprint’, for each amino acid type, allowed unambiguous assignment of most sidechain atoms. Figure 37 provides a visual representation of all the atoms in Znk2, with assigned resonances highlighted. A complete list of chemical shift values for the assignments is available in Appendix Three – Protein Resonance Assignments

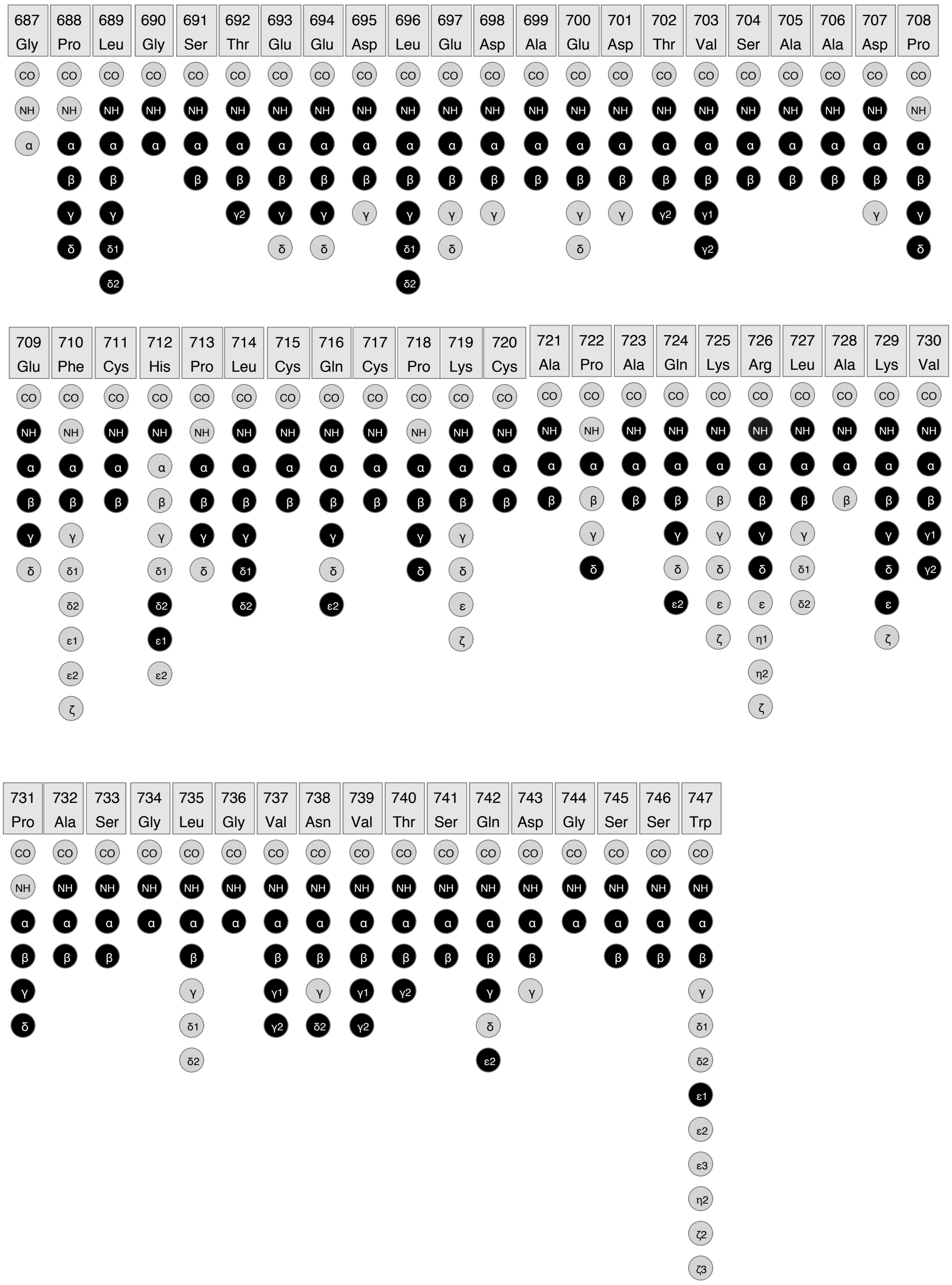


Figure 37: Assignment graph of free Znk2; atoms with a resonance assigned to them are shown in black.

4.3 Preparation of Vps29 samples for NMR

Murine Vps29 was recombinantly expressed from a pGEX4T2 vector in *E.coli.*, as described in Chapter 2, Section 2.2.2.3. The protein was expressed with an N-terminal Glutathione-S-Transferase tag with a Thrombin cleavage site (LVPR/GS) between. An SDS-PAGE gel of the initial purification steps can be seen in Figure 38, after which the sample was gel filtered to remove remaining contaminants and concentrated using a centrifugal concentrator. Murine Vps29 differs from human Vps29 by only 1 residue, with the final C-terminal residue being Serine in mVps29 and Proline in hVps29.

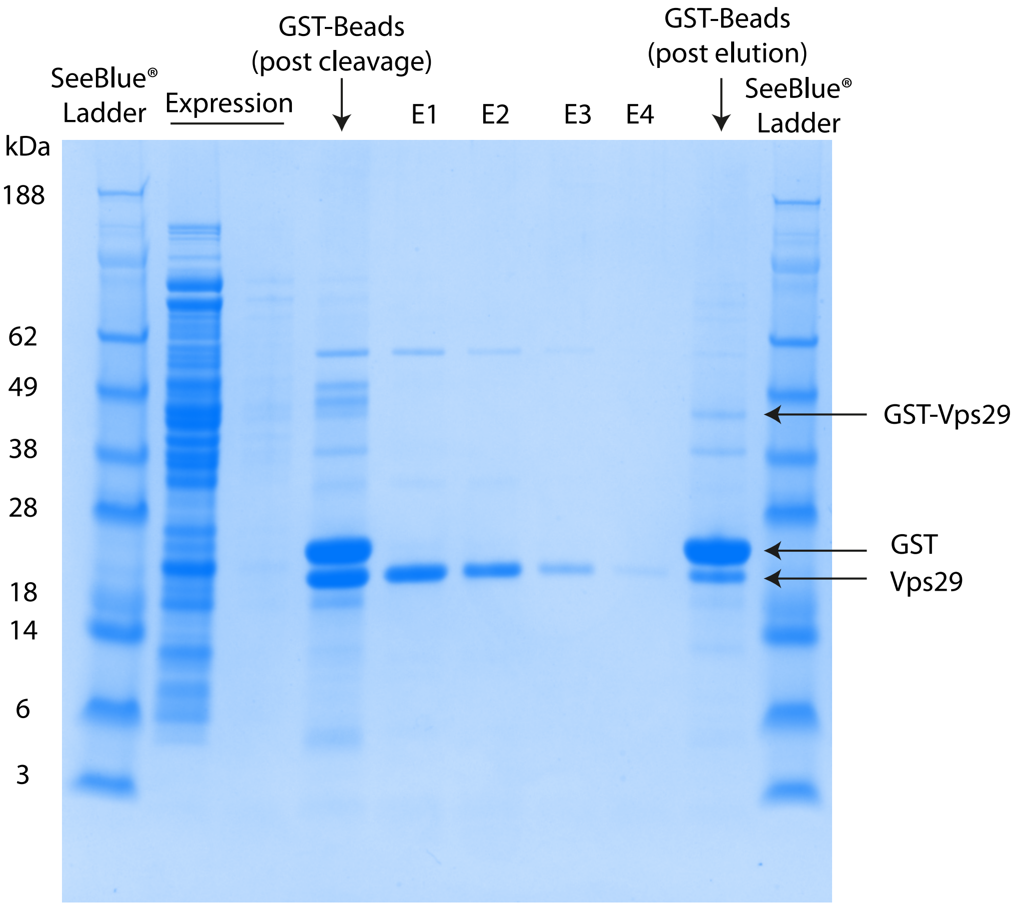


Figure 38: An SDS PAGE gel of the initial purification steps of Vps29. Vps29 has a molecular weight of 20.5kDa, GST of 26.9kDa, and GST-Vps29 fusion protein has a combined molecular weight of 47.4kDa. E1-4 refer to the elution steps of washing cleaved Vps29 from the GST sepharose beads, leaving GST on the beads (GST-beads (post-elution).

4.4 Resonance Assignment of Vps29.

A previous partial assignment of amide resonances for Vps29 was available prior to this project from the work of Swarbrick et al. [133] (who are collaborators with the Owen group). Whilst those authors did carry out NMR experiments with Vps29, and did include some amide resonance assignments in their publication, no resonance information was deposited in the BMRB database. They did offer to provide us with their resonance information, however given that the conditions used in the work described in this thesis for forming the complex with Znk2 differed from those used by Swarbrick et al., and given also that the spectra were quite crowded and the previous assignment was incomplete, it was necessary to undertake our own independent assignment of the Vps29 free protein to ensure accuracy and to provide reliable comparison data for our later assignment work on the complex with Varp Znk2. The assignment of Vps29 was undertaken in much the same way as described above for Znk2, although of course the significantly larger number of resonances for Vps29 made the task substantially more involved. Notably, even though there were chemical shift differences between the new data and that previously obtained for the assignment by Swarbrick et al [133], the previously recorded assignment did still provide a very useful comparison for some ^{15}N chemical shifts.

Wild type Vps29 has 182 residues, however the recombinant version used in this study has an additional 10 N-terminal residues resulting from the cloning. Being a larger protein than Znk2 means that there is more overlap of both backbone and sidechain resonances, as is clear from the ^{15}N-HSQC spectrum shown in Figure 39. Assigning the backbone resonances required the use of ^{15}N-HSQC, CBCACONH, CBCANH, HBHACONH and HBHANH experiments, as for Znk2. To assign the sidechain resonances, the spin system information from the previous experiments was used in conjunction with ^{13}C-HSQC, ^{13}C-CT-HSQC, HCCH-TOCSY, HCCH-COSY and DQF-COSY experiments. An assignment graph depicting which atoms have resonances assigned to them is shown in Figure 40, whilst a complete list of chemical shift values can be found in Appendix Three – Protein Resonance Assignments .

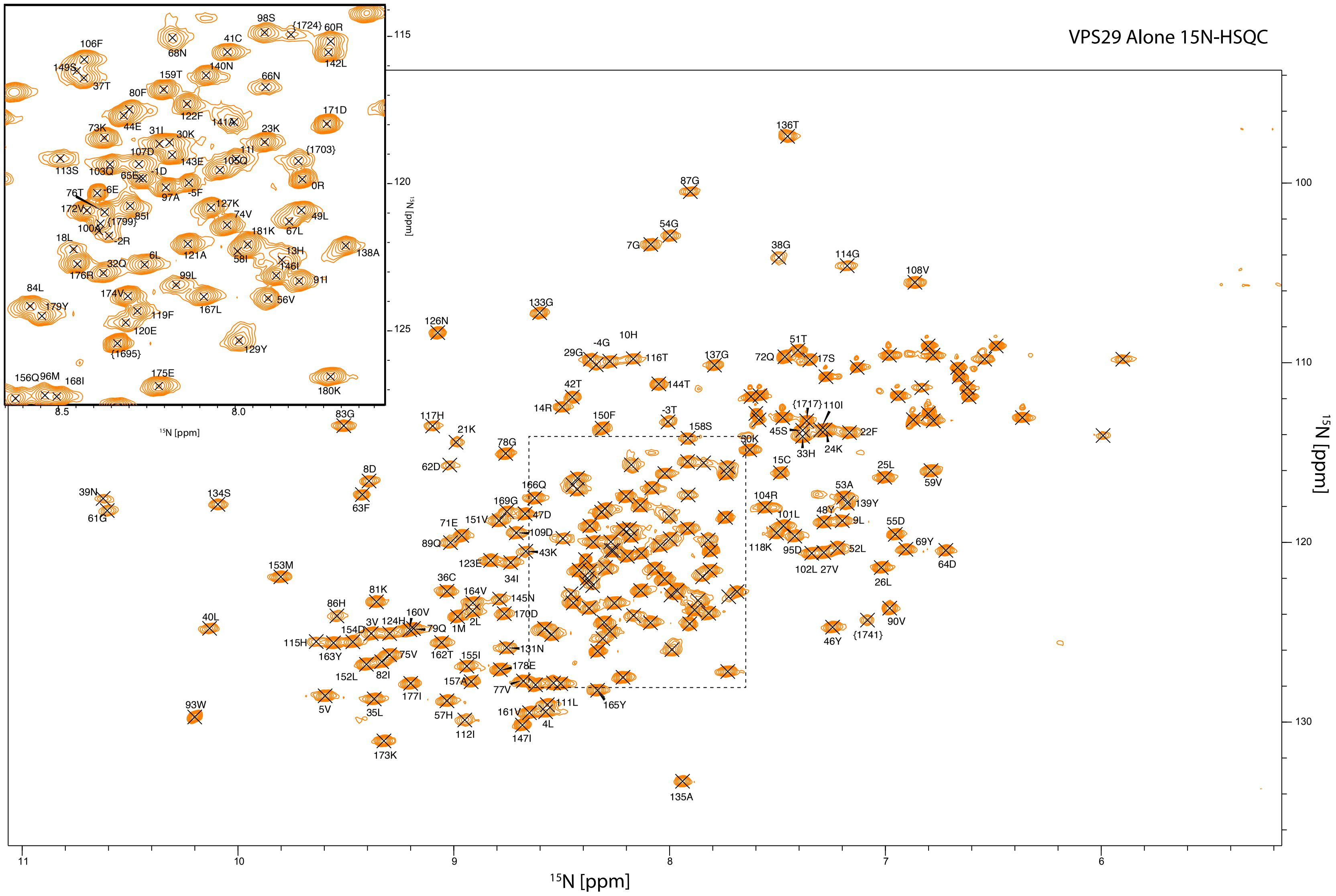


Figure 39: ^{15}N-HSQC of Vps29 in 20mM deuterated Tris, 20mM deuterated DTT, 200mM NaCl, pH7.0. Recorded using a Bruker Avance 800MHz spectrometer at 298K. The great majority of picked peaks with no marked assignment arise from sidechain CONH$_2$ groups of Asn and Gln.

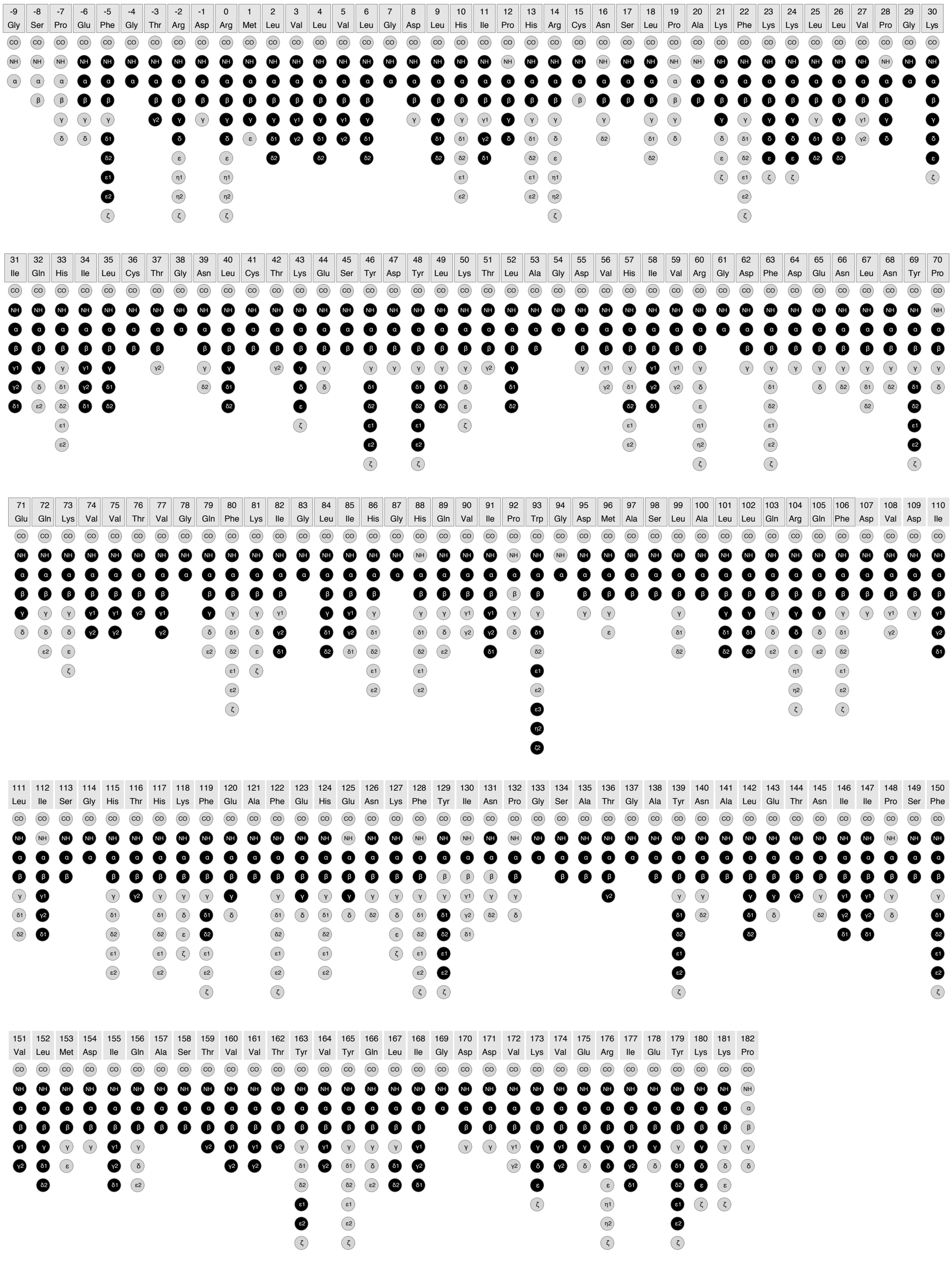


Figure 40: Assignment graph of free Vps29, atoms with a resonance assigned to them are shown in black.

4.5 Resonance assignment of Vps29 and Znk2 in complex.

Two types of samples of Vps29 and Znk2 in a 1:1 complex were created, each having one of the components (Vps29 in one case, Znk2 in the other) labelled with ^{15}N and ^{13}C isotopes and the other component at natural isotopic abundance. A suite of experiments was recorded on these different complex samples to enable resonance assignment of both components, carried out as in the case of the free proteins (see Sections 4.2 and 4.4). It was quickly apparent that many resonances in both proteins undergo large changes in chemical shift as a result of complex formation (see Figure 41 and Figure 47 for ^{15}N-HSQC overlays of free and bound spectra for Vps29 and Znk2 respectively). It was therefore necessary to carry out re-assignment of the affected resonances *de novo* in both proteins. This was done in essentially the same manner as described for the free proteins, and was greatly helped by using the free protein assignments as a reference point. By the end of this process, an essentially complete assignment of Vps29 in complex with Znk2 was achieved.

From the spectra shown in Figure 41 , it is clear that in the presence of Znk2, there are many substantial chemical shift changes of resonances for residues in Vps29, indicating that these residues are at, or near, points of interaction, or at the very least are affected by the binding, perhaps by a structural change within the Vps29 molecule [237].

For Znk2, many residues in the central region of the sequence had large chemical shift changes, see Figure 47. However, the assignment process was made very difficult in this case as the critical residues from Varp in the interface of the complex experienced severe line-broadening. Some ^{1}H signals also shifted into the immediate vicinity of the water signal, necessitating the preparation of samples in D_2O to allow these signals to be detected and assigned. Despite these various difficulties, an essentially complete assignment was also eventually achieved for Znk2 in the complex.

4.5.1 Chemical Shift Perturbations

Chemical shift perturbations help establish which regions of the protein, and which specific residues, may be involved in ligand binding. Comparing the ^{15}N-HSQC spectra of the free and bound states allows calculation of how much the amide group resonances of any specific residue have moved. The ^{15}N and ^{1}H chemical shift perturbations experienced by each amide group can then be combined into a single number (in this work this was done using the formula given in Section 2.3.2), normalised and plotted onto a histogram. This is a common method of identifying binding sites on proteins. It can be performed using either ^{13}C-^{1}H or ^{15}N-^{1}H groups, although the amide resonances from ^{15}N-HSQC are more often used as they are generally more sensitive to binding events [237] (probably because they respond to small perturbations of the hydrogen bonding network).

Measuring CSPs does not always require an independent assignment of the bound state. Titrations can sometimes be used to follow the movements of assigned peaks across successive additions of ligand, however this is only the case if the system is in fast exchange and the spectra are not too crowded. In instances where the kinetics of complexation results in a slow or intermediate exchange of peaks, or very crowded spectra, additional experiments are required to make an independent assignment of the bound state. As discussed in Section 4.5, for the Vps29:Varp Znk2 complex, both free and bound states were almost fully assigned for both proteins; for a list of chemical shift assignments see Appendix Three – Protein Resonance Assignments .

4.5.1.1 Titration of unlabelled Znk2 into labelled Vps29

^{15}N-HSQC spectra were recorded using samples of labelled Vps29 with unlabelled Znk2 at Vps29:Varp Znk2 ratios of 1:0, 1:0.25, 1:0.5, 1:0.77 and 1:1, with the CSPs being calculated from the final 1:1 sample. The CSPs were plotted in the histogram in Figure 42, which shows that two parts of the Vps29 sequence are particularly affected; residues Leu 25, Leu 26, Leu 152, Val 165 and Val 174 all show large changes in chemical shift. The chemical shift perturbations were also mapped onto a structure of Vps29 using a “colour ramp”, running from grey for residues experiencing no change to bright yellow for residues experiencing large changes. Because such colour ramps can easily be dominated by the largest CSP values, which would make the many informative but smaller variations that occur for other residues difficult to see, it is useful to define an “upper threshold” for the colour ramp, whereby it reaches its maximum colour saturation at a value that is scaled down relative to the largest CSP value in the dataset. In this case the scale factor was set to 0.5, meaning that the colour ramp reaches its maximum of bright yellow at a CSP that is 0.5 times the largest CSP value in the dataset; for the Vps29 data, the largest CSP value is 0.661) for Leu26, meaning the upper threshold for the colour ramp is 0.331.

As shown by Figure 43, the residues that have large chemical shift perturbations are mainly clustered in one region of the folded structure, suggesting that there is just one main binding site on Vps29 for Znk2. The histogram in Figure 42 also illustrates that residue I91 does not appear to be involved in Varp Znk2 binding, as previously reported by Hesketh et al [7].

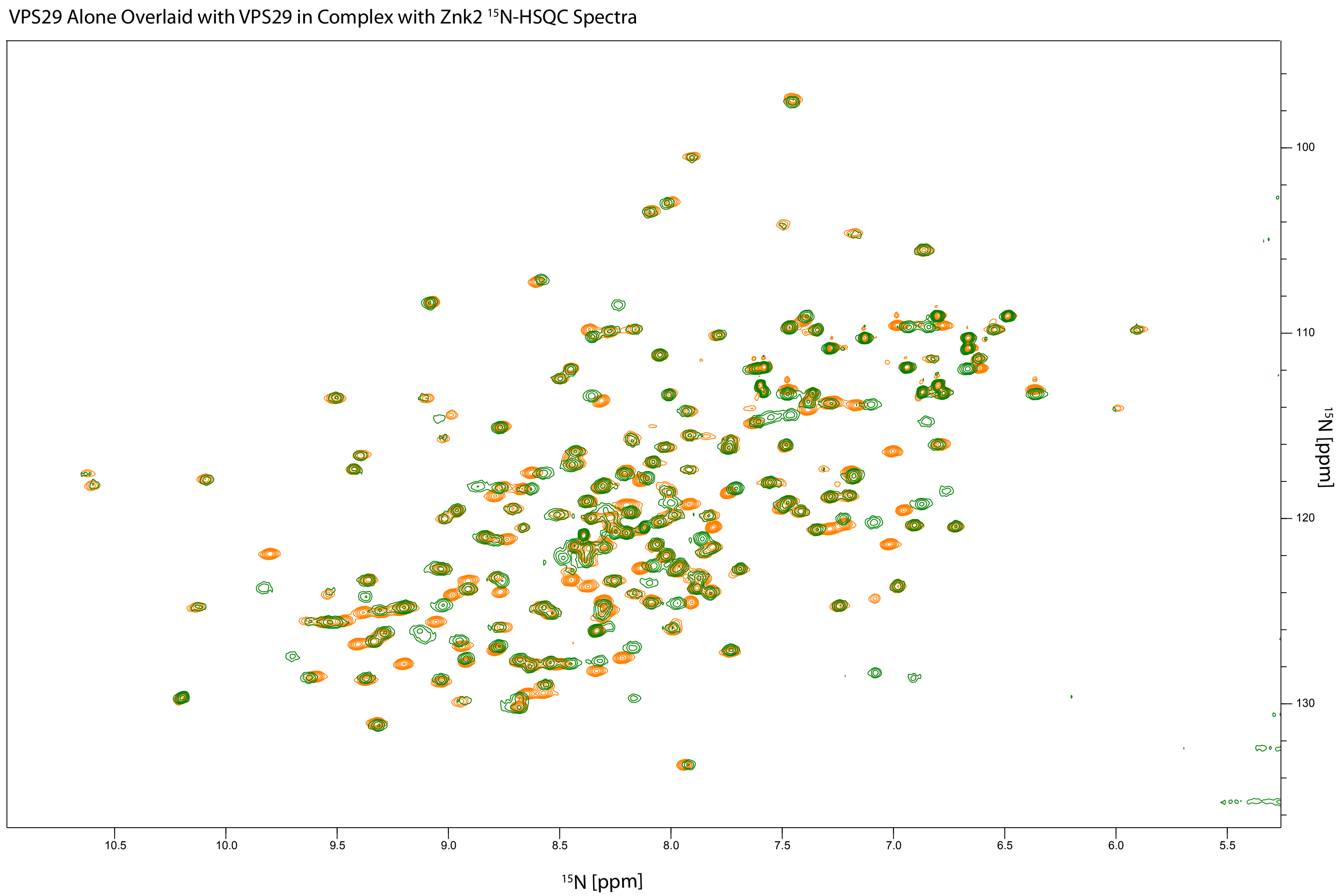


Figure 41: Overlaid ^{15}N-HSQC spectra of 300µM free Vps29 (orange) and 244µM Vps29 in complex with 244µM Znk2 (green). Both samples contained 20mM deuterated Tris, 20mM deuterated DTT, 200mM NaCl, pH7.0 and the spectra were recorded using a Bruker Avance 600MHz spectrometer at 298K.

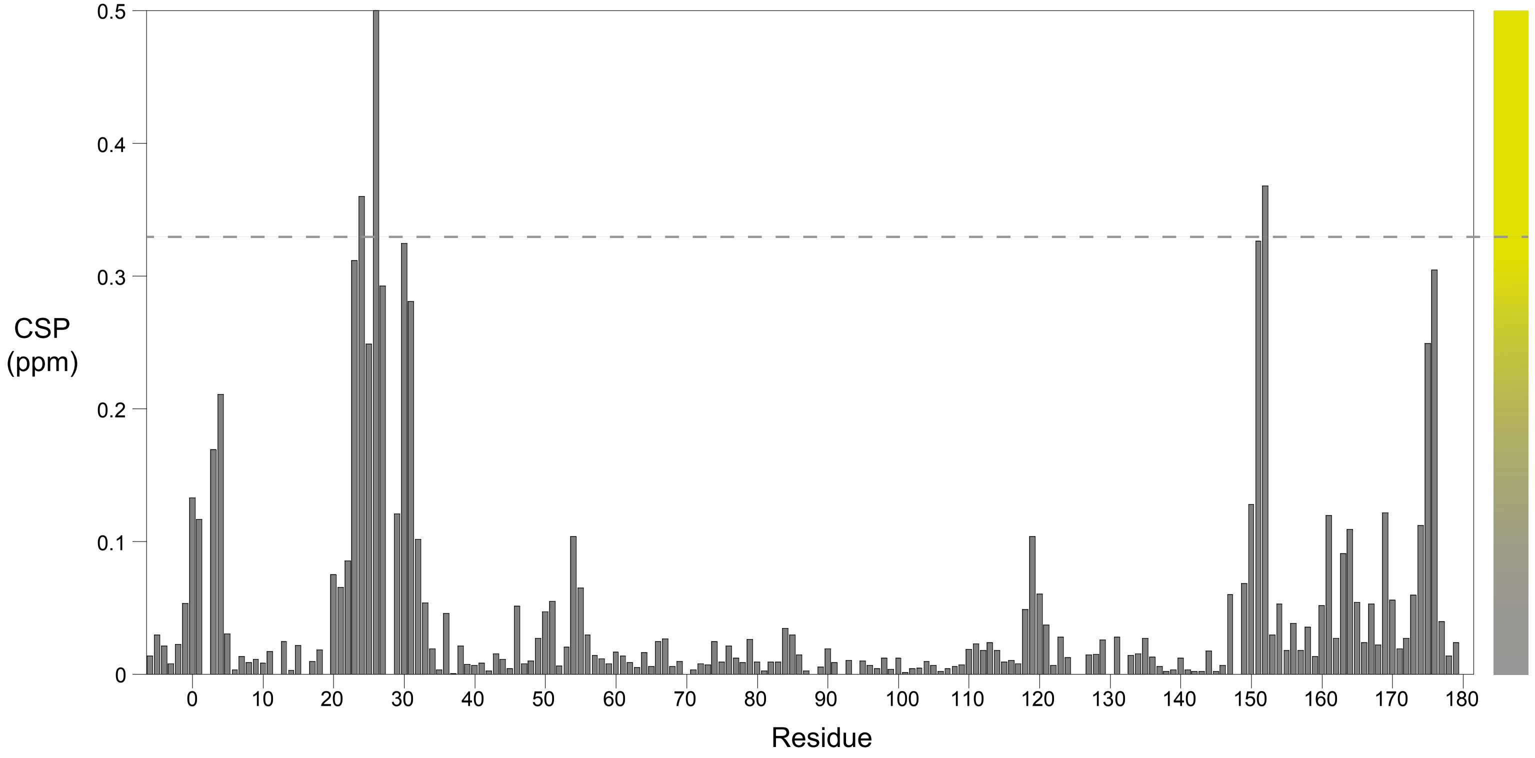


Figure 42: Histogram showing CSPs calculated from ^{15}N-HSQC chemical shifts of both free Vps29 and Vps29 in complex with Znk2. The graphical colour bar on the right depicts the intensity of the CSP. The largest CSP in the dataset is 0.662 ppm (for residue Leu 26, truncated in the figure); the colour gradient runs from grey to yellow and reaches its maximum at 0.5 x the largest CSP in the dataset, with values above this shown uniformly in bright yellow. This colour scheme has been used in Figure 43 to represent the intensity of CSPs.

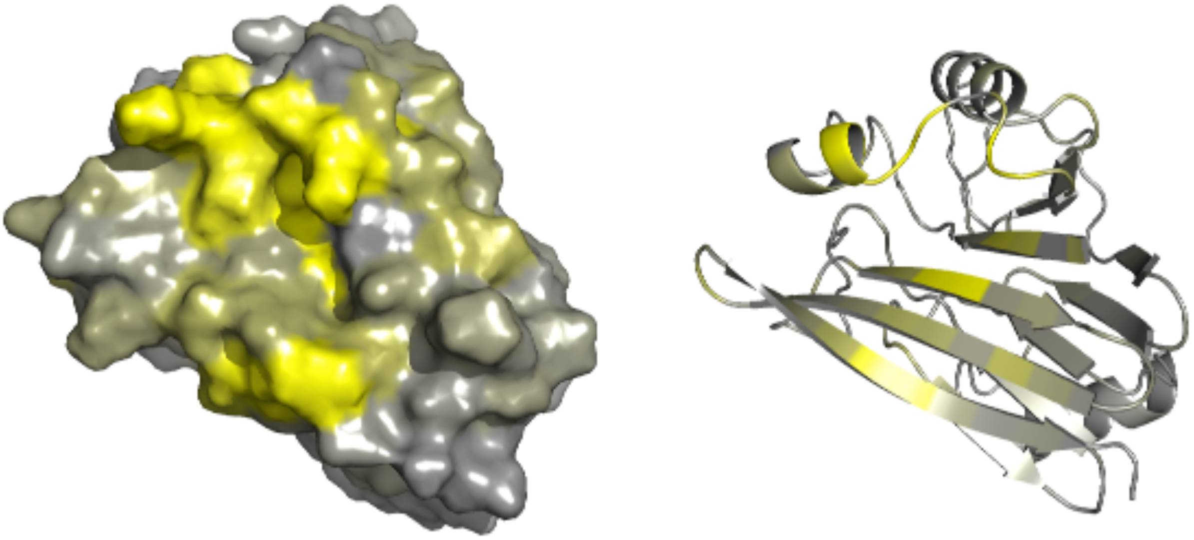
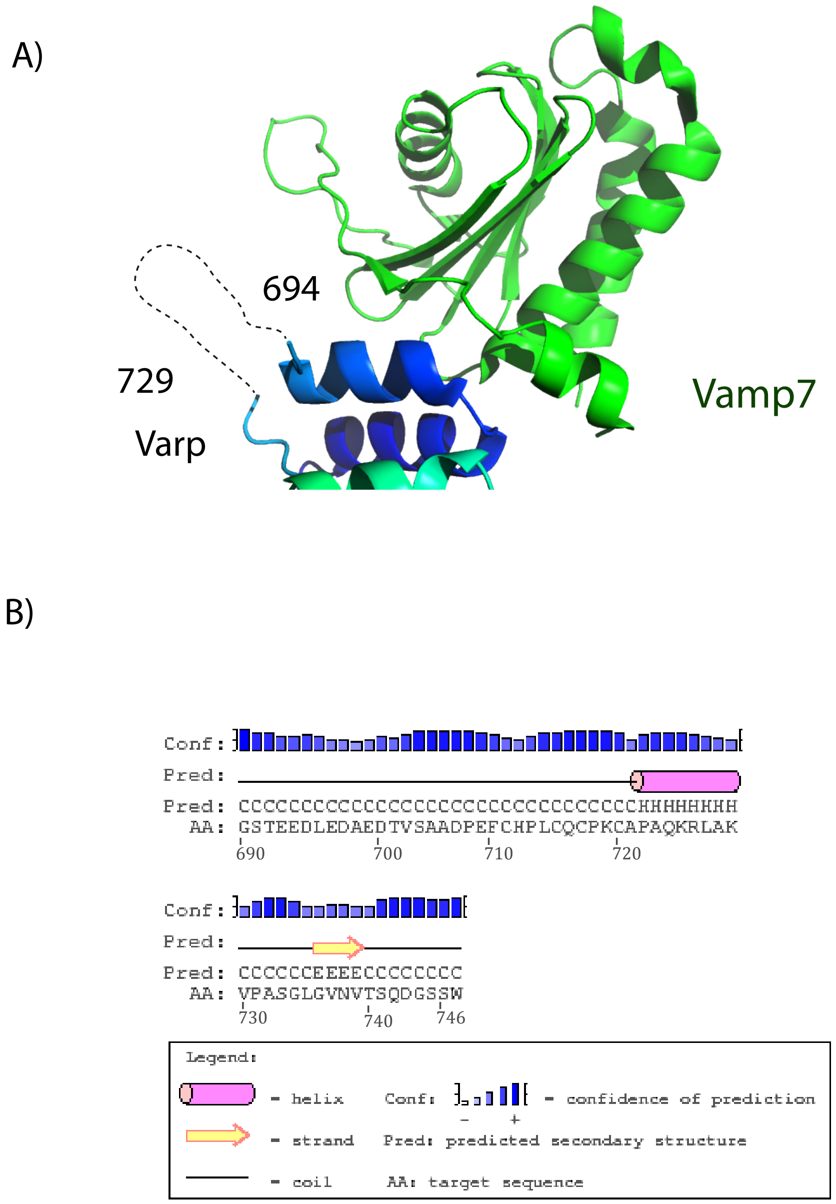


Figure 43: Amide group chemical shift perturbations (CSPs) for Vps29 resulting from addition of Varp Znk2 to a 1:1 ratio. The CSP values (calculated as $(\Delta\delta(^{1}H)^{2} + 0.2(\Delta\delta(^{15}N)^{2}))^{0.5}$) are mapped onto the structure of Vps29 using a colour ramp defined as described in the text and shown in Figure 42.

4.5.1.2 Titration of unlabelled Vps29 into labelled Znk2

^{15}N-HSQC spectra of labelled Znk2 with unlabelled Vps29 at Varp-Znk2:Vps29 ratios of 1:0, 1:0.25, 1:0.5 and 1:1 were recorded. As for the previous titration that used labelled Vps29 and unlabelled Znk2, the 1:1 Complex sample was used to calculate CSPs, using the same formula as before. From the overlaid ^{15}N-HSQC spectra in Figure 47 it is clear that there are large changes in chemical shift between the free and bound states. This was to be expected as the small Varp-Znk2 protein was predicted to be mostly unstructured (see prediction in Figure 44) and accessible to the solvent, whereas once bound to the Vps29 protein the exclusion of solvent, together with through-space effects such as ring-current effects from aromatic residues within the Vps29 structure, would create a strongly different environment. Indeed, by analysing secondary chemical shift data, shown in Figure 45, it can be determined that the N- and C-terminal regions of Varp-Znk2 are lacking in structure. Both free and Vps29-bound states of Varp-Znk2 demonstrate very similar Δδ13Cα - Δδ13Cβ values, which suggests that no structure is either gained or lost within the loop upon Vps29 binding. There is a region of large positive values around residues 708-731 in both free and Vps29-bound Varp-Znk2, which suggests these residues are largely structured, and that this is maintained in both free and bound states. The binding of Vps29 to Znk2 caused a great deal of line broadening for residues in the interface region of the complex, which made resonance determination and assignment very difficult, but also served to help identify the residues that are involved in complex formation. It is likely that the line-broadening results from intermediate rate exchange between the free and bound states of each component in the mixture, since the relatively weak affinity of this complex means that even at a 1:1 ratio a substantial proportion of each component remains unbound. For instance, for a 1:1 mixture in which both components are present at 200μM, given the K_D value of approximately 6μM for the interaction, about 16% of each would be expected to be in the unbound state, which would easily be enough to cause substantial line-broadening if the off-rate of the complex was in an unfavourable regime. At lower total concentrations, or if one of the components was in excess, the proportion of the unbound state would probably be higher, resulting in a still greater degree of line-broadening. As can be seen in both the secondary chemical shift graphs in Figure 45 and the histogram of CSPs from Znk2 in Figure 46, there are missing values for some residues in the bound state, which is a direct result of the bound state resonances remaining undetected due to significant line broadening in those cases.

It is clear from the CSP values that although other residues are affected by Vps29 binding to some extent, the strongest effects are seen for the region 708-729, which coincides well with the secondary chemical shift determined structured region of the Varp Znk2 loop.



```
Conf:
Pred:
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCHHHHHHHH
  AA: GSTEEDLEDAEDTVSAADPEFCHPLCQCPKCAPAQKRLAK
      690          700        710         720

Conf:
Pred:
Pred: CCCCCCEEEECCCCCCCC
  AA: VPASGLGVNVTSQDGSSW
      730         740  746
```

```
Legend:
 - helix     Conf:        - confidence of prediction
                   -  +
 - strand    Pred: predicted secondary structure
 - coil      AA: target sequence
```

Figure 44: Znk2 is located on a flexible loop. A) Structure of VID and ANKRD2 of Varp with VAMP7, PDB 4B93. Dashed line represents missing electron density for residues 694-729 of Varp. B) A structure prediction for residues 692-746 of the Varp Znk2 protein sequence, obtained using the PSIPred server [238].

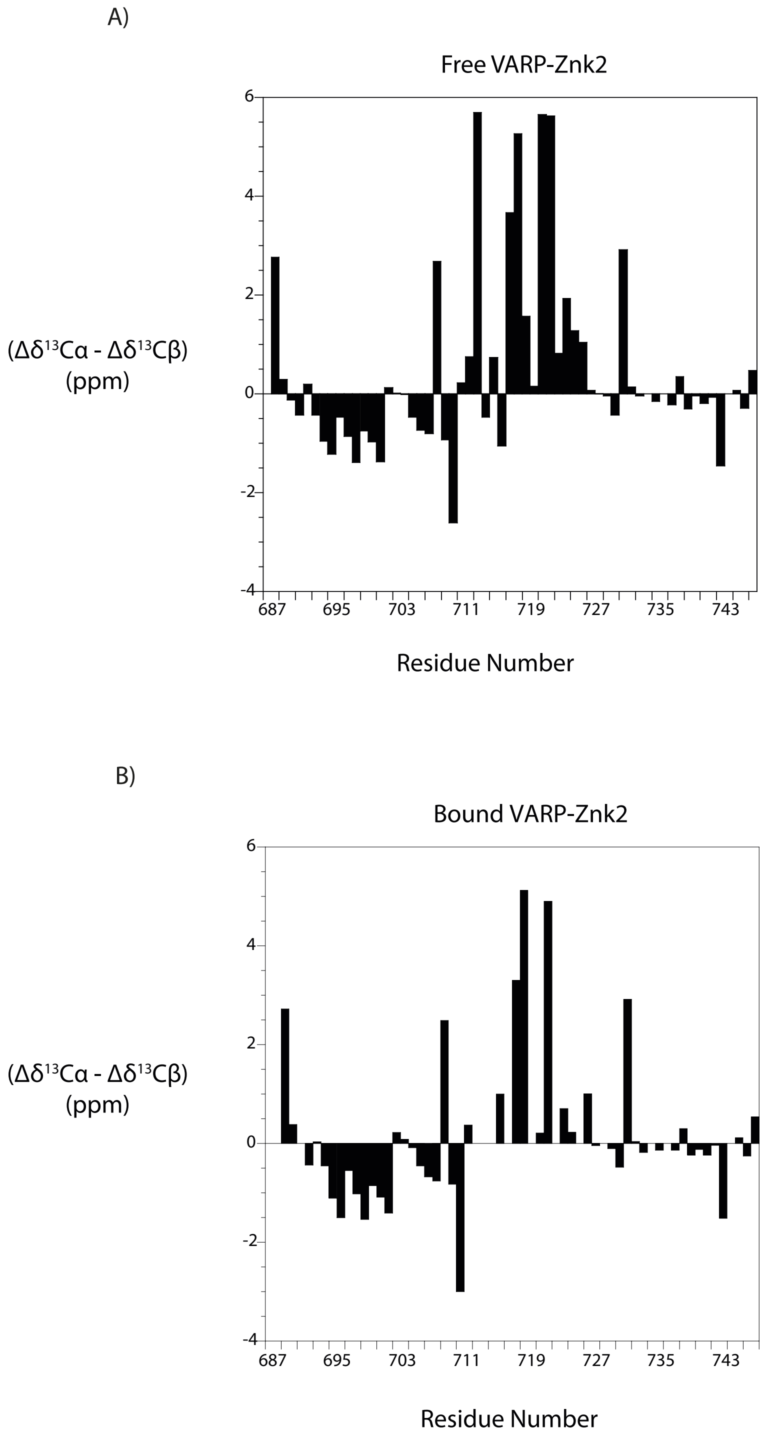


Figure 45: Secondary chemical shift data for Varp-Znk2, in both the free and Vps29 bound states.$\Delta\delta^{13}C\alpha$ is the difference between the experimentally measured shift of Cα for a given residue and the corresponding random coil value, while $\Delta\delta^{13}C\beta$ is the equivalent quantity for Cβ (values were calculated using CCPN analysis). A) $\Delta\delta^{13}C\alpha$ - $\Delta\delta^{13}C\beta$ values for residues in free Varp-Znk2. B) $\Delta\delta^{13}C\alpha$ - $\Delta\delta^{13}C\beta$ values for residues in Varp-Znk2 bound to Vps29.

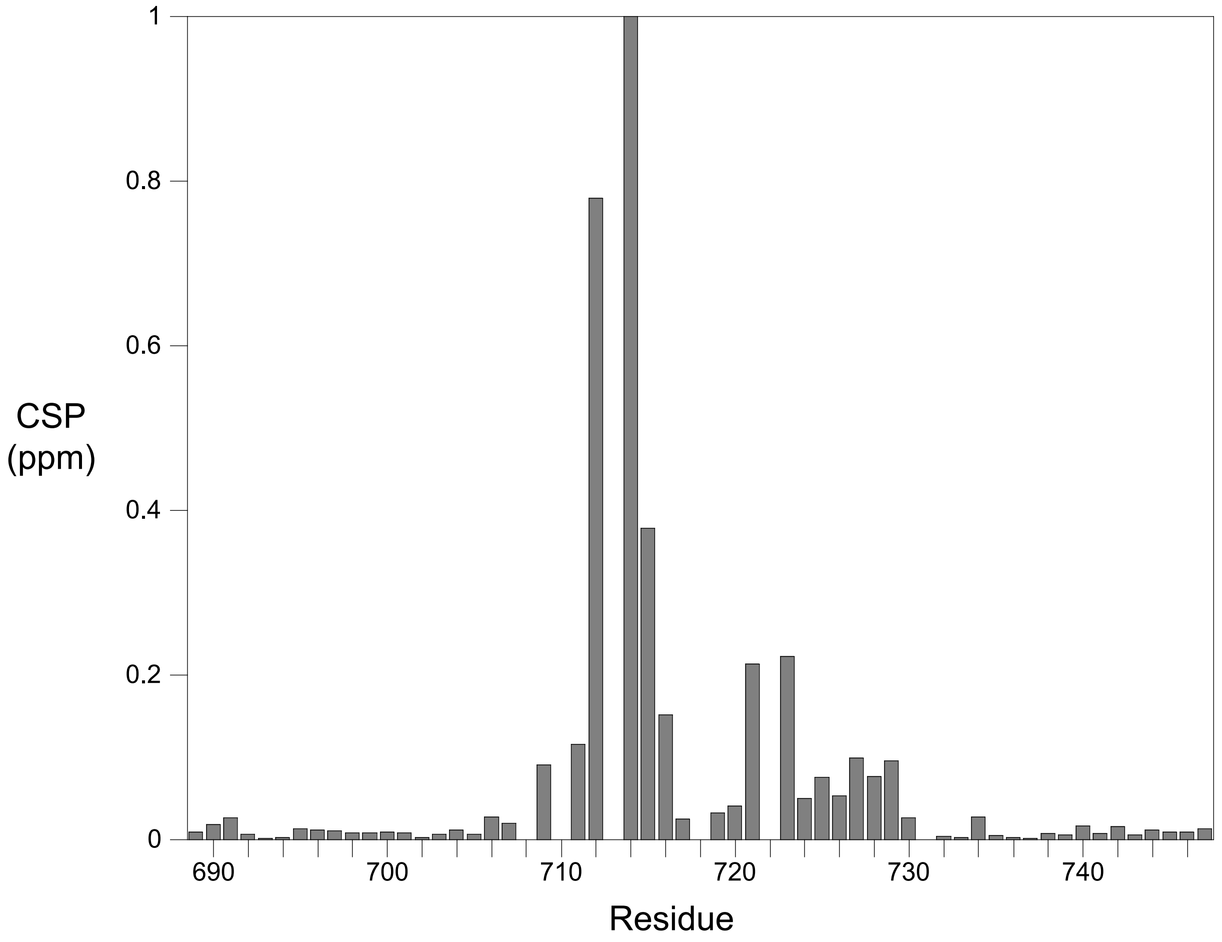


Figure 46: Histogram of chemical shift perturbations (CSPs) measured from ^{15}N HSQC resonances arising from Znk2 in free and Vps29 bound states.

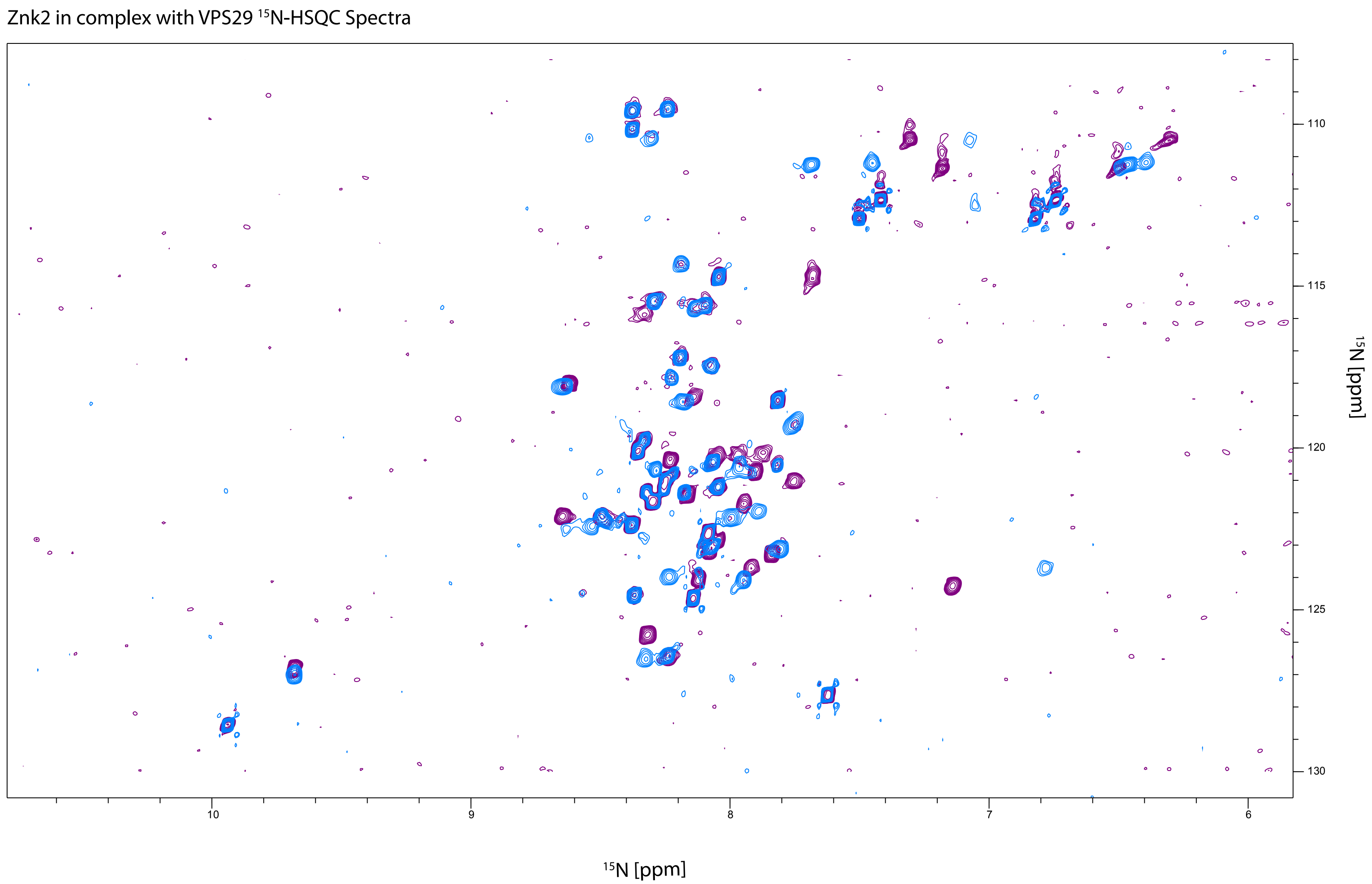


Figure 47: Overlay of 15N-HSQC of free Znk2 and Znk2 bound to Vps29. The purple spectrum is 272μM free Znk2 and the blue spectrum is 247μM Znk2 with 247μM Vps29 at 1:1. Recorded on Bruker Avance 800MHz at 298K, both in 20mM deuterated Tris, 20mM deuterated DTT, 200mM NaCl, pH7.0.

Chapter Five

5 NMR studies of Varp and Vps29; structure of the Vps29:Znk2 complex

Whilst CSP data can provide some information on the binding site for Varp Znk2 in Vps29, structural characterisation provides the most detailed view of how complex formation is mediated. The structure of the Vps29:Varp Znk2 complex was generated using the protocol described in Chapter 2, section 2.3.3. Briefly, this involved generating a template for Vps29 from an X-ray crystallography derived structure (PDB 2R17), and using the previously determined resonance assignments to determine intra- and inter-molecular NOE cross peak information to derive atom to atom distance constraints, and thus calculating complex structure ensembles using XPLOR. The structural information was then used to suggest mutant forms of Vps29, which were subsequently made and used for further SPR studies.

5.1 NOE derived distance constraints

The position and intensity of Nuclear Overhauser Effect (NOE) cross peaks between various atoms allows the determination of distance constraints, based upon the strength of the observed cross peak. NOE cross peaks arise between atoms that are close in space, and generally speaking, the stronger the NOE peak appears, the closer the atoms are, though other factors such as local flexibility can complicate this relationship. For complexes, preparation of samples in which only one component is isotopically labelled allows NOESY-type experiments to be filtered to show either intra- or inter-molecular NOE peaks.

Creating a distance constraint list for intermolecular atomic distances requires unambiguous assignment of NOESY cross peaks, which for protein-protein complexes are usually measured from an experiment in which two dimensions arise from the labelled protein, and one dimension arises from the unlabelled protein. In the case of the Vps29:Varp Znk2 complex, NOESY-type experiments were either filtered to exclusively show intermolecular NOEs, or unfiltered, in which case they showed both inter- and intra-molecular NOEs. These experiments were conducted on multiple samples of 1:1 Vps29:Varp-Znk2 with varied labelling schemes, varied mixing times and optimised for detecting signals from either aliphatic or aromatic residues.

For ^{15}N, ^{13}C-Znk2:Vps29 samples, intermolecular NOE cross peaks were particularly difficult to detect and assign. Key residues in the interfacial region experienced line-broadening effects that degraded the intensity of many cross peaks. Altering the experimental mixing time did, in unfiltered NOESY-type experiments, produce more readily detectable NOEs. However, due to the much lower sensitivity of filtered experiments (which are specific for intermolecular cross peaks) determining the origin of the NOEs arising in this situation was not straightforward. For ^{15}N, ^{13}C-Vps29:Znk2 samples, due to the larger range of residues that are involved from Vps29 in forming the complex interface, more intermolecular NOEs were detectable, although, the line-broadening effects seen for interfacial residues may have impacted the intensity of measurable NOEs, possibly obscuring some altogether. Again, filtered and unfiltered NOESY-type experiments, with different mixing times and carbon shift optimisations were needed to detect the widest possible range of intermolecular NOEs. By the end of the process, a total of 49 intermolecular NOEs had been assigned.

To assist complex structure generation, some intramolecular distance constraints were also required. In the case of Vps29, the bulk of the structural information used was supplied by the NCS constraints to enforce similarity to the crystal structure 2R17; however, as these NCS constraints were applied much more loosely in the interfacial region, it was useful to include some intramolecular NOE constraints for this part of the protein. As previously mentioned, residues in Znk2 experienced line broadening to a great extent which results in difficulty detecting cross peaks. Some intramolecular NOEs were determined for these residues in Znk2 however, and thus included in the constraints used to generate the ensemble structures. For details of the constraints used when calculating the final ensemble see Appendix Four – Structure Statistics.

Once a list of assigned NOEs had been compiled from these data, these were classified into very strong (0-2.3 Å), strong (0-2.9 Å) medium (0-3.5 Å) and weak (0-5.0 Å) intensity categories. The upper distance bounds used for these categories were calibrated using assigned NOE cross peaks in the unfiltered NOESY spectra of samples in which the Vps29 component was labelled; specifically, the intensity corresponding to the very strong category (0-2.3 Å) was set using sequential $d_{\alpha N}$ peaks in regions of regular anti-parallel β -sheet, that of the strong category (0-2.9 Å) was set using sequential d_{NN} peaks in regions of regular α -helix, that of the medium category (0-3.5

Å) was set using (i, i+3) $d_{\alpha N}$ peaks in regions of regular α -helix, and those in the weakest category were set to be consistent with the expected approximate detection limit for NOES, set as 5.0 Å. This calibration was extended for use in various types of filtered spectra, in each case by comparing the measured intensities for particular specific NOE connections as they appeared in both the filtered and unfiltered spectra, using only signals that could be observed free of overlap in both spectra. As a check that the calibration of the intermolecular NOE-based restraints was consistent with the structure calculations and force field, test calculations were also run in which the upper bounds for just these restraints in the final set were either tightened or loosened slightly; it was found that tightening them caused a significant increase in their violations, whereas loosening had relatively little effect since violations were already very few. For a list of NOESY type experiments used for structure complex determination, see Appendix Two – NMR .

5.2 Vps29 Template – Chemical shift and NOE-derived peptide bond conformations.

As previously described in Chapter 2, the generation of the Vps29 template structure was an iterative process. The X-ray crystallographic structure of Vps29 in PDB 2R17 was used as a starting point; this structure was chosen in preference to PDB 1Z2X that had previously been determined in the Owen group [147], as 2R17 lacks the protruding conformation of helix 2 that is apparent in 1Z2X, and which results from a crystal contact to its counterpart in a symmetry-related molecule. This conformation seen in 1Z2X is inconsistent with our NOE data (there are clear contacts between Trp93 (HZ2, HZ3, HE3 and HH2) and both Leu142 (HD*, HG) and Ala141 (HB*) that could not occur if helix 2 was in the conformation of 1Z2X).

To generate the template structure, chain A was extracted, the Seleno-Met residues changed to Met, and the peptide bonds 11-12, 40-41 and 91-92 specified as having the *cis* conformation. Information from both NOESY spectra and other NMR experiments can be used to confirm the *cis* confirmation of these bonds in solution. It is well accepted that certain patterns of NOEs are characteristic for either *cis* or *trans* conformations of Xaa-Pro peptide bonds [239]. Not only this but the chemical shift difference between the Proline Cβ and Cγ are also diagnostic in this situation, with a difference of 0-7.5 indicating a *trans* conformation and >8.0 indicating a *cis* conformation [240]. Both Ile11-Pro12 and Ile91-Pro92 are examples of Xaa-Pro peptide bonds and so these criteria can be applied.

From the NOESY cross peak assignment performed in this study, it was evident that residues 11 and 12 display an NOE pattern characteristic of a *cis* peptide bond, (namely, there is a strong NOE cross peak between 11Leu Hα and 12Pro Hα and no NOE cross peak from 11Leu Hα to 12Pro Hδ, see Figure 48) the chemical shift difference is also consistent with a *cis* bond, with Cβ shift of 33.7 and a Cγ shift of 23.8. Proline 92 of Vps29 shows a characteristic chemical shift difference of 8.5 between the Cβ (34.0) and Cγ (25.5) chemical shifts, and the NOEs arising between residues Ile91 and Pro92 are characteristic for a *cis* bond also.

Cis peptide bonds are more unusual for non-Proline sequences. In this instance, it seems that the bond between Leu40 and Cys41 is likely to be in a *cis* conformation, unlike in the X-Ray PDB structure in deposition 2R17. Unfortunately, no NOE cross peaks were detected between Leu 40 Hα and either Cys 41 Hα or Cys 41 HN, and ^{13}C chemical shift statistics for non-proline cis peptide bonds have not been complied, so neither conformation could be confirmed by the experimental NMR data. However, in other high-resolution crystal structures of Vps29 PDB structure files this bond is modelled as *cis* (PDB 5GTU, 5WYH, 1W24). Also, it became apparent during rounds of template generation (as described in Chapter 2) that a *trans* conformation of this bond resulted in large Ramachandran violations during the XPLOR calculations, and therefore the decision was made to alter this bond to a *cis* conformation in the Vps29 template used in this study.

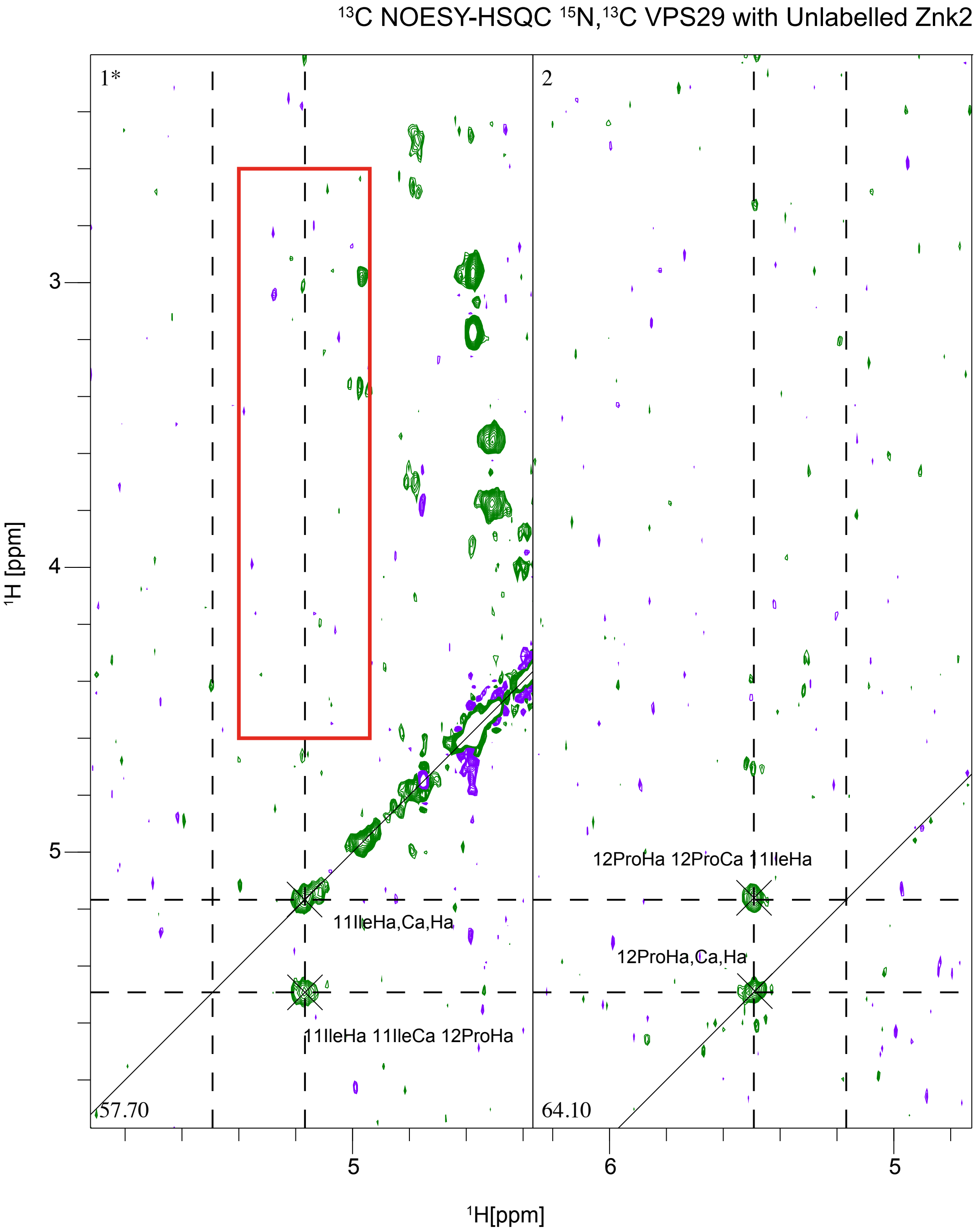


Figure 48: ^{13}C-NOESY HSQC experiment gathered from a sample containing ^{15}N, ^{13}C labelled Vps29 in complex with unlabelled (natural abundance) Znk2. Peaks on the diagonal arise from the 'self' residue, whilst those in the vertical column arise from other atoms close in space. These strips clearly show NOE cross peaks between 11Ile Hα and 12Pro Hα. The red box indicates the region in which one would expect to see an NOE cross peak from 11Ile Hα to 12Pro Hδ, had the peptide bond been *trans* so that these atoms would have been close enough in space to one another to generate a cross peak.

5.3 Chi1 angles of Zinc co-ordinating Cysteines in Varp

To determine the Chi1 angles of the zinc co-ordinating cysteines in Varp Znk2, HBNH and HACAHB experiments were used to estimate the relative sizes of $^3J\alpha\beta$ and $^3JN\text{-}\beta$ coupling constants from the relative intensities of the cross peaks in these experiments. Using this information, the predominant rotamer about the Cα-Cβ bond in each Cys residue can be determined, as in [241], and used to define approximate structural constraints.

	60.0° **Peak intensity**	**180.0°** **Peak intensity**	**-60.0°** **Peak intensity**
$^3J\alpha\beta2/3$	Weak/Weak	Weak/Strong	Strong/Weak
$^3JN\text{-}\beta2/3$	Strong/Weak	Weak/Weak	Weak/Strong

Table 5: A table describing the expected observed peak intensities for Chi1 angle conformations as reported by [241].

The following rotamers were determined for the zinc binding cysteines in Varp and used in the subsequent complex structure calculations: Cys711 -60.0°, Cys715 180.0°, Cys717 60.0°, Cys720 -60.0°. These data relate to Varp Znk2 in the free state, since the severe line-broadening of the Varp Znk2 Cys Hβ signals in the complex with Vps29 precluded the use of similar experiments for the complex. While it is an assumption to suppose that these rotameric states would necessarily remain the same in the complex as they are in the free state, it was thought not unreasonable to apply loose Chi1 angle restraints to maintain them in the structure calculations.

5.4 Complex Structure

The structure calculations, carried out using the protocol described in the Materials and Methods section 2.3.3, resulted in the generation of an ensemble of 50 final structures.

The Vps29:Varp Znk2 complex structures were ordered according to their XPLOR energies, and the 25 structures with the lowest total energy were used to create a PDB ensemble file, superimposed on either the Vps29 molecule or the Znk2 molecule. Figure 49 depicts both the single lowest energy state of the complex, as well as an ensemble of 25 structures superimposed on the Vps29 molecule. For the full statistics for the final calculated structures, see Appendix Four – Structure Statistics.

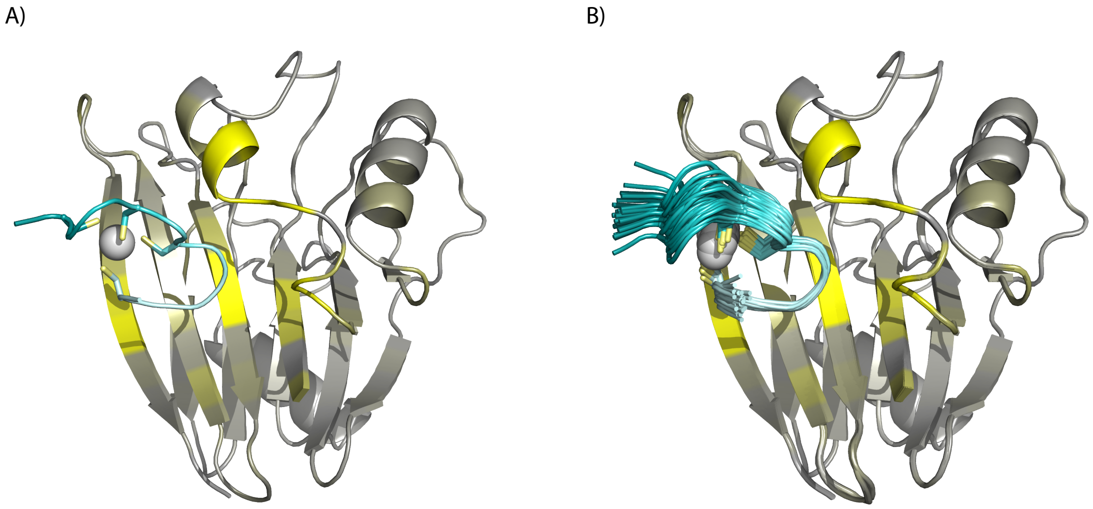


Figure 49: Structure of the Varp Znk2:Vps29 complex. A) The lowest energy structure from the ensemble of the calculated Varp Znk2:Vps29 complex with Varp Znk2 in cyan, and Vps29 in gray. Chemical shift perturbations of Vps29 are displayed as in Figure 43 ranging from gray to bright yellow, depending on intensity. B) An ensemble of 25 Varp Znk2:Vps29 structures, superimposed on Vps29. Chemical shift perturbation information is displayed on Vps29, the colour ramp shown here is identical to that in Fig 43.

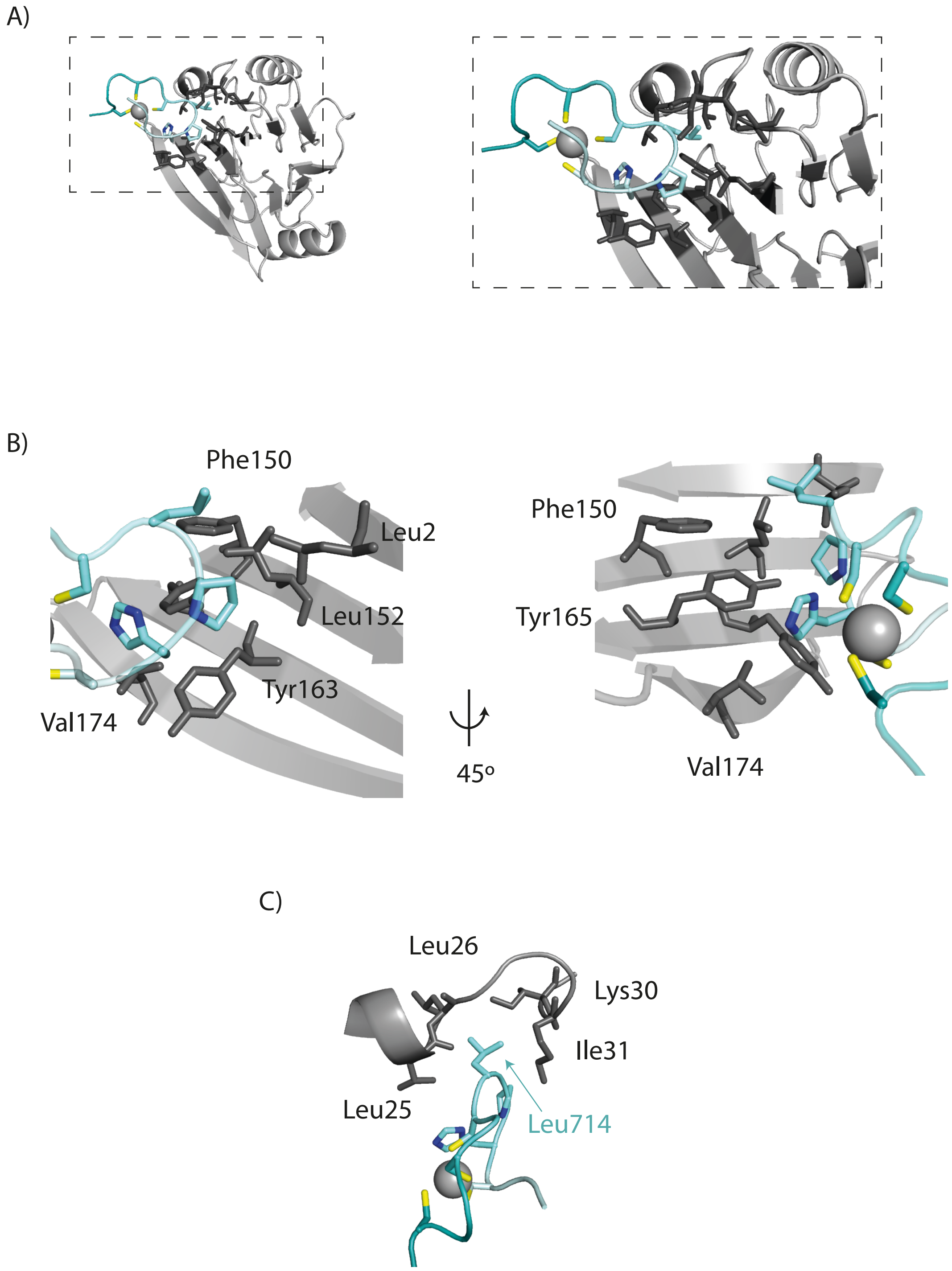


Figure 50: A closer view of the solution structure of Varp-Znk2 in complex with Vps29. The N- and C- terminal tails of Varp-Znk2 are omitted for clarity, as is the cloning artefact tail on the N-terminal end of Vps29. A) The ordered region of Varp-Znk2 is shown in aquamarine, Vps29 in grey, with residues from which distance constraints to Varp-Znk2 were determined (using NOESY cross peaks) in stick format, darker grey. B) Close up of the Varp-Znk2 binding residues in the beta sheet of Vps29. C) A close up of the Varp-Znk2 binding residues located on the helix and loop in the binding site of Vps29.

A single structure from the ensemble, that with the lowest total energy, is shown in Figure 50. This structure clearly demonstrates that Varp-Znk2 binds to a small patch on Vps29, formed by a small helix, and a beta sheet. The Vps29 residues shown in 'stick' style in Figure 50 are those for which defined intermolecular distance constraints based on NOEs were used in calculating the structure. In total, 49 intermolecular NOEs were assigned and used to define distance constraints for the calculation. Residues across the surface of the Vps29 beta sheet are seen interacting with residues in Varp-Znk2; Figure 50B shows a close up of these residues. The curvature of the beta sheet and the local sequence together create a pocket for Varp-Znk2 to bind into. The total buried surface area of the interface is only 659 ± 22 Å^2 (calculated using the CCP4 program PISA, after first deleting the atoms of the disordered tails of both Vps29 and Varp), and this small interface requires very few residues from Varp-Znk2 to form the complex. The coordination of zinc in the Varp-Znk2 peptide plays a key indirect role, as it is this which produces the conformation of His-Pro-Leu necessary for the interaction. As depicted in Figure 50, the main residues from Znk2 that contact Vps29 are His712, Pro713, and Leu714. The projection of Leu714 into a cavity produced by Leu25, Leu26, Lys30 and Ile31 (Figure 50C) suggests that it is mainly this hydrophobic interaction that is critical in complex formation. It is possible there may also be a contribution from two intermolecular hydrogen bonds linking the ring OH groups of two well-conserved Tyr residues on Vps29 to two backbone carbonyl oxygens on Varp-Znk2; these H-bonds comprise Tyr 163 OH (Vps29) – Cys 711 O (Varp-Znk2) and Tyr 165 OH (Vps29) – Pro 713 O (Varp-Znk2). It appears that both of these Tyrosine residues are well conserved with invariant spacing across many species, including Vps29p from *S.cerevisiae* [164]. Another very well conserved residue is Leu152 from the beta sheet which helps to form the binding cavity for Varp Znk2, as can be seen in Figure 50B. Figure 50 shows only the ordered part of Varp-Znk2, whilst the entire 61 residue protein was included in the calculations, only sequential NOESY cross peaks were detected for the N- and C-terminal tails of the protein and so the tails must both be disordered with no secondary structure as previously predicted (Figure 44 and Figure 45). These tails have been excluded from Figure 50 to make it easier to view the important binding residues. Figure 51 depicts the Varp-Znk2:Vps29 complex, with Vps29 shown in surface representation, the area in red represents the residues in Vps29 that have NOE cross peaks to residues in Varp-Znk2, the same as those shown in stick formation in Figure

50. The close up in Figure 51 clearly shows the very compact space into which Leu714 of Varp-Znk2 binds.

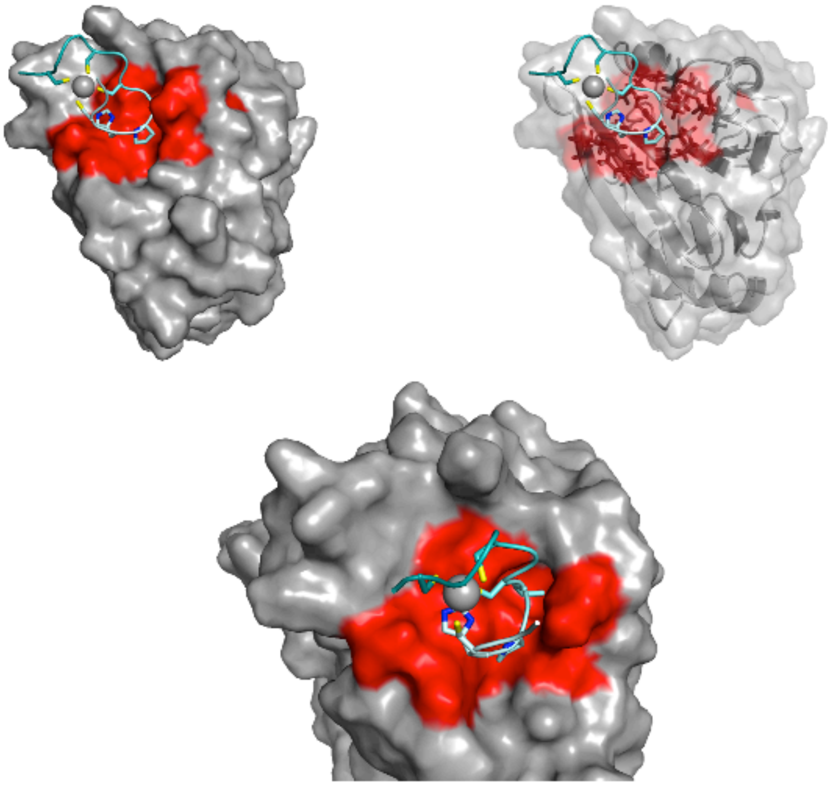


Figure 51: Surface representation of Vps29 binding to Varp-Znk2. Residues in Vps29 demonstrating NOE cross peaks to Varp-Znk2 are depicted in red. The close up shows the pocket in Vps29 that Leu714 of Varp-Znk2 fits into.

As the presented structure demonstrates, a very small number of residues from Varp are critical for creating the functionally important complex with Vps29. This suggests that Varp can be involved in a variety of interactions whilst also being bound to Vps29, and therefore the retromer core complex more generally. Additionally, as the knuckles are located within long, unstructured peptide chains, the interaction of Varp with Vps29 is unlikely to cause any steric hinderance that might interfere with other interactions of Varp. Particularly important in this respect is that the interaction with Varp should not prevent Vps29 from forming the retromer core complex by binding to Vps35. Figure 52B shows Vps29 from the calculated structures superimposed onto Vps29 in complex with Vps35. It is clearly evident that the two interaction sites are structurally separate and there should be no hinderance to Vps29 binding both Vps35 and Varp at the same time. This is important as it shows that Varp can be bound to functional retromer core complex in a cellular context, as will be discussed further in Chapter 7.

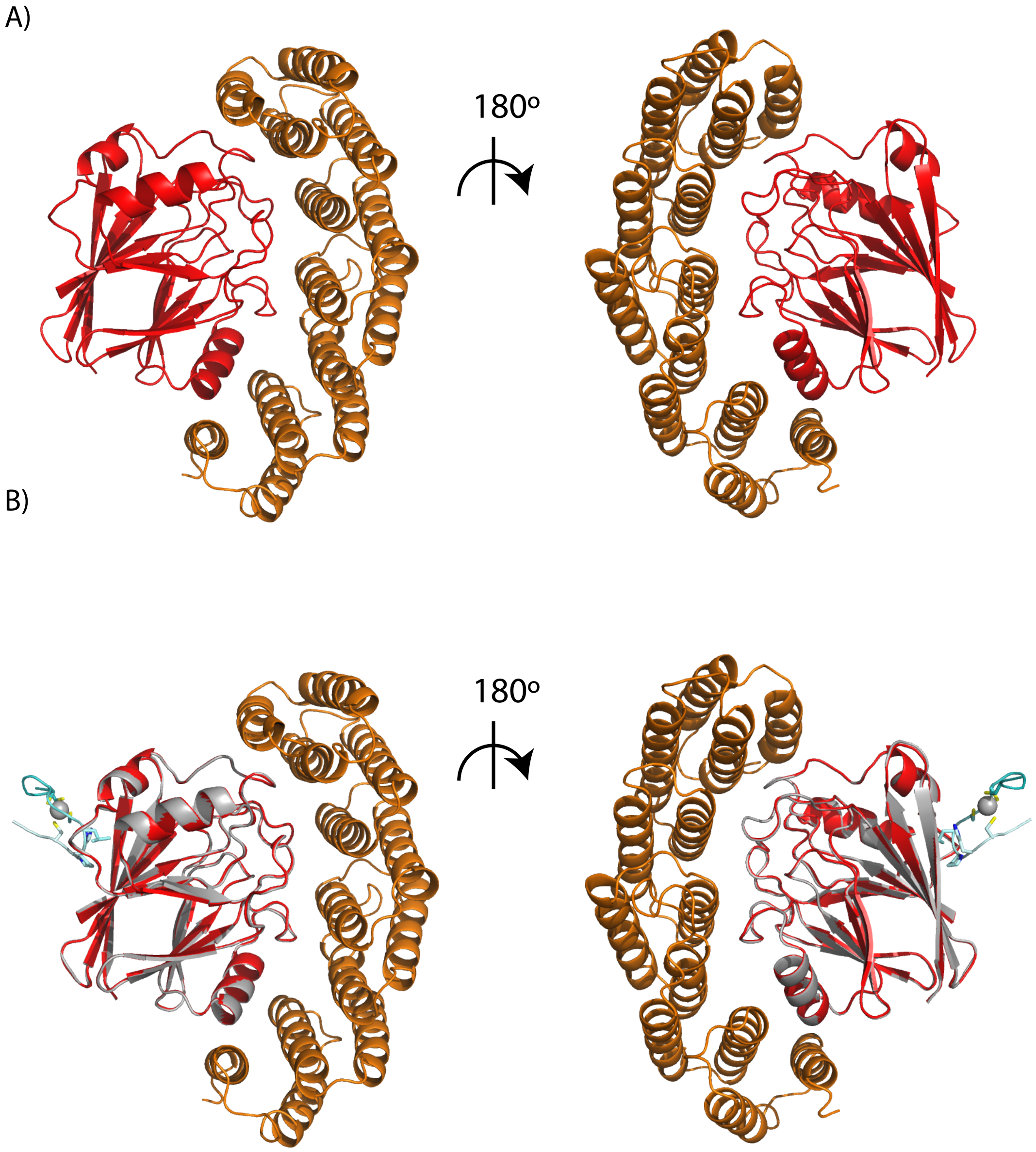


Figure 52: Structures showing Vps29 bound to Vps35, with the Varp-Znk2:Vps29 complex structure superimposed, demonstrating that it is possible for Vps29 to bind Varp-Znk2 and Vps35 at the same time. Two views of the Vps35:Vps29 complex (PDB 2R17), showing Vps35 in orange and Vps29 in red. B) Superposition of the Vps35:Vps29 and Vps29:Varp-Znk2 complexes, using N, Cα and C' of residues 2-181 of Vps29 in the two structures for fitting in PYMOL. Vps35 is shown in orange, Vps29 from PDB 2R17 in

red, Varp-Znk2 in cyan, and Vps29 from the calculated structure of Vps29:Varp Znk2 in gray.

5.5 Mutations in Vps29 alter the GST-Znk2:Vps29 interaction as detected by SPR assay.

It has been reported previously that some mutations of residues in the Vps29 protein impact its binding affinity to Varp [7]. For instance in Yeast-2-Hybrid, GST-pulldown and in vivo siRNA knockdown and rescue experiments, the L152E mutation was shown to cause a large reduction in binding to the 'cys-rich' regions of Varp (referred to as Znk1 (residues 396-460) and Znk2 (residues 692-746) in this thesis), whereas the I91S mutation of Vps29 had very little impact on the binding of Varp [7]. The non-binding mutant was fortuitously discovered by the Owen group by evaluating an unliganded X-ray crystal structure. The published experiments could not provide a quantitative value for the difference in binding affinity between the two Vps29 mutants. Using the SPR assay as detailed for the wtVps29:Znk1/2 interaction in Section 3.2, the mutant Vps29 proteins could be assessed for their Varp binding activity. The two mutant versions of Vps29 used in this experiment have been shown to be fully folded in a previous study [147]. The capture plot in Figure 53D demonstrates that GST –Znk2 binding was within 20RU for all sample injections (excepting cycle 1, which was a zero-concentration control point for I91S), and as only positive concentration measurements are used for the affinity fit, this means that no results were excluded from the affinity calculations.

The I91S and L152E mutants of Vps29 were injected at multiple concentrations, see Figure 53, with an additional 46μM injection point for L152E, and with duplicate repeats for each mutant. As demonstrated by the sensorgram in Figure 53A, there is a distinct difference between the responses of the two mutants. Whilst the injection of L152E (blue) produced very little response, showing that binding was very much reduced relative to wild-type, the response for I91S (red) binding was much higher. A K_D of 7.5μM was calculated for the I91S mutant, which, compared to the average K_D value at 25°C for wtVps29 K_D of 6.0μM (+/-1.78) can be classed as unaltered. From these results, it is clear that the I91S mutant of Vps29 had little or no impact on the binding affinity of the interaction.

The L152E mutant still displayed some binding activity to Znk2, albeit with a much-reduced response level. In this case, the calculated K_D of 25μM for the L152E Vps29:Varp Znk2 interaction is quite likely to be incorrect. From the steady state affinity fit shown in Figure 53C, it is clear that the interaction had not reached equilibrium, showing a much more linear binding curve than other experiments. Biacore recommends that concentration points of at least 2X K_D should be measured to accurately calculate the K_D of an interaction. In this case, with the calculated K_D of 25μM, the 46μM data point represents only 1.84x K_D (see Table 6). It is likely that repeating the experiment with higher concentration values of the L152E mutant, and ensuring the interaction reaches equilibrium, would lead to a much higher K_D value (weaker binding) for the interaction.

For a final table comparing the calculated K_D results for wtVps29 and mutant versions of Vps29 with Znk2, see Table 7.

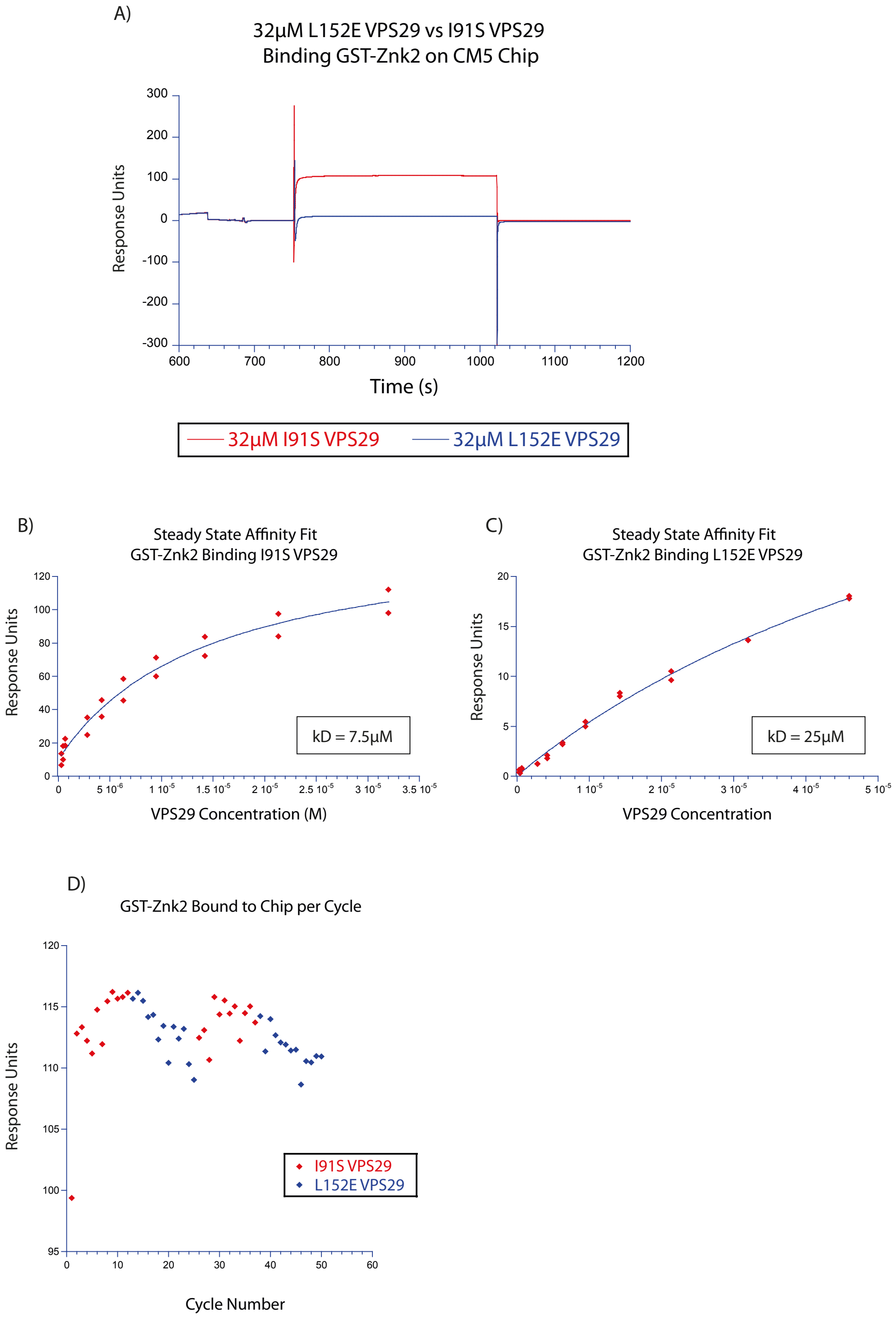


Figure 53: An SPR experiment showing mutant Vps29 binding to Znk2 at 25°C. A sensorgram depicting the 32µM analyte injection phase (phase z) for both I91S and L152E mutants at 25°C. B) Steady state affinity fit for I91S. C) Steady state affinity fit for L152E. D) a Capture level plot demonstrating the level of GST-Znk2 bound to the sensor surface for all cycles, including I91S and L152E experiments. All are within 20RU difference, apart from cycle 1, which was a zero control point for the experiment and thus not required for the steady state affinity fit. AS for Figures 26-28, it should be noted that in this and subsequent capture level plots that the Y-axis ranges does not begins at zero, and is different in each case so as to correspond to the data range.

Concentration of analyte (μM)	x kD I91S	x kD L152E
0.00	0.00	0.00
0.29	0.04 x	0.01 x
0.44	0.06 x	0.02 x
0.67	0.09 x	0.03 x
1.00	0.13 x	0.04 x
2.81	0.37 x	0.11 x
4.00	0.50 x	0.16 x
4.21	0.56 x	0.17 x
6.32	0.84 x	0.25 x
8.00	1.10 x	0.32 x
9.48	1.26 x	0.38 x
14.22	1.90 x	0.57 x
21.33	2.84 x	0.85 x
24.00	3.20 x	0.96 x
32.00	4.30 x	1.40 x
46.00	- - -	1.84 x

Table 6: A table of the concentrations of mutant Vps29 used, expressed as a multiplier of the final calculated K_D values for each mutant. This is a useful measure of whether the calculated K_D is likely to be accurate, as Biacore recommends using a concentration range representing up to **at least** 2x the K_D of the interaction.

Analyte	Ligand	kD
wtVPS29	GST-Znk2	6.00µM (+/- 1.78)
L152E VPS29	GST-Znk2	25µM
I91S VPS29	GST-Znk2	7.5µM

Table 7: A table of calculated K_D values for the interaction of either wtVps29, L152E Vps29 or I91S Vps29 with GST-Znk2, all at 25°C.

5.6 Additional SPR Experiments analyse the Vps29 : Znk2 interaction further.

Nathan Zaccai of the Owen lab (CIMR) conducted additional SPR experiments with Vps29 and Znk2 to gather further information about the interaction and help finalise the data for later publication. In line with the previously calculated affinity of 6.00μM (+/-1.78μM) between WT Vps29 and WT GST-Znk2 WT, in Zaccai's experiments WT Vps29 was found to bind to WT GST-Znk2 with an affinity of 4.4μM +/- 0.8μM at 25°C.

5.6.1 EDTA treatment and Vps29 mutation both reduce the apparent binding affinity of the interaction.

Zaccai also repeated SPR experiments designed to perturb the interaction. For instance, the binding interaction between Vps29 and GST-Znk2 in the presence of EDTA was re-attempted. The apparent K_D of the interaction was calculated to be 45.5μM (+/-12μM), as shown in Figure 54A. It should be noted that this study was undertaken using 'GST-CLA' as a binding reference, instead of GST as used in the previous experiments. It was felt that the reference protein should be somewhat more similar in size to GST-Znk2; the GST-CLA protein consists of GST linked to residues 623-632 of the β2 adaptin subunit of human AP-2 (UniProt P63010) and does not bind Vps29 specifically. The calculated K_D of 45.5μM in this instance is lower than the previously calculated K_D of 86μM. The reason for this discrepancy is unclear, however, the previous experiment was conducted at 25°C, whilst Zaccai repeated the experiment at 12°C. Indeed, thermodynamic experiments presented in Section 3.2.3 demonstrate that the temperature does impact the calculated K_D, (albeit minimally). Not only this, but the use of a different binding reference protein may have impacted the apparent K_D. If there was non-specific binding to the GST alone reference surface in the previous experiment, then subtracting this from the active surface data would have resulted in an artifactually high value of K_D for the interaction in that case.

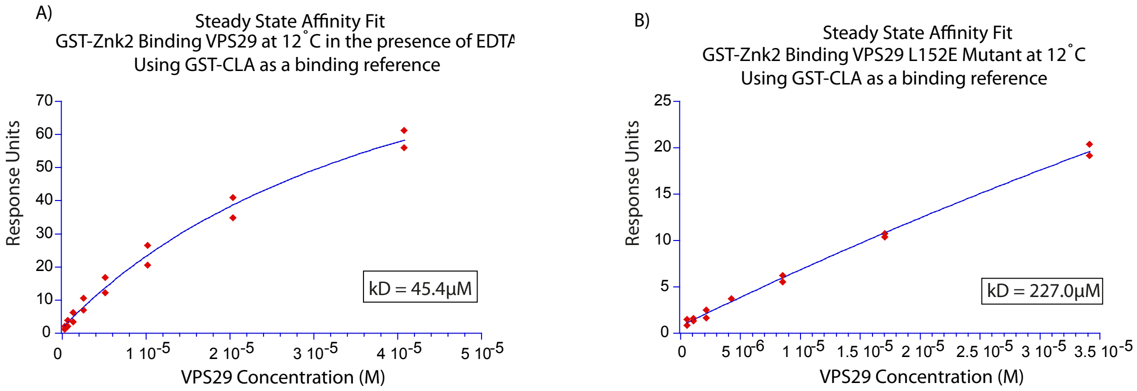


Figure 54:Steady state affinity fits of repeated SPR experiments. A) Steady state affinity fit of GST-Znk2 binding Vps29 in the presence of 10mM EDTA, at 12°C. B) Steady state affinity fit of GST-Znk2 binding the L152E Vps29 mutant at 12°C.

Additionally, Nathan Zaccai determined that the L152E Vps29 mutant bound to GST-Znk2 with a K_D of 227μM (+/- 4.9μM) at 12°C, Figure 54B. This result is in stark contrast to the earlier reported result of 25μM (at 25°C). Unfortunately, neither of the experiments conducted with L152E Vps29 used a protein concentration greater than 48μM, which means that the binding data collected stopped well short of saturation in both cases. This experiment will need to be repeated again, using a much higher range of concentrations to produce a more accurate result.

5.6.2 Structure directed mutants of Vps29 show a reduced binding affinity compared to WT.

Once the structure of the Vps29:Znk2 structure was solved, Vps29 mutants were designed in an attempt to disrupt the interaction site. Several mutants were designed, including L26S, Y165S, F150V and F150A. Unfortunately, the F150A and F150V Vps29 mutants were unable to be purified, and whilst protein expression was achieved in *E.coli*, upon attempting to purify the proteins, it became apparent that the recombinant proteins were insoluble and precipitated into the cell pellet, and were not soluble in the cell supernatant fraction.

In contrast, both L26S and Y165S mutants were expressed in *E.coli* , successfully purified, and used in subsequent SPR experiments. The L26S and Y165S mutants were shown to bind to GST-Znk2 with affinities in the micromolar range. The L26S Vps29 mutant bound GST-Znk2 with an apparent K_D of 22.3μM, and the Y165S mutant bound GST-Znk2 with an apparent K_D of 46.7μM, as shown in Figure 55. These experiments were undertaken with GST as a binding reference, and interestingly, they display affinities in a similar range to that of the L152E Vps29 mutant K_D of 25μM determined previously, when GST was used as a binding reference.

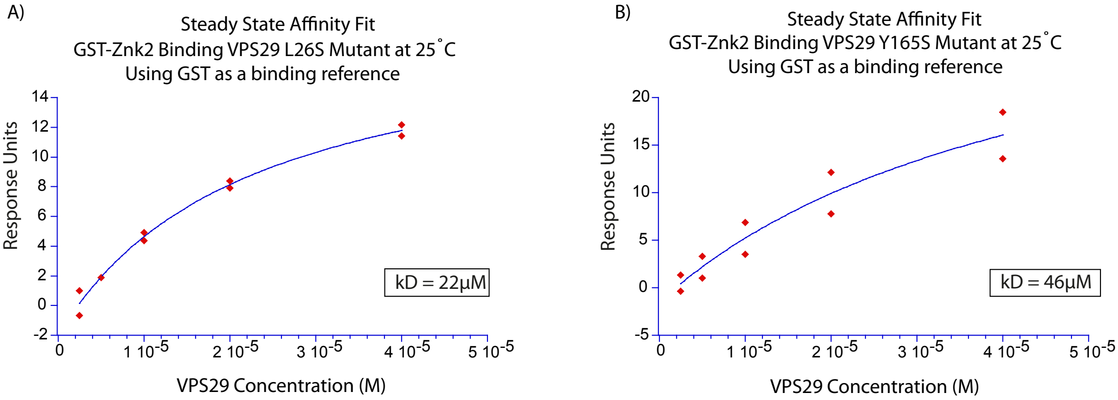


Figure 55: Steady state affinity fits of mutant Vps29:WT GST-Znk2 SPR experiments. A) Steady state affinity fit for GST-Znk2 binding L26S Vps29 mutant, at 25°C. B) Steady state affinity fit for GST-Znk2 binding Y165S Vps29 mutant, at 25°C.

5.6.3 Mutations in Varp Znk2 abolish interaction with Vps29 as measured by SPR.

Zaccai carried out further SPR experiments using mutants of Znk2, with either the Histidine and Leucine of the central triplet sequence (HPL) mutated to Alanines or the 4 Cysteines mutated to Serines. These mutants were coupled to the CM5 chip using the anti-GST antibody method as described for WT GST-Znk2 in Sections 3.2.1 and 3.2.2. In this case however, GST-CLA was used as a binding reference.

The 4C mutant of GST-Znk2 showed a reduced binding affinity, compared to WT, of 54.5μM (+/- 1.9μM) whilst the K_D of the interaction between the HL mutant of GST-Znk2 and WT Vps29 could not be determined, shown in Figure 56. Although the steady state affinity fit did produce a curve (data not shown), the sensorgram data revealed that very little Vps29 was binding to the ligand surface and therefore that any K_D calculation should be considered unreliable.

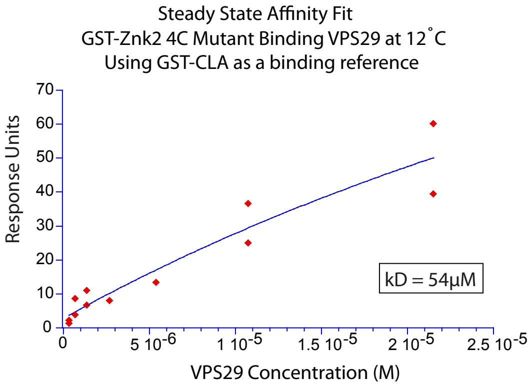


Figure 56: Steady state affinity fit for GST-Znk2 4C mutant binding WT Vps29, at 12°C.

The HL mutant version of Znk2 can be considered to have no binding to wtVps29 and this is consistent with what is known about the interaction between Vps29 and Znk2 from the structure. Whilst the HPL motif directly interacts with Vps29, the cysteine residues co-ordinate zinc to create a loop structure. It is therefore unsurprising that the GST-Znk2 4C mutant had a smaller impact on the binding affinity than the HL mutant. The cysteines are not directly involved in binding to Vps29, but do create the structure of the loop, which most likely exists to present the HPL motif in an optimal configuration for binding to Vps29. Interestingly, the binding of Vps29 to ETDA treated GST-Znk2 was similar to that of the GST-Znk2 4C mutant (45.4μM vs 54.5μM respectively), which suggests that the EDTA treatment did effectively remove the zinc and destroy the loop structure in these experiments.

5.7 Structure-directed mutants affect interaction in vivo.

Lena Wartosch (formerly CIMR) expressed several structure-directed mutants of Vps29 in vivo to investigate whether the interaction of Varp and Vps29 would be affected in a cellular context. The mutants used in this case include: L26A L26S, F150V, F150A, L152E, Y165S, V174D. As shown in Figure 57, these are some of the key residues in Vps29 that mediate the interaction and one would expect that the mutation of any of these residues to a sufficiently different residue would severely reduce the interaction.

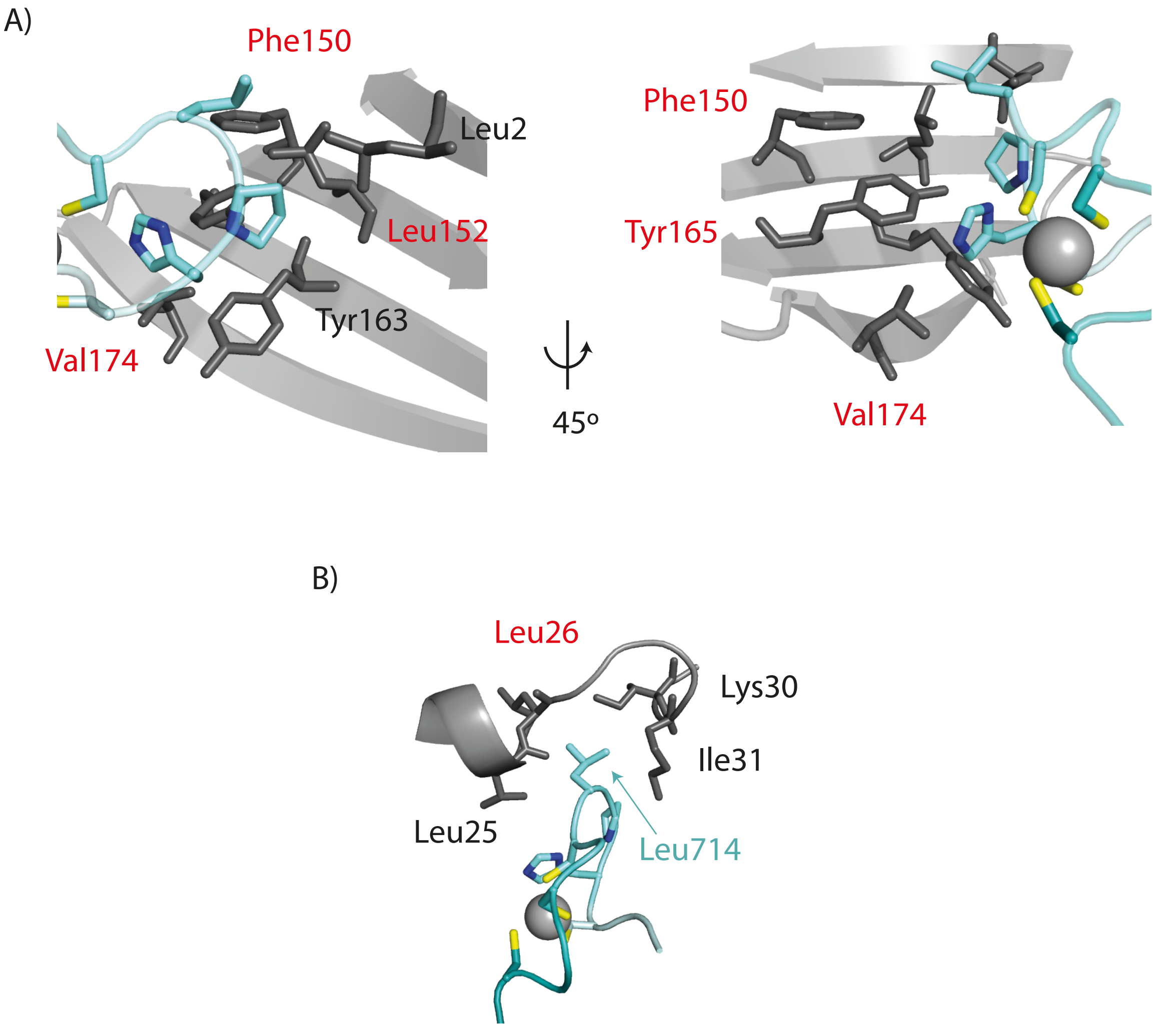


Figure 57: A structure diagram showing mutated residues labelled in red. A) A close up of the Varp binding residues in the beta sheet of Vps29. The residues in the beta sheet of Vps29 that were mutated are labelled in red. B) A close up of the Varp binding residues in a loop in Vps29, with Leu 26 labelled in red.

Figure 58 depicts the effect of several Vps29 mutants on the co-localisation of Varp-GFP and Vps29-tagRFP, in HeLa cells. From Figure 58, it is clear that the well-established loss-of-binding Vps29 mutant, L152E, results in a loss of Vps29:Varp co-localisation, which is very evident in saponin treated (cytosol extracted) cells. No Varp remains bound to membranous structures, and none is found colocalized with L152E Vps29, which is in contrast to the WtVps29-tagRFP-containing cells. Two of the structure-directed mutants, L26S and Y165S, also display this loss of co-localisation. Interestingly, the other structure directed mutations do not seem to have a large impact on the co-localisation between Varp and Vps29. Perhaps in the case of the F150A/V mutants, it is not altogether surprising that the mutations have very little effect on the interaction as the structure shown in Figure 57 shows that F150 is significantly more distant from the interaction site than either L152 or L26. This also suggests that F150 is not critically involved in structuring the binding site. As a control, Wartosch also demonstrated that for all the mutants used in the *in vivo* experiments, the co-localisation with Vps35 was not affected (data not shown).

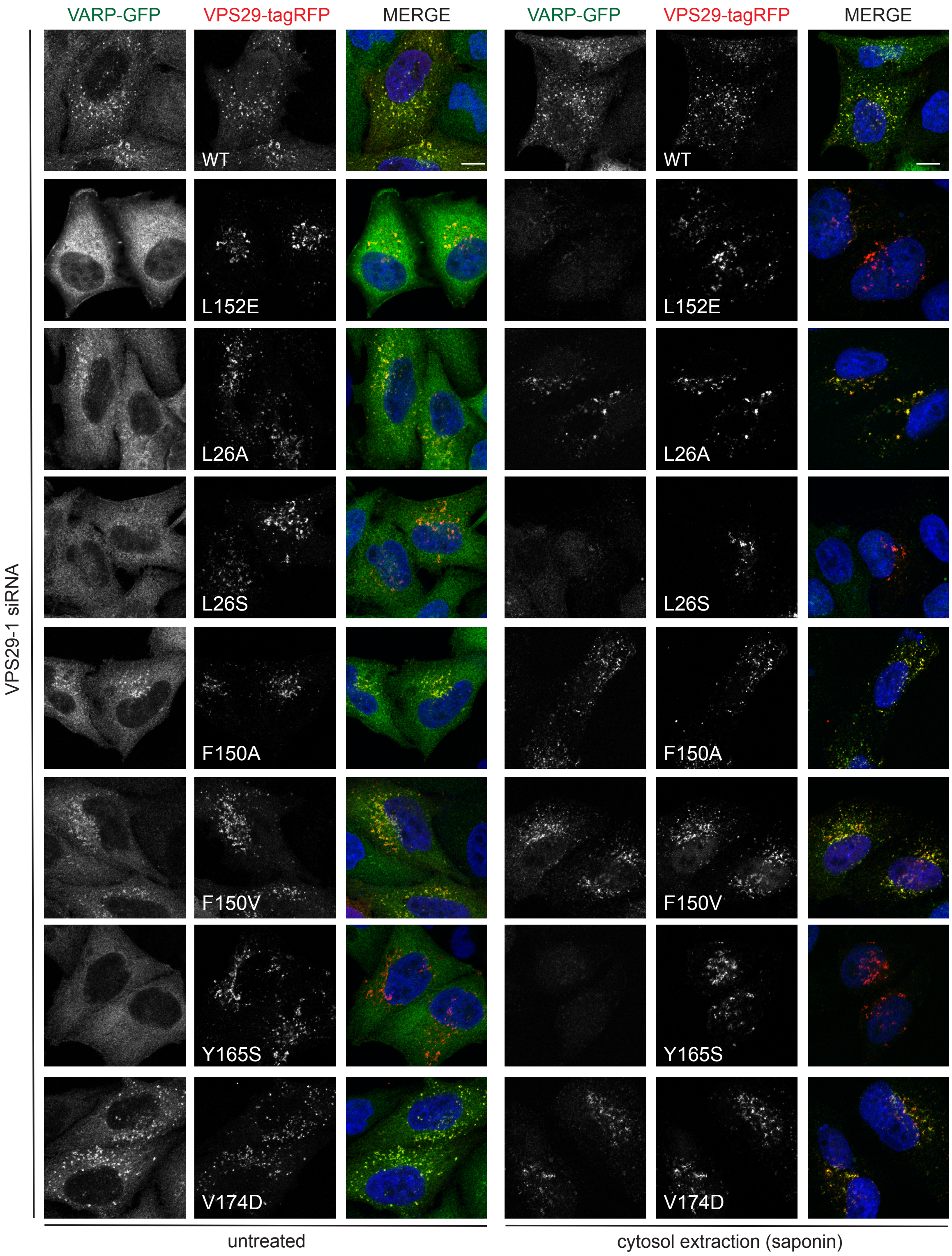


Figure 58: Immunofluorescence confocal Microscopy images of Vps29-siRNA treated HeLa cells.Varp-GFP in left-hand panels, various Vps29-tagRFP constructs in middle panels, with the two merged in the right panels. The three left-hand columns represent untreated (cytosol intact) cells, whilst the three right-hand columns represent saponin treated (cytosol extracted) cells.

Chapter Six

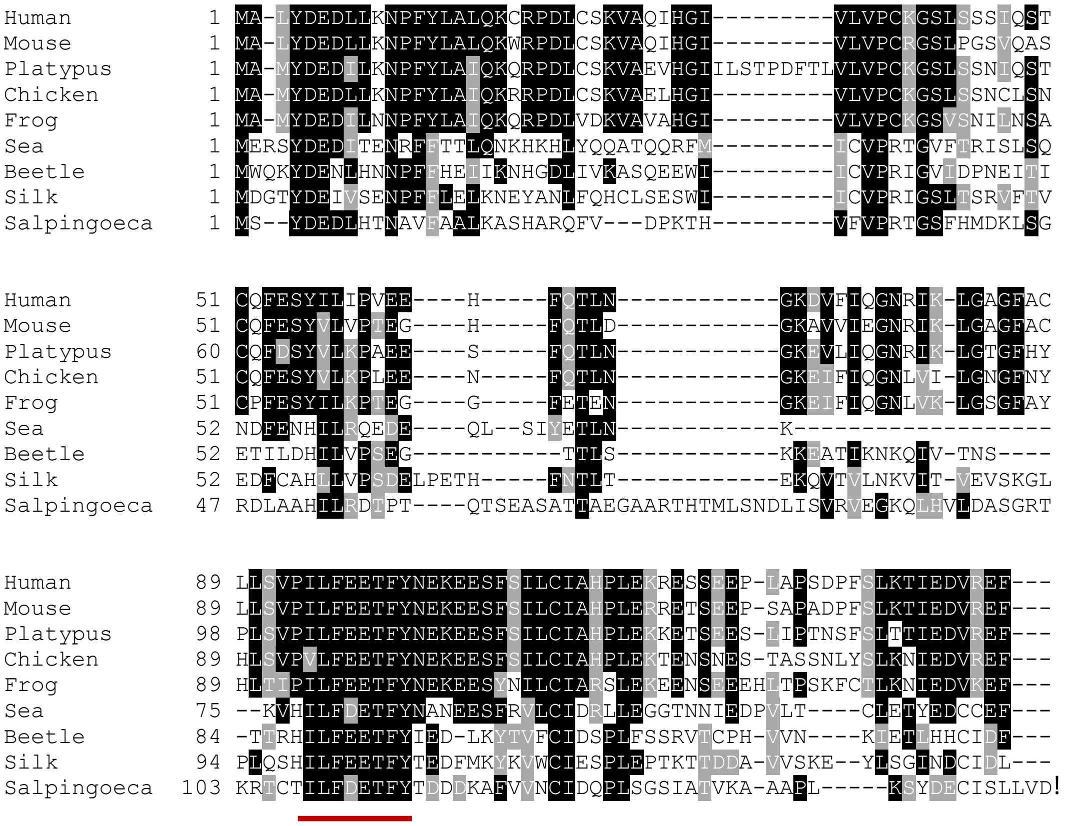
6 Investigation of hVarp N-terminal domain

In addition to the main part of the project that concentrated on the zinc-knuckle sequences of Varp, some exploratory work was carried out on the N-terminal region (1-136aa). Varp has been reported to have several other binding partners (See Chapter 1, Section 1.4.1), but many of these interactions have not been localised to a specific area of Varp. As most domains of Varp are fairly well studied, the N-terminal region provides an as-yet unstudied region that could be a binding site for other important binding partners of Varp or be required for another biological function.

6.1 The N-terminal domain of hVarp is highly conserved

A multiple sequence alignment of the N-terminal region was performed using a variety of vertebrates and invertebrates, using a similar panel of species as in [7]. It is clear to see from the sequence alignment in Figure 59 that there are large sections of very well conserved sequence, suggesting the possibility that this region has some functional significance.

The N-terminus of hVarp has some sequence similarity to the newly described Varp-like protein Vrl1 from *Saccromyces Cerevisiae*, but Vrl1 is not as well conserved as those from other species, as may be seen in Figure 60. Most notably, Vrl1 does not have the otherwise widely conserved 'ILFEETFY' motif, although the Ile, Leu and Tyr of this motif are absolutely maintained with identical spacing.

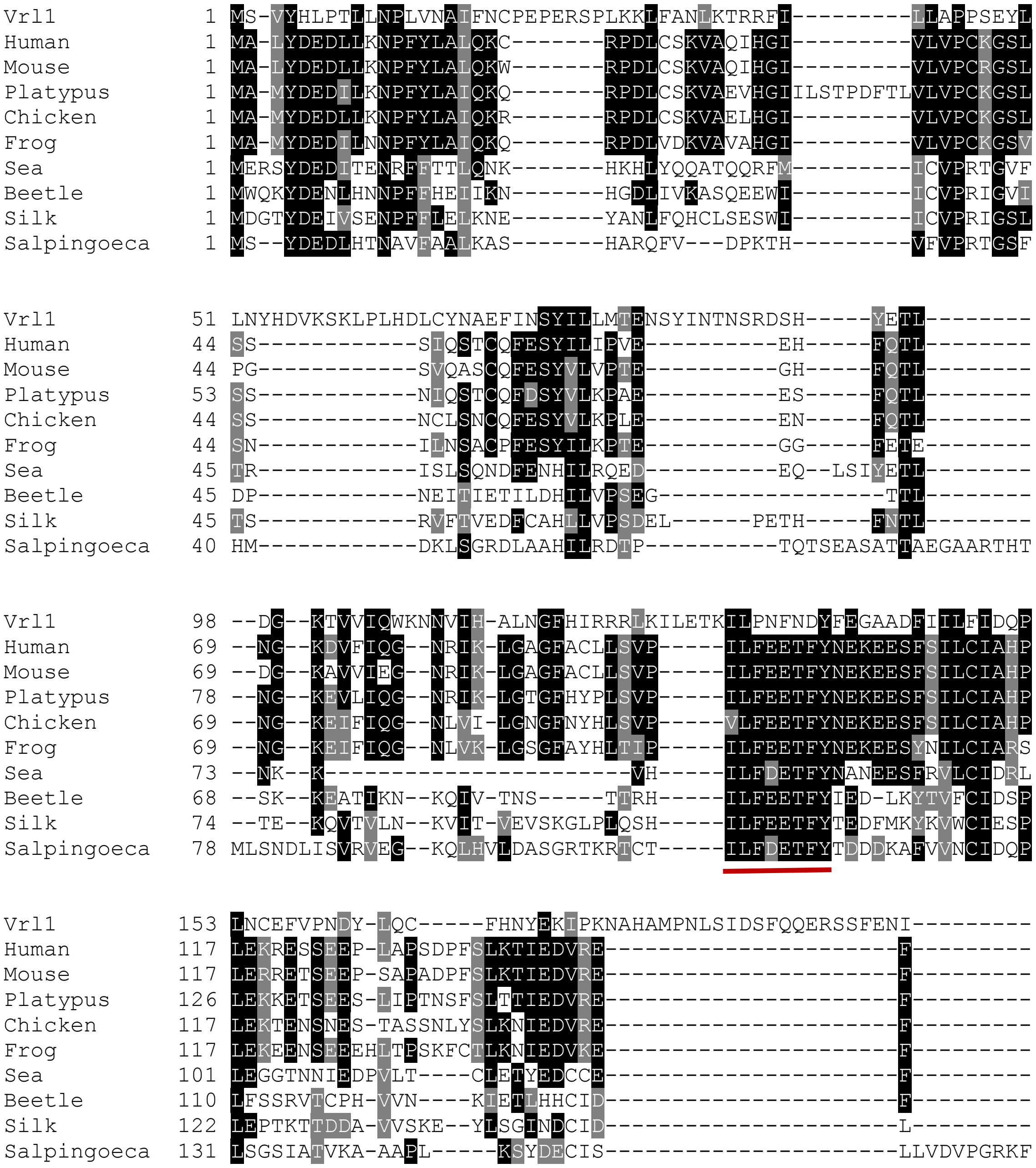


```
Human           1 MA-LYDEDLLKNPFYLALQKCRPDLCSKVAQIHGI---------VLVPCKGSLSSSIQST
Mouse           1 MA-LYDEDLLKNPFYLALQKWRPDLCSKVAQIHGI---------VLVPCRGSLPGSVQAS
Platypus        1 MA-MYDEDILKNPFYLAIQKQRPDLCSKVAEVHGIILSTPDFTLVLVPCKGSLSSNIQST
Chicken         1 MA-MYDEDLLKNPFYLAIQKRRPDLCSKVAELHGI---------VLVPCKGSLSSNCLSN
Frog            1 MA-MYDEDILNNPFYLAIQKQRPDLVDKVAVAHGI---------VLVPCKGSVSNILNSA
Sea             1 MERSYDEDITENRFFTTLQNKHKHLYQQATQQRFM---------ICVPRTGVFTRISLSQ
Beetle          1 MWQKYDENLHNNPFFHEIIKNHGDLIVKASQEEWI---------ICVPRIGVIDPNEITI
Silk            1 MDGTYDEIVSENPFFLELKNEYANLFQHCLSESWI---------ICVPRIGSLTSRVFTV
Salpingoeca     1 MS--YDEDLHTNAVFAALKASHARQFV---DPKTH---------VFVPRTGSFHMDKLSG

Human          51 CQFESYILIPVEE----H-----FQTLN------------GKDVFIQGNRIK-LGAGFAC
Mouse          51 CQFESYVLVPTEG----H-----FQTLD------------GKAVVIEGNRIK-LGAGFAC
Platypus       60 CQFDSYVLKPAEE----S-----FQTLN------------GKEVLIQGNRIK-LGTGFHY
Chicken        51 CQFESYVLKPLEE----N-----FQTLN------------GKEIFIQGNLVI-LGNGFNY
Frog           51 CPFESYILKPTEG----G-----FETEN------------GKEIFIQGNLVK-LGSGFAY
Sea            52 NDFENHILRQEDE----QL--SIYETLN------------K-------------------
Beetle         52 ETILDHILVPSEG-----------TTLS------------KKEATIKNKQIV-TNS----
Silk           52 EDFCAHLLVPSDELPETH-----FNTLT------------EKQVTVLNKVIT-VEVSKGL
Salpingoeca    47 RDLAAHILRDTPT----QTSEASATTAEGAARTHTMLSNDLISVRVEGKQLHVLDASGRT

Human          89 LLSVPILFEETFYNEKEESFSILCIAHPLEKRESSEEP-LAPSDPFSLKTIEDVREF---
Mouse          89 LLSVPILFEETFYNEKEESFSILCIAHPLERRETSEEP-SAPADPFSLKTIEDVREF---
Platypus       98 PLSVPILFEETFYNEKEESFSILCIAHPLEKKETSEES-LIPTNSFSLTTIEDVREF---
Chicken        89 HLSVPVLFEETFYNEKEESFSILCIAHPLEKTENSNES-TASSNLYSLKNIEDVREF---
Frog           89 HLTIPILFEETFYNEKEESYNILCIARSLEKEENSEEEHLTPSKFCTLKNIEDVKEF---
Sea            75 --KVHILFDETFYNANEESFRVLCIDRLLEGGTNNIEDPVLT----CLETYEDCCEF---
Beetle         84 -TTRHILFEETFYIED-LKYTVFCIDSPLFSSRVTCPH-VVN----KIETLHHCIDF---
Silk           94 PLQSHILFEETFYTEDFMKYKVWCIESPLEPTKTTDDA-VVSKE--YLSGINDCIDL---
Salpingoeca   103 KRTCTILFDETFYTDDDKAFVVNCIDQPLSGSIATVKA-AAPL-----KSYDECISLLVD!
```

Figure 59: Multiple sequence alignment of hVarp residues 1-136 generated using T-Coffee, and formatted using BoxShade. Black shading represents a widely conserved amino acid, whilst gray shading indicates amino acid similarity within a defined group. See Section 2.1.1 for details.



```
Vrl1           1 MS-VYHLPTLLNPLVNAIFNCPEPERSPLKKLFANLKTRRFI---------LLAPPSEYL
Human          1 MA-LYDEDLLKNPFYLALQKC-------RPDLCSKVAQIHGI---------VLVPCKGSL
Mouse          1 MA-LYDEDLLKNPFYLALQKW-------RPDLCSKVAQIHGI---------VLVPCRGSL
Platypus       1 MA-MYDEDILKNPFYLAIQKQ-------RPDLCSKVAEVHGIILSTPDFTLVLVPCKGSL
Chicken        1 MA-MYDEDLLKNPFYLAIQKR-------RPDLCSKVAELHGI---------VLVPCKGSL
Frog           1 MA-MYDEDILNNPFYLAIQKQ-------RPDLVDKVAVAHGI---------VLVPCKGSV
Sea            1 MERSYDEDITENRFFTTLQNK-------HKHLYQQATQQRFM---------ICVPRTGVF
Beetle         1 MWQKYDENLHNNPFFHEIIKN-------HGDLIVKASQEEWI---------ICVPRIGVI
Silk           1 MDGTYDEIVSENPFFLELKNE-------YANLFQHCLSESWI---------ICVPRIGSL
Salpingoeca    1 MS--YDEDLHTNAVFAALKAS-------HARQFV---DPKTH---------VFVPRTGSF

Vrl1          51 LNYHDVKSKLPLHDLCYNAEFINSYILLMTENSYINTNSRDSH-----YETL--------
Human         44 SS------------SIQSTCQFESYILIPVE----------EH-----FQTL--------
Mouse         44 PG------------SVQASCQFESYVLVPTE----------GH-----FQTL--------
Platypus      53 SS------------NIQSTCQFDSYVLKPAE----------ES-----FQTL--------
Chicken       44 SS------------NCLSNCQFESYVLKPLE----------EN-----FQTL--------
Frog          44 SN------------ILNSACPFESYILKPTE----------GG-----FETE--------
Sea           45 TR------------ISLSQNDFENHILRQED----------EQ--LSIYETL--------
Beetle        45 DP------------NEITIETILDHILVPSEG----------------TTL--------
Silk          45 TS------------RVFTVEDFCAHLLVPSDEL------PETH-----FNTL--------
Salpingoeca   40 HM------------DKLSGRDLAAHILRDTP----------TQTSEASATTAEGAARTHT

Vrl1          98 --DG--KTVVIQWKNNVIH-ALNGFHIRRRLKILETKILPNFNDYFEGAADFIILFIDQP
Human         69 --NG--KDVFIQG--NRIK-LGAGFACLLSVP-----ILFEETFYNEKEESFSILCIAHP
Mouse         69 --DG--KAVVIEG--NRIK-LGAGFACLLSVP-----ILFEETFYNEKEESFSILCIAHP
Platypus      78 --NG--KEVLIQG--NRIK-LGTGFHYPLSVP-----ILFEETFYNEKEESFSILCIAHP
Chicken       69 --NG--KEIFIQG--NLVI-LGNGFNYHLSVP-----VLFEETFYNEKEESFSILCIAHP
Frog          69 --NG--KEIFIQG--NLVK-LGSGFAYHLTIP-----ILFEETFYNEKEESYNILCIARS
Sea           73 --NK--K-------------------------VH-----ILFDETFYNANEESFRVLCIDRL
Beetle        68 --SK--KEATIKN--KQIV-TNS-----TTRH-----ILFEETFYIED-LKYTVFCIDSP
Silk          74 --TE--KQVTVLN--KVIT-VEVSKGLPLQSH-----ILFEETFYTEDFMKYKVWCIESP
Salpingoeca   78 MLSNDLISVRVEG--KQLHVLDASGRTKRTCT-----ILFDETFYTDDDKAFVVNCIDQP

Vrl1         153 LNCEFVPNDY-LQC-----FHNYEKIPKNAHAMPNLSIDSFQQERSSFENI---------
Human        117 LEKRESSEEP-LAPSDPFSLKTIEDVRE----------------------F---------
Mouse        117 LERRETSEEP-SAPADPFSLKTIEDVRE----------------------F---------
Platypus     126 LEKKETSEES-LIPTNSFSLTTIEDVRE----------------------F---------
Chicken      117 LEKTENSNES-TASSNLYSLKNIEDVRE----------------------F---------
Frog         117 LEKEENSEEEHLTPSKFCTLKNIEDVKE----------------------F---------
Sea          101 LEGGTNNIEDPVLT----CLETYEDCCE----------------------F---------
Beetle       110 LFSSRVTCPH-VVN----KIETLHHCID----------------------F---------
Silk         122 LEPTKTTDDA-VVSKE--YLSGINDCID----------------------L---------
Salpingoeca  131 LSGSIATVKA-AAPL-----KSYDECIS----------------------LLVDVPGRKF
```

Figure 60: Sequence alignment of hVarp 1-136 including the same sequences as in Figure 59, but now also including Vrl1. Black shading represents a widely conserved amino acid, whilst gray shading indicates amino acid similarity within a defined group. See Section 2.1.1 for details.

One can find further similarities between corresponding positions in ‘ILFEETFY’ from hVarp and ‘ILPNFNDY’ in Vrl1; both Proline and Phenylalanine are hydrophobic and bulky while Asparagine and Glutamic acid are both polar amino acids.

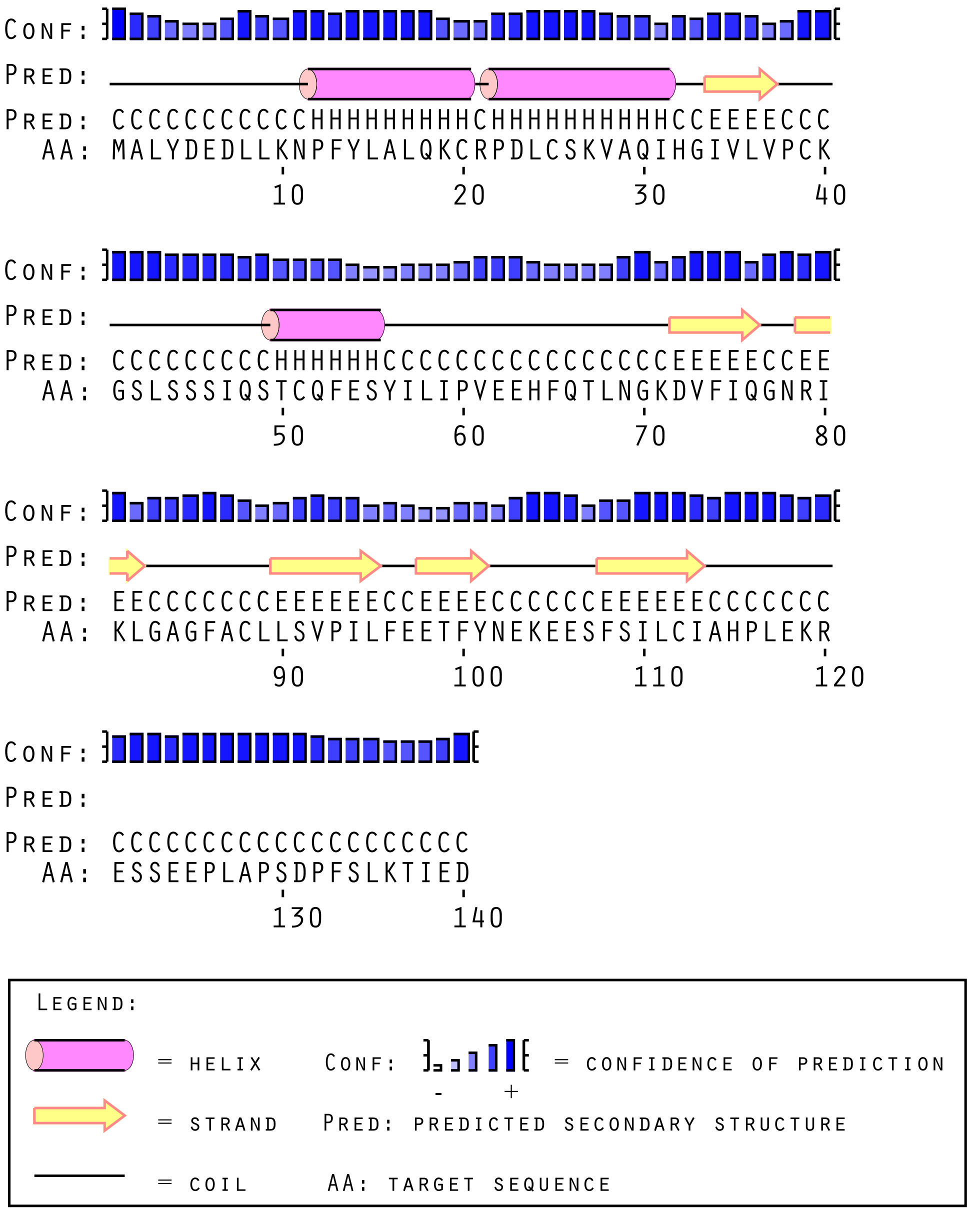


Figure 61: A prediction of the likely structure formed by the N-terminus of hVarp, as predicted by PsiPred [238].

From the structure prediction in Figure 61, it is apparent that the N-terminal region is quite likely to have significant areas of secondary structure and is unlikely to be an unstructured N-terminal extension. It could therefore be the case that the sequence of the N-terminal region is conserved to ensure the fidelity of this structure and possibly the overall structure of Varp more generally. However, in many cases even the tertiary structure of a protein can be retained without much sequence conservation.

It is possible that the conserved N-terminal region of Varp is responsible for mediating interactions with other proteins. Varp has been reported in the literature to bind to both golginA4 (aka golgin245) and the microtubule motor Kif5A [6], however to date, these interactions have not been mapped to a specific region of Varp. Not only this, but there may be as-yet undiscovered proteins that interact with Varp, perhaps other trafficking cargo or regulatory proteins.

6.2 Identifying protein interactions of the conserved N-terminus using a Yeast 2 Hybrid screening approach.

To attempt to elucidate the function of the N-terminal region of hVarp, a Yeast-2-Hybrid screen was undertaken to determine if there were any binding partners in the human proteome. The yeast diploids were created as described in Chapter 2, Section 2.8.

After three days of incubation on a quadruple dropout (QDO) plate selective for two of four reporter genes, 761 single colonies had grown. Of these, 300 colonies were picked as a representative sample and were further tested. The following rounds of selection enabled the exclusion of colonies that did not activate all four GAL4 promoter genes, or show a PCR product when screened with pGADT7 primers. Therefore, 148 colonies that grew but which were not blue on the QDO/X/A plate (meaning that the MEL1 gene for X-a-Galactose metabolism has not been activated by GAL4 BD/AD interaction) were excluded. Of the remaining 152 colonies that grew and in addition were blue (for an example of positive colonies arising on a QDO/X/A plate, see Figure 62a), 66 colonies did not produce a PCR product when screened with the pGADT7 primers and 28 colonies produced multiple PCR products; these colonies were therefore all excluded from further analysis.

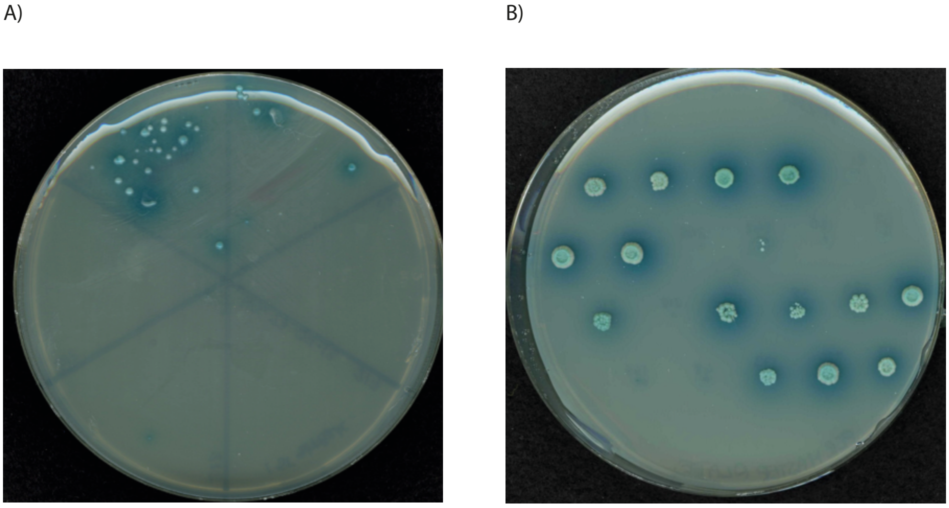


Figure 62: Images of plates from a Y2H screen. An example of a 9cm QDO/X/A screening plate with colonies re-streaked from a screening QDO plate. Only two of the re-streaked colonies (in triangular sections) grew. B) An example of a 9cm QDO/X/A PCR master plate on which some colonies continued to successfully grow, whereas others did not.

After the exclusion of the above colonies, 58 colonies were left that produced single bands on a SYBR™ Safe stained agarose gel on PCR and were sent for sequencing with pGADT7 primers. Of these, 44 colonies produced good quality sequencing data that provided information on the human partial gene fragments that were supplied cloned in frame with the GAL4-AD.

Using the NCBI Blastx function [235] at https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastx&PAGE_TYPE=BlastSearch (accessed 12/2013), the nucleotide sequences were searched as a translated protein sequence against a database of known protein sequences. Using this approach, a total of 18 individual proteins were detected in the screen with 8 proteins being detected in more than one colony in the screen, see Table 8.

Protein	No. of Hits	Function
AP-1 mu1-B	5	Clathrin coat adaptor protein involved in TGN to endosome transport
UB conj. E	5	Attaches Ubiquitin to lysine residues on proteins for degradation
TALPID3	4	Cillia formation and Hedgehog signalling
M3K19 iso 1	4	Phosphorylates MAP kinase kinase
K+/Na+ Beta 3	4	Subunit of larger protein that maintains Na+ and K+ gradients
Kininogen	3	Inhibits thiol protease activity, involved in coagulation
Ribosomal L31	2	Forms part of the 60S subunit of ribosomes
Rho DI2	2	Inhibits the dissociation of GDP from Rho proteins

Table 8: A table of positive hits that were found more than once in the Y2H screen performed on the N-terminal domain of hVarp.

All these results would need further investigation to verify whether the indicated interactions are actually real, since Yeast-2-Hybrid screens are not absolute indicators of true interactions. Many false positives can be picked up by this system due to non-specific interactions, so further evidence from different techniques is usually required to validate any potential hits. False negatives are also an issue with this technique. Certainly, a broad Y2H screen may not pick up all the possible interaction partners for a certain protein. For example, in a normalized human library (as used in this instance), copy numbers of certain specialized proteins may be very low and therefore one is much less likely to detect interactions with specialized proteins. This situation can be improved by selecting alternative libraries, for instance it could be that using a melanocyte library would be more suitable in this case, as Varp is known to have a role in transporting the TYRP1 protein to melanosomes [4]. Also, membrane proteins are inherently difficult to detect with this method and considering the known functions of Varp, this could be a serious limitation in this instance. Other potential issues include steric hindrance and transient interactions.

6.3 Independent confirmation of a positive Y2H interaction.

In an attempt to confirm the newly discovered interaction between Varp and Talpid3, Sally Gray and Paul Luzio (CIMR, Cambridge) undertook directed Y2H screens replating the relevant fragments of the two molecules, 1-136 of Varp and 1454-1644 of Talpid3, as determined by the initial Yeast 2 Hybrid screen.

In a first experiment, using the same type of Y2H screen as had been used to discover the interaction initially (Mate and Plate™), they were able to demonstrate that the interaction between Varp and Talpid3 results in strong growth on QDO/X/A media, as shown in Figure 63. There is some slight growth of the Talpid3/empty vector (pGBKT7) diploid also, but this is far less than diploids containing either a Varp 1-136 or Varp 1-117 plasmid.

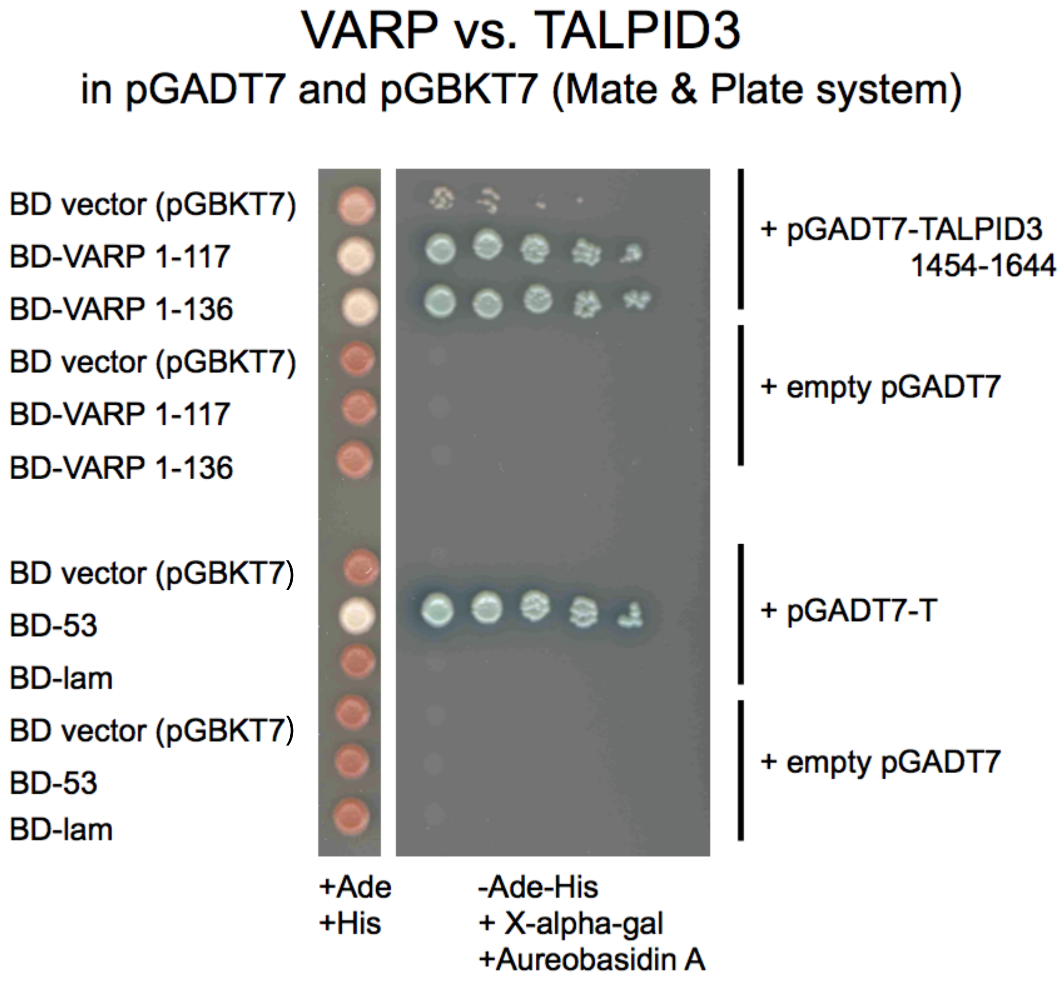


Figure 63: Growth of colonies in directed Y2H screen undertaken by Sally Gray. The colonies represent 3 days growth at 30°C. BD-53 (in pGBKT7) and pGADT7-T represent a positive control in this system.

In further work, Luzio and Gray showed that the conserved 'ILFEETFY' motif, identified as a result of the previous sequence conservation analysis, is responsible for mediating this interaction. By mutating LFEETFY to AASSSAS they completely abolished the interaction between Talpid3 and Varp, see Figure 64.

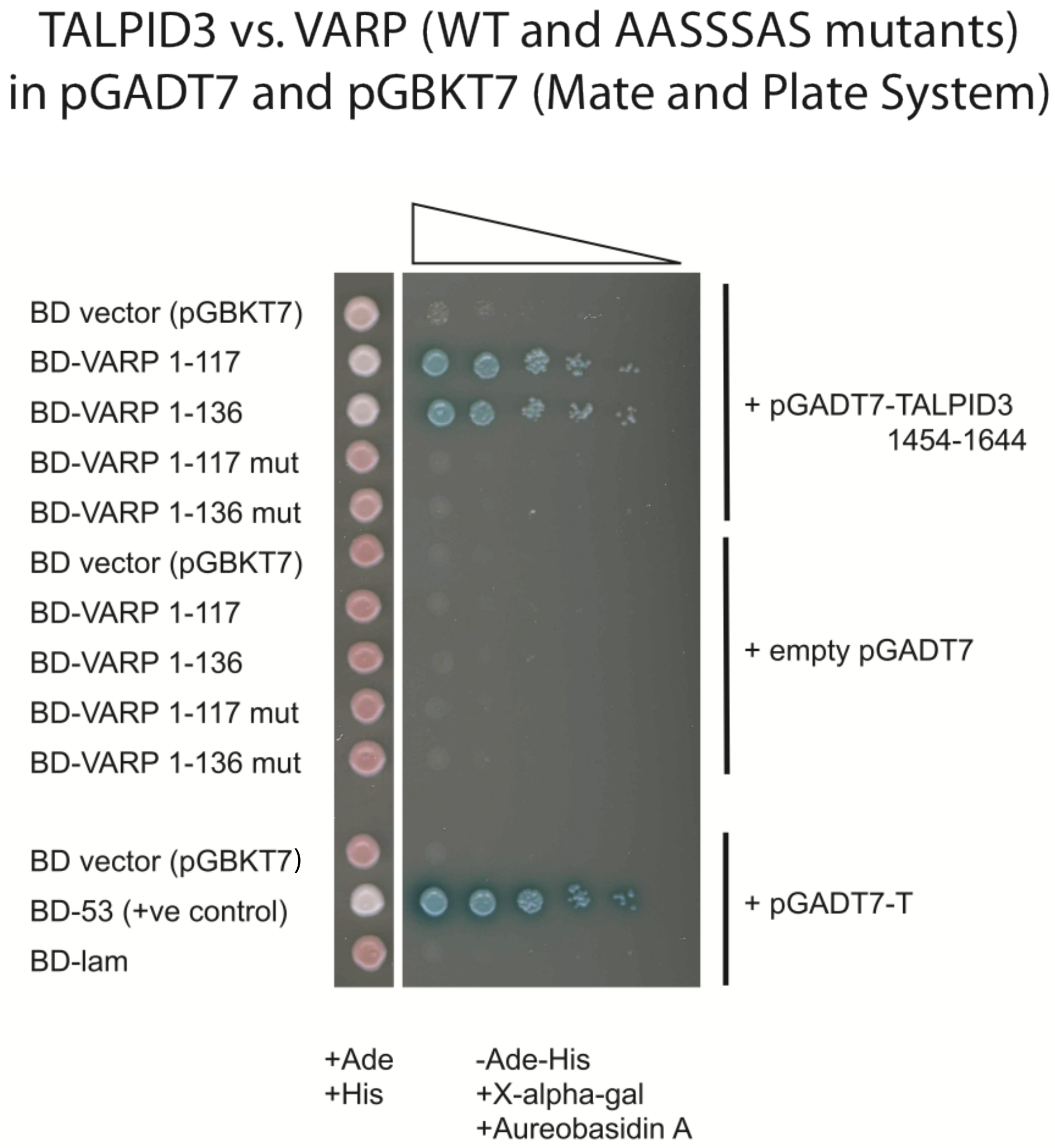


Figure 64: Growth of colonies in directed Y2H screen undertaken by Sally Gray. The colonies represent 2 days growth at 30°C. BD-53 (in pGBKT7) and pGADT7-T represent a positive control in this system.

Other reporter systems can be used in a Y2H screen to detect protein-protein interactions. This includes using the alternative DNA-binding domain, LexA, from *E.coli* with the GAL4 activation domain (AD) to induce reporter gene expression in specialized

S.cerevisiae strains (here referred to as the Hybrigenics system). As there are some differences between the Mate and Plate™ and Hybrigenics systems, including expression levels of bait proteins, interactions may be detected differently. Several attempts were made to confirm this interaction in the Hybrigenics system. Plasmids containing either Varp 1-136, 1-117 or a longer construct with 1-396 (that contains the Vps9 domain of Varp), all auto activated in this system, although, greater growth was evident for Varp/Talpid3 double positive diploids (See Figure 65 for an example experiment from this system). Interestingly, the 1-396 Varp construct demonstrated the best growth in this system, possibly as a result of the Vps9 domain stabilising the N-terminal region. The growth of the 1-396 construct far exceeds the growth of the autoactivating diploids (any pB29 – VARP with empty p6).

Whilst the Y2H data does support the idea of an interaction between Talpid3 and Varp, it is clear that further studies will be required to more definitively confirm this interaction, utilizing methods such as pulldown assays and *in vivo* studies.

VARP vs. TALPID3 in pP6 and pB29 (Hybrigenics system)

pP6 has N-terminal Gal4 AD, pB29 has C-terminal LexA

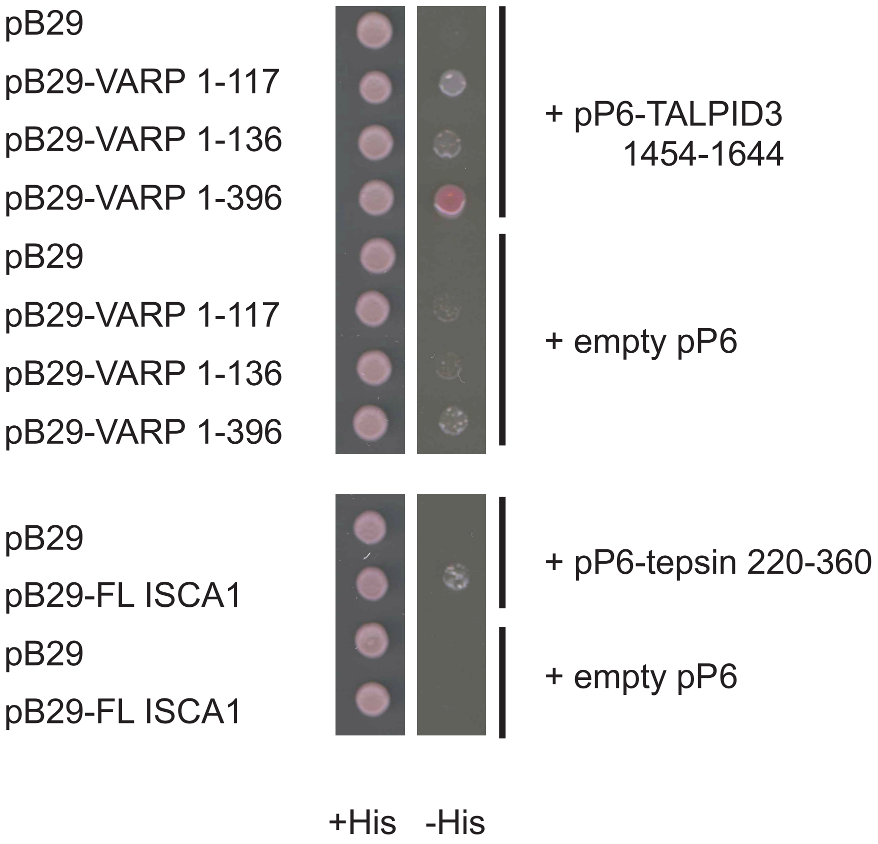


Figure 65: Growth of colonies in directed Y2H screen undertaken by Sally Gray. The colonies represent 3 days growth at 30°C. pB29-FL ISCA1 and pP6-tepsin 220-360 represent a positive control in this system.

Chapter Seven

7 Conclusion and discussion

In this study, the interaction of mammalian Varp with Vps29 was investigated at a structural and functional level. The NMR/X-ray crystallography hybrid structure of the Varp Znk2:Vps29 complex combined with biochemical studies as presented in this thesis, demonstrate that the interaction can be mediated by either of two short, homologous zinc binding sequences located at 428-443aa and 709-723aa of Varp binding to a hydrophobic patch on the surface of Vps29. The structure shows how the second Varp zinc knuckle projects a small zinc/4 cysteine stabilised 3-residue loop that is bound between a beta-sheet and a small helix of Vps29. A *de novo* SPR-based binding assay was developed and used to confirm the mechanism of interaction through the design of specific point mutations. When transfected *in vivo* these mutants, in contrast to wtVps29, fail to rescue defects in endosome to plasma membrane trafficking caused by Vps29 deletion. Finally, a Yeast 2 Hybrid study revealed a multitude of potential novel interaction partners for the N-terminus of Varp, possibly expanding the range of known Varp functions, which could form the basis of future study.

7.1 Varp and Vps29 interaction

The results presented in Chapters 3 and 4 describe the biophysical and structural characterization of the interaction between Varp and Vps29.

7.1.1 Surface Plasmon Resonance

A Varp construct, consisting of residues 692-746 of Varp, and termed Znk2, was recombinantly expressed in *E.coli*, and purified as either as a fusion protein with GST, or alone as a cleavage product. Initial pulldown studies confirmed the direct interaction between Vps29 and GST-Znk2, in a concentration dependent manner. A Varp Znk1 construct (consisting of residues 369-460 of Varp) was expressed in *E. coli* as a GST fusion protein but showed inferior biophysical properties. Subsequently, Surface Plasmon Resonance data was collected and, as presented in Chapter 3 confirms the finding that Varp Znk1 and Varp Znk2 have slightly different binding affinities for Vps29

[7]. However, the interpretation of this is unclear as the SPR technique clearly has some limitations; the difference between the calculated K_D values for Znk1 and Znk2 is small and could therefore be within experimental error.

Overall, the data suggest that the interaction of Varp Znk2 with Vps29 has a 1:1 stoichiometry and, although fewer experiments were conducted with Znk1, the same can be considered to be true for Varp Znk1. The SPR data shows no signs of a two-stage binding curve that might be expected if there were two binding sites with different affinities. Given the very strong sequence similarity of Znk2 and Znk1, (see Figure 66), this further suggests that in the context of binding of full-length Varp to Vps29, one molecule of Vps29 binds to Varp Znk2 whilst a second molecule of Vps29 could bind in a similar fashion, but independently, to Znk1.

Figure 66: Alignment of the sequences of zinc binding loops of Znk1 and Znk2 of Varp.

As well as binding to Varp, Vps29 also has a critical functional interaction with Vps35, and binding experiments with Vps29 mutants demonstrate that the particular binding sites for Varp and Vps35, both of which are highly conserved, are highly selective for their respective partners. Thus the mutation L152E, which is in the centre of the Varp binding pocket of Vps29, reduces the interaction affinity for Varp Znk2, (although it is not yet clear by how much the binding interaction is reduced see Sections 5.5 and 5.6 for details) but has been reported to have no effect on the interaction with Vps35 (as reported in [147]). Conversely, an I91S mutant of Vps29 almost entirely abolishes Vps35 binding [147], but has no impact on the binding of Varp Znk2.

This experiment thus demonstrates that Vps29 has the ability to bind both Vps35 and Varp simultaneously, meaning that Varp can be bound to Vps29 as part of the larger retromer core complex, and also most likely as part of SNX-containing retromer coats as well. Upon repeating the GST-Znk2:Vps29 binding experiments in the presence of zinc chelating agent EDTA, the affinity of the interaction was demonstrably 7- to 15-fold lower (45.5μM (+/-12μM)/86μM with EDTA vs 6μM without EDTA). It should be noted that there are limitations to the experiments performed that may affect the absolute value of the affinities determined in the case of both L152E Vps29 and the interaction in the presence of EDTA. These arise primarily from the interactions in question not reaching a steady state, and therefore saturation, over the concentration range used. This limits the accuracy of the calculated affinity values, and due to this the actual affinity of these interactions is likely to be much lower.

As a direct result of the findings from the SPR experiments presented in this thesis, Nathan Zaccai (Cambridge Institute of Medical Research) performed additional SPR experiments using the Vps29 mutants published prior to this study, structure directed mutants designed in this study and also some Znk2 mutants (See Section 5.6). A repeated study of the binding interaction between L152E Vps29 and GST-Znk2 demonstrated that the affinity of this interaction was approximately 227μM. This work also revealed that structure directed Vps29 mutants Y165S and L26S reduce the affinity for GST-Znk2 *in vitro* whilst other Vps29 mutants, F150A and F150V, could not be successfully expressed in *E. coli* for biophysical analysis. A mutant form of GST-Znk2 with all four zinc co-ordinating cysteine residues mutated to serine, was also shown to have an approximately 9-fold reduction in binding ability, similar to the EDTA treatment reported in this thesis. This further supports the idea that a four cysteine-coordinated zinc is important for pre-organising/rigidifying the Vps29 binding site of Znk2 residues HPL and thus avoiding entropic penalties upon binding. Additionally, a mutant of Znk2 with Histidine mutated to Serine and Leucine mutated to Alanine, completely abrogated binding to Vps29. This supports the structural information that illustrates that the HPL residues are the critical residues that contact Vps29.

Taken together, these binding interaction studies identify key residues that are required for the low micromolar affinity of the Varp Znk2: Vps29 interaction. In addition, they also show the key role played by the zinc coordinating cysteines in the optimal interaction of the two components. It is evident that the interaction between Varp and Vps29 is strong enough to be consistent with a genuine *in vivo* interaction, but also weak enough to enable the timely dissociation of the complex when needed. These studies additionally show that both the rate of association and dissociation of the complex are fast. These attributes can be considered to be important for proteins involved in trafficking pathways in the cell, as spatial and temporal activity must be very tightly regulated over short timescales. However, regulation by zinc seems highly unlikely as the concentration of cellular zinc doesn’t alter significantly with biological stimuli.

7.1.2 Structure of the complex

As presented in Chapter 5, Varp has been shown to interact with a conserved hydrophobic patch on Vps29, using three key residues in a conformation stabilised by zinc coordinating cysteines.

As discussed in section 1.4.2, zinc binding motifs can be classified according to sequence, structure and coordination geometry. From both the sequence and the structure of Znk2 in Varp, it is clear that it fits into the ‘zinc binding loop’ category as described by Krishna et al [220] whilst a newer classification assembly by Andreini et al [222] would put the Varp zinc binding loop into a broader category that they named ‘zinc necklaces’. Although much study into zinc binding proteins, and domains, has been conducted in the time between these papers, there are still many zinc domains with unknown functions. Most of the ‘zinc necklace’ category have structural, dimerization, regulatory, catalytic, substrate or even unknown roles [222]. For this reason, this study likely presents a novel function for a short ‘zinc binding loop/zinc necklace’ as it is possibly the first such motif to be shown to be entirely sufficient for a heterologous protein-protein interaction, not for multimerization purposes. Although this structure has been determined using the second zinc knuckle in Varp (here called Znk2), it is very likely that the binding mode would be highly similar for Znk1, which, as shown in Figure 66, is essentially identical in the organisation of the four zinc binding cysteines, and the critical HPL residues.

7.1.2.1 Comparison of the Varp Znk2:Vps29 complex structure with other Vps29 protein complex structures

The conserved hydrophobic patch on Vps29 that is required for Varp binding can also bind to TBC1D5, which is a Rab7 GTPase activating protein (GAP) [162], or RidL, which is an effector protein from the infectious agent *Legionella pneumophila* [163]. In two separate studies by others, TBC1D5 has been reported to bind to the Vps35/Vps26/Vps29 complex with an apparent K_D equal to 220 ± 10 nM [197] and

200nM [163] whilst RidL binds the same complex with ~400nM affinity [163]. Yao et al go further to show that in the presence of Vps29 alone, TBC1D5 binds with 10μM affinity, whilst RidL binds with 200nM affinity [163]. When compared to the binding affinities of Varp Znk1 and Znk2 for Vps29 alone as determined in this study, with K_D values of 4.8μM and 6.0μM respectively (see Chapter 3, section 3.3), TBC1D5 binds with very similar affinity whilst RidL has a significantly higher affinity. For pathogenic organisms to outcompete the endogenous binding partners for proteins in the cell, one would expect pathogen-derived proteins that function by displacing endogenous protein interaction partners would need to have a much higher affinity, and these results agree with this principle. However, it must be noted that both Yao et al [163] and Jia et al [197] used isoform 2 of Vps29 for their studies, which has the alternate, longer, MAGHR- sequence N –terminal tail, rather than only M as in the canonical sequence for the first-identified isoform 1 used in this thesis. For RidL binding, this seems to be of no consequence, as no residues in the determined structure interact with the additional residues [163]. In contrast, the PDB structure of the Vps29:TBC1D5 complex (5GTU), shows that residue -3Met in the longer N-terminal tail of Vps29 isoform 2 interacts extensively with 150Trp and 153Phe in an α-helix of TBC1D5, as shown in Figure 67; therefore the affinity measurements in that case may not be directly comparable to those in this thesis for binding to the zinc knuckles of Varp. The relative expression levels of Isoform 1 Vps29 and Isoform 2 Vps29 *in vivo* have not been determined, so it is unclear which isoform is more relevant biologically.

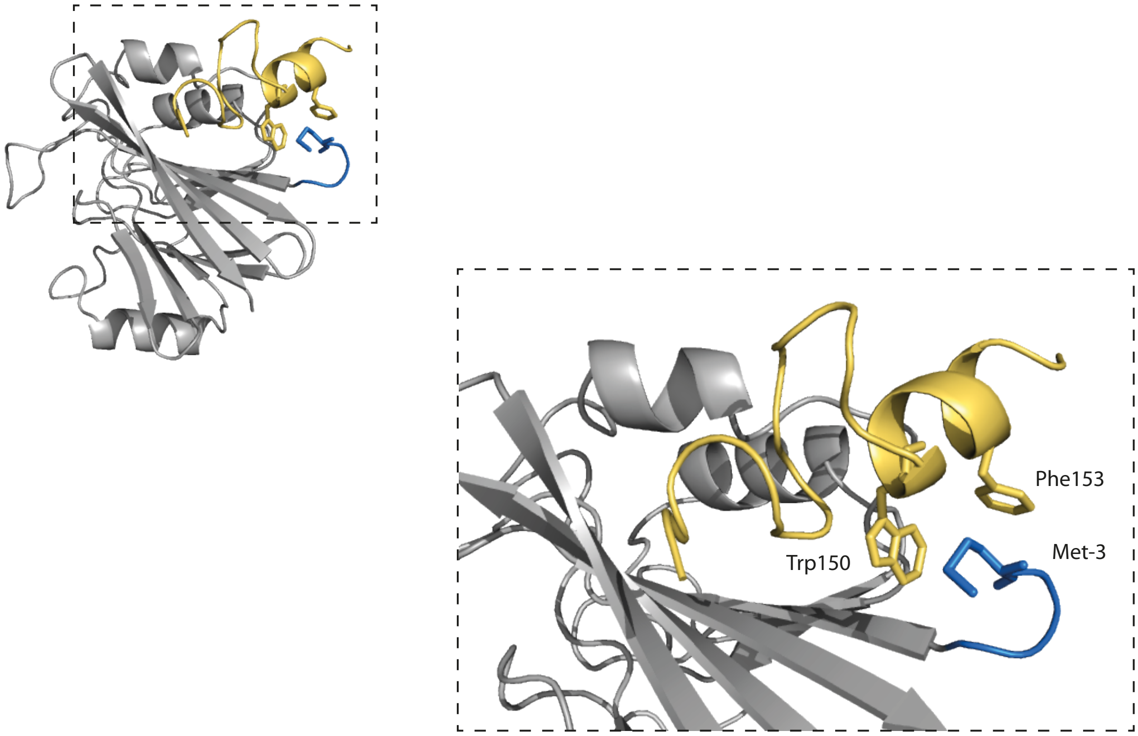


Figure 67: Structure of Vps29:TBC1D5 complex from PDB 5GTU. The TBC1D5 peptide is shown in yellow with Trp150 and Phe153 in stick format, Vps29 in grey, with the alternative tail shown in blue, and Met -3 in stick format.

Yao et al note that for the RidL:Vps29 interaction, key residues in RidL that are required for Vps29 binding are located in a hairpin loop and those residues that form an extensive hydrogen bond network around it, and that this network creates the correct loop conformation without directly contacting the Vps29 molecule, are required for Vps29 binding [163], see Figure 68. Much the same is seen in the structure with Varp Znk2, however in that case the loop conformation is maintained by the zinc co-ordinating cysteine residues, rather than by a hydrogen bond network. It is likely that the need for loop-stabilization in these proteins represents avoidance of an entropic penalty upon formation of the complex.

The Vps29-binding loop of TBC1D5 also contains three key residues that contact the hydrophobic site on Vps29, namely N140, P141, L142, which are analogous to the HPL motif in Varp Znk2, but how these residues are stabilized to create a conformation conducive to Vps29 binding is unclear, see Figure 68. In contrast to TBC1D5 and Varp Znk2, RidL uses a larger number of residues in a loop to contact Vps29: Y166, P168, I170, P171, P172, with RidL making a contact with Arg176 in the beta sheet in Vps29, as shown in Figure 68, which is not seen for the Varp Znk2:Vps29 or TBC1D5:Vps29 complex structures.

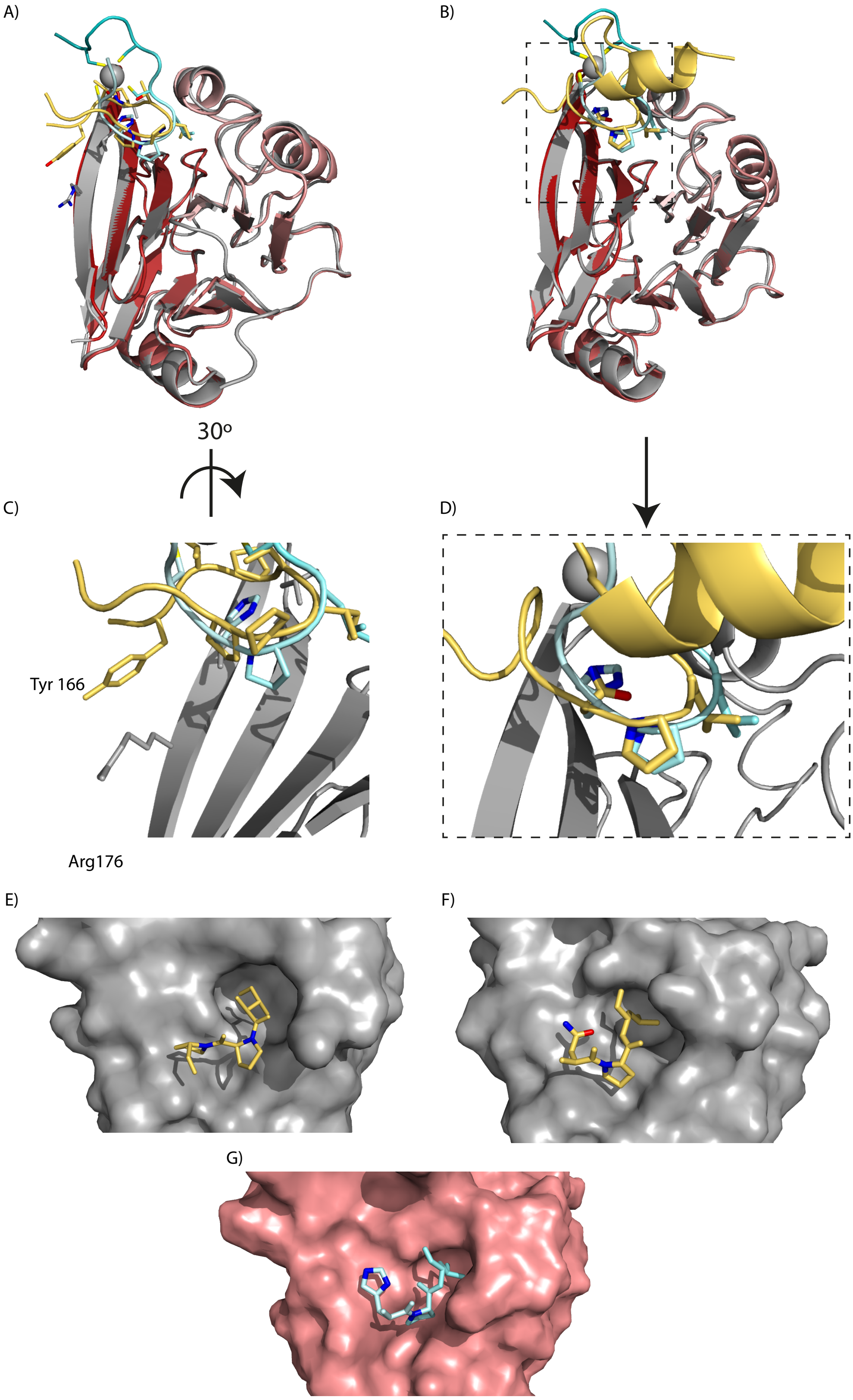


Figure 68: Structures of other Vps29:protein interactions for comparison. A) Structure of RidL binding Vps29 from PDB 5WYH. RidL peptide shown in yellow, Vps29 from 5WYH shown in gray, Vps29 from the Varp Znk2:Vps29 structure in red, Varp

Znk2 peptide shown in cyan. B) Structure of TBC1D5 binding Vps29 from PDB 5GTU. TBC1D5 peptide shown in yellow, Vps29 from 5GTU shown in gray, Vps29 from the Varp Znk2:Vps29 structure in red, Varp Znk2 peptide shown in cyan. C) Close up of the RidL:Vps29 binding site, with Varp in cyan. D) Close up of the TBC1D5:Vps29 binding site, with Varp in cyan. E) Surface representation of Vps29 from PDB 5WYH. RidL peptide shown in yellow, Vps29 shown in gray. F) Surface representation of Vps29 from PDB 5GTU. TBC1D5 peptide shown in yellow, Vps29 shown in gray. G) Surface representation of Vps29 from the Varp Znk2:Vps29 structure. Vps29 shown in red, Varp Znk2 peptide shown in cyan.

During the process of investigating the evolution of Vps9 domains in eukaryotes, Herman et al discovered that Varp (defined by the N-terminal ~400 residues that encompass Rab21 GEF activity) is an ancient protein, found widely in opisthokonts (a subset of eukaryotes, encompassing both Holozoa and Holomycota) however the zinc knuckles are relatively recent acquisitions, appearing only in Holozoa [242]. This is consistent with the lack of any Cys-rich region in the *S.cerevisiae* protein Vrl1. In collaboration with the Dacks group, we were able to show that the key Varp and TBC1D5 binding residues in Vps29 (Leu26, Leu 152 and Tyr165), and also the NPL (or HPL, DPL) residues in TBC1D5 appear much earlier in evolution than the zinc knuckles, prior to the Last Eukaryotic Common Ancestor (LECA) and, although Varp was present in the LECA, the zinc knuckles were not a feature at this point. Clearly then, the interaction between TBC1D5 and Vps29 pre-dates the interaction between Vps29 and Varp during evolution, and it is possible that Varp evolved to enable retromer to interact with early, Rab21-bearing endosomes: certainly, the acquisition of Vps29 binding was facilitated by the binding site, centered on L152E, already having evolved.

7.1.2.2 Varp Znk2:Vps29 structure in wider context

The structure of Varp Znk2 and Vps29 in complex apparently confirms a 1:1 interaction model. Why then does Varp have two zinc knuckles? It could be the case that two Vps29 molecules (and therefore retromer core molecules) are linked by interaction with Varp, and indeed this has been shown to be the case with retromer and Fam21 [200]. It remains to be seen how the interaction between Varp and Vps29 is mediated in the context of full length Varp, and this will probably require investigation using techniques such as electron microscopy and cryo-electron tomography. Varp Znk2 is located on a flexible, disordered loop as seen in the NMR structures, the Varp ANKRD2 crystal structure (where no electron density is present for the region encompassing the entire Znk2 region) and also the PSIPred prediction of secondary structure. It is most likely that Znk1 is also located on a (different) flexible polypeptide loop, and if this is the case this could allow the two zinc knuckles of Varp to bind to two Vps29 molecules that are some significant distance from each other.

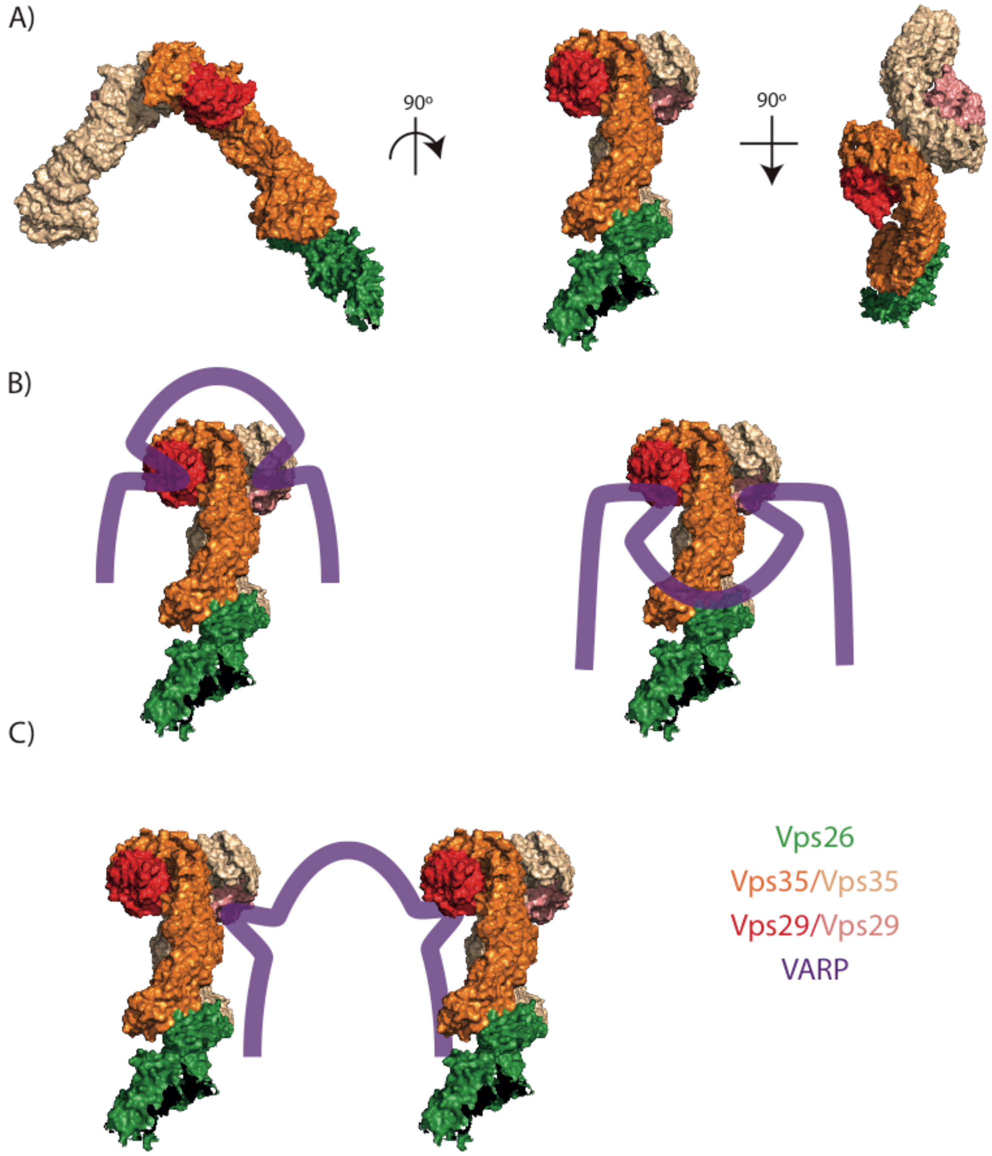


Figure 69: Potential structural models for the mode of interaction between full-length Varp and retromer core complexes. A) The structure of a partially complete retromer core complex dimer as determined by Kovtun et al [149]. Two Vps29 molecules are shown in red and pink, two molecules of Vps35 are shown in dark and pale orange and only one Vps26 molecule is shown, in green. The retromer core dimer complex can be seen to be mediated by two Vps35 molecules that interact on the reverse of the Vps29 binding interface. B) Two alternative schematic models of how Varp might bind via the two zinc knuckles to the two copies of Vps29 in a retromer

dimer. C) A schematic model of how the two zinc knuckles of Varp could act to join two dimers of retromer core complex.

Research published recently by Kovtun, Leneva et al. has shown, using cryo-electron tomography, that on tubules generated from large liposomes decorated with retromer core/SNX-BAR proteins, retromer is oriented in an elongated form protruding from the surface, with Vps26 closest to the membrane and Vps29 "reaching out" towards the cytosol, in an arch-like structure (see Figure 69) [149]. They have also shown that the retromer core complex is capable of dimerizing via the C-terminus of Vps35, ensuring that two Vps29 molecules are close in space, yet both molecules have the Varp binding patch accessible to the cytosol. It is intriguing to hypothesise that Varp contains two zinc knuckles in order to bind to arched dimeric retromer core complexes, or to possibly to bridge between different retromer arches, thereby helping to form a lattice, and this will certainly be investigated in the future. Also, as Varp contains two copies of the zinc knuckle motif, the effective affinity will be increased by avidity effects. Modelling the Varp protein onto the cryo-electron tomography structure would be a good first step in visualising how Varp achieves Vps29 binding in a cellular context. However, no full-length structure of Varp yet exists, although approximately 75% of the structure of Varp has been determined in smaller sections.

In order to test the functional significance of the Varp Znk2:Vps29 complex structure, several structure-directed mutants of Vps29 were prepared and tested *in vitro* as previously described, and these were also tested *in vivo*; the structure-directed mutants of Vps29 for expression in HeLa cells were cloned by Lena Wartosch (formerly Cambridge Institute for Medical Research). HeLa cells were depleted of endogenous Vps29 using siRNA technology and transfected with RFP-tagged mutant Vps29 in fluorescence imaging experiments. These experiments confirmed that the localisation of Varp is disrupted when L152E, L26S or Y165S mutants replace WT Vps29 under these conditions (see Section 5.7). Other mutants L26A, F150A, F150V and V174D had no effect on Varp localisation in HeLa cells *in vivo*. Whilst the effect of these mutations on Varp localisation are mixed, with some being more effective than others, it does clearly show that Varp binding to Vps29 can in some cases be prevented using single mutants that were designed based on the complex structure presented in this thesis. The fact that all mutants colocalised with Vps35 similarly shows that they were all folded, and able to be incorporated into membrane associated retromer complexes. As the trafficking of GLUT1 from endosomes to the plasma membrane requires the interaction between Varp and Vps29 [7], we are currently investigating how this pathway is perturbed by different Vps29 mutations, as perhaps more subtle phenotypes could be observed using this method (although currently the HeLa cells traditionally used in this assay are proving problematic to work with).

7.2 Varp N-terminus

The results presented in Chapter 3 describe the discovery of a conserved LFEETFY sequence in the Varp N-terminus (residues 1-136), as well as the identification of potential interaction partners for this region.

The Yeast 2 Hybrid screen of the Varp N-terminus against a normalised human library produced many hits of potential Varp-interacting proteins. Since Yeast 2 Hybrid experiments can produce a number of false positives, interactions detected by Y2H should be further tested; this is commonly done by using directed Y2H with only the two proteins of interest, in addition to other biochemical techniques such as pulldowns, immunoprecipitations and binding assays using recombinant proteins. Not only this, some interactions could be missed, possibly due to steric hinderance by the GAL4 DNA-binding domain attached to the bait, or low copy numbers of specialised proteins in the library used for screening and hence other tests such as mass spectrometry following immunoprecipitations or pulldowns from cell cytosols should be performed.

7.2.1 Varp and AP-1?

Despite the potential limitations just discussed, in this study some interesting possible interaction partners were detected that may provide more insight into Varp functions within the cell. In particular, the mu1-B (AKA AP-1 mu2) subunit of the trafficking protein AP-1, which is involved in trafficking between the endosome and the Trans Golgi Network, AP-1 was detected multiple times in the screen, specifically, the AP-1 mu1-B subunit. This interaction does seem plausible, considering that AP-1 is involved in trafficking to melanosomes and in melanosome biogenesis, together with Rab32 and Rab38 that are known interaction partners of Varp [243]. Indeed, it has been shown that abolishing the interaction of Varp with Rab32/38 restricts the peripheral localisation of TYRP1 in melanosomes [216]. It has been proposed that AP-1, AP-3 and

BLOC-2 all act independently at the endosome to sequester cargo destined for the maturing melanosome, based on the trafficking of tyrosinase and TYRP1 [243]. Also, it was demonstrated that both Rab32 and 38 show no colocalization with SNX1, so the role of Varp in melanosome biogenesis is most likely not mediated by SNX-BAR/retromer coats [243]. This certainly gives support to the idea that Varp may interact with AP-1 to mediate this specialised trafficking pathway in melanocytes. If the interaction between Varp and AP-1 mu1-B can be confirmed, it will certainly be worth undertaking *in vivo* experiments in future work to determine the impact of abolishing this interaction, by measuring the impact on TYRP1 trafficking.

7.2.2 Varp in ciliogenesis?

Unexpectedly, a centrosome-associated protein, Talpid3 was also detected multiple times in the Y2H screen with a Varp 1-136 construct. A C-terminal fragment of Talpid3 was detected 4 times in the screen and the 1454-1644 amino acid sequence of Talpid3 detected had one mutation (P to L at position 1568). Subsequent further investigation by Sally Gray and Emeritus Professor Paul Luzio (Cambridge Institute of Medical Research) established, using a directed Y2H screen, that the interaction was between the N-terminus of Varp (both aa 1-117 and aa 1-136 constructs) and Talpid3 (aa 1454-1644), as shown in Section 6.3. Talpid3 is a ubiquitously expressed, centrosome-associated protein with two putative coiled coil regions, however no structural data is yet available for any region of Talpid3. Four isoforms of the protein are produced by alternative splicing, and it is a C-terminal fragment of the longest of these, isoform three (1644aa long), that has been demonstrated to interact with the 1-136aa or 1-117aa N-terminal fragments of Varp. Talpid3 is also known to interact with the proteins CP110, Cep290, Cep97 and Kif24 [244], all of which have roles in the initial stages of primary cilium formation. Talpid3 is also required for hedgehog signalling, as hedgehog signalling is reliant on proper primary cilia formation [245, 246]. Although Varp has never been shown to be directly required for ciliogenesis, the Varp interaction partner VAMP7 has been shown to be required for proper formation of primary cilia in MDCK cells (the interaction between Varp and VAMP7 was discussed in Section 1.4.1) [247], so a role for Varp in this context could also be plausible, especially when considering that a significant proportion of HA-tagged Varp is seen to localise to the plasma membrane [129]. Also, as will be discussed later, the Rab8 GEF Rabin8 seems to cooperate with Varp to promote the ciliary trafficking of VAMP7 [248].

-

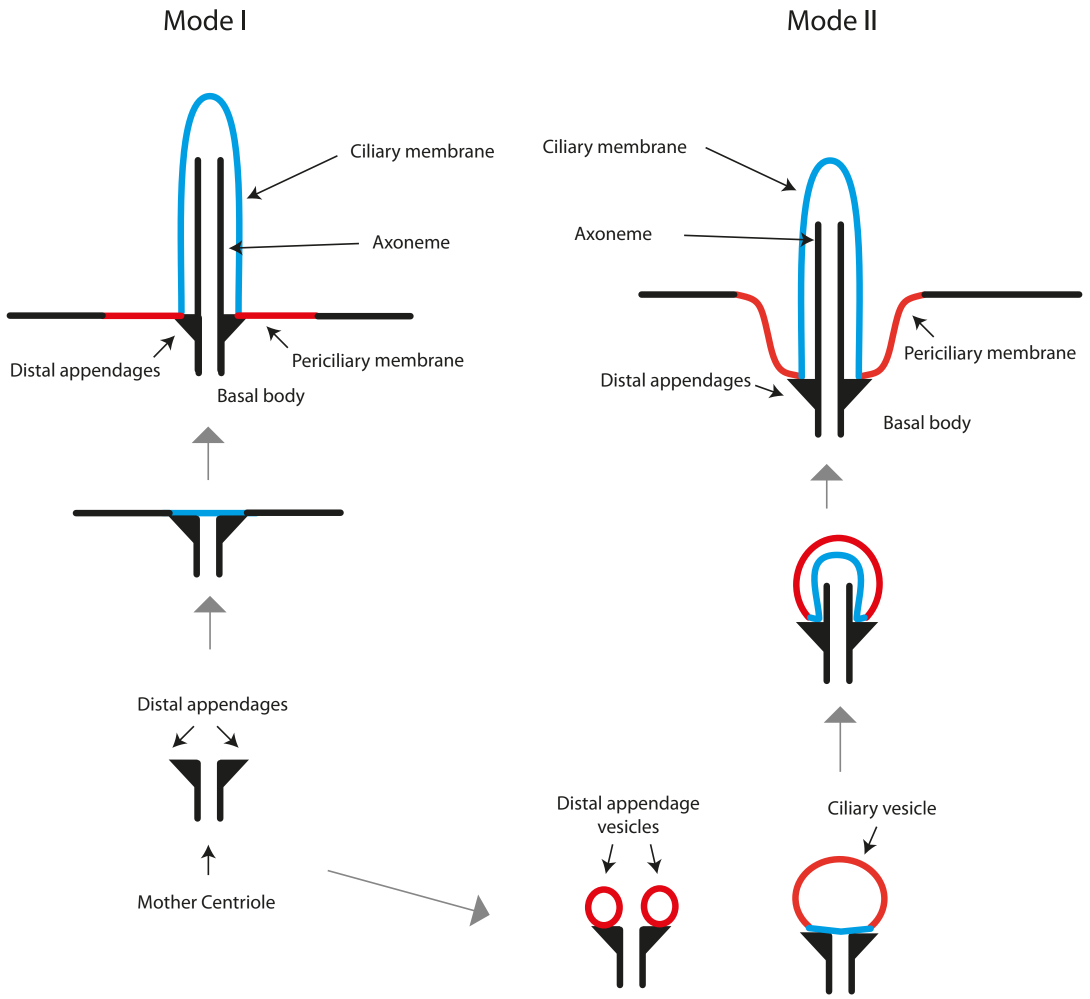


Figure 70: A diagram of the two modes of ciliogenesis. Adapted from [249]. Mode I is thought to occur in polarized epithelial cells whilst mode II is thought to occur in fibroblasts.

There are thought to be two modes of primary cilia formation. Currently it is thought that in polarised epithelial cells, a differentiated mother centriole, termed a basal body, docks onto the apical plasma membrane using transition fibres/distal appendages formed from a complex of proteins, mode I in Figure 70. From here, the ciliary axoneme is extended and using post-Golgi vesicles, the membrane is elongated into a ciliary membrane with specialised proteins [249]. In fibroblasts however, it is thought that the initial stage of primary ciliogenesis relies on the formation of a ciliary vesicle that attaches to the distal end of the mother centriole, mode II in Figure 70 [249]. This process is achieved by the fusion of a multitude of smaller 'distal appendage vesicles' into a larger vesicle, in a process that is regulated by removal of the regulatory protein CP110 [250]. In both instances, the basal body resides within a specialised "transition zone" in the membrane [251]. A brief overview of these modes of ciliogenesis is shown in Figure 70.

Primary cilia are an almost universal feature of human cells. In zebrafish photoreceptor cells, where the primary cilium is specialised to form a light sensing organelle, the knockout of Talpid3 has been shown to alter both the positioning and the docking of the basal body required for proper photoreceptor cell formation and function [252]. Additionally, knockdown of Talpid3 has been shown to cause the aberrant localisation of centriolar satellites in retinal pigmented epithelial cells (RPE1 cells) and also to have an impact on initial ciliary vesicle formation [244]. Centriolar satellites are seen as non-membranous trafficking vehicles for cargo destined for the centrosome or the basal body, and therefore any membrane proteins required for cilium formation must either be inserted into the apical cell membrane and cross the diffusion barrier to enter the periciliary membrane, or be inserted directly into the periciliary membrane at the base of cilia [253, 254]. The small GTPase Rab8 has many functions within the ciliogenesis pathway, including a role in the creation of the ciliary vesicle in some cells [244]. Subsequent to basal body docking, Rab8 also acts to fuse membranous carriers containing essential ciliogenesis cargo to the membrane at the ciliary base and, accordingly, perturbing Rab8 function in frog photoreceptors leads to the accumulation of rhodopsin transport carriers (RTCs) at the base of the cilium [255].

Whilst Talpid3 is known to be involved in the localisation of the non-membranous carriers termed centriolar satellites, one of the main functions of Talpid3 in ciliogenesis is thought to be related to either the activation or localisation of Rab8, and presumably therefore the delivery of Rab8 positive membranous trafficking vesicles also. This seems to be cell dependent, for example in RPE1 cells the knockout of Talpid3 had an effect upon the localisation of Rab8 [244] however Rab8a localisation was not affected in Talpid3 knockout zebrafish photoreceptor cells [252]. Interestingly, the expression of constitutively active Rab8a did negate the effect of a Taplid3 knockout in these cells, that is normal basal body docking and positioning were achieved [244, 252]. It has been demonstrated that residues 701-1533 of Talpid3 are required for function in both ciliogenesis and Rab8a vesicle docking [256], and this overlaps with the 1454-1644 residue region that was detected in the Varp-baited Y2H screen in the work presented in Section 6.2 and 6.3.

Moreover, a role for the SNARE protein VAMP7 in ciliogenesis is currently emerging. A VAMP7 knockout study showed that the loss of VAMP7 does not completely prevent cilia formation, yet it does reduce the frequency and the length of cilia in MDCK cells [247]. Upon further investigation, it appears that VAMP7 interacts directly with the Rab11-Rabin8-Rab8 ciliogenesis machinery to assist in the delivery of Golgi-derived rhodopsin transport carriers (RTCs) laden with ciliary cargo to the ciliary base [248]. At the ciliary base, VAMP7 interacts with SNAP25 and syntaxin3 to fuse with the periciliary membrane, however the partial effect of a VAMP7 knockout on cilium length and frequency does suggest some SNARE protein redundancy [248]. This study also postulates that Rab8 and Rab11 interact with the longin domain of VAMP7 when it is in the open conformation, and Rabin8 was shown to interact with the VAMP7 SNARE domain; interestingly, these interactions are unlikely to exclude Varp from interacting with VAMP7 concurrently [248]. Naharros et al propose that Talpid3 acts to localise Rab8 effectors to the basal body, including Rabin8 [252], and Kandachar et al use GST-Pulldown experiments to suggest that Rabin8 binds to Varp [248]. Considering this, the most likely scenario is that Varp is acting as a link between VAMP7 and the ciliogenesis machinery, either delivering or retrieving VAMP7 from the site of ciliogenesis.

As Talpid3 seems to act upstream of Rab8, promoting recruitment and/or activation, and VAMP7 seems to be required at a stage when the Rab8 effector Rabin8 has been recruited to the forming cilia, i.e Rab8 is active, then the point at which Varp is involved in this process is unclear. Talpid3 does persist at the ciliary base throughout ciliogenesis and this suggests the hypothesis that Varp binds Talpid3 at the ciliary base to hold closed VAMP7 at the transition zone of forming primary cilia, until such time as it is required to participate in fusion events. In this scenario, Varp is acting as linker between the VAMP7-positive cargo laden membranous carriers and the ciliary base. This theory seems to fit the available evidence, as Varp and Rabin8 were detected on VAMP7 positive Golgi-derived vesicles and Varp was also detected on rhodopsin laden

Golgi derived structures termed RTCs destined for the cilium [248]. However, Dennis et al have previously demonstrated a role for Varp in the recycling of VAMP7 away from melanosomes [218]. Thus, it could be hypothesised that the function of the Varp - Talpid3 interaction is to hold Varp at the ciliary base to ensure a supply of Varp to recycle post-fusion VAMP7. It could even be the case that both of these are true. That is, Varp binds to Talpid3 at the ciliary base to link membranous carriers to the ciliary base pre-fusion, and then Varp persists at the ciliary base via interaction with Talpid3 to recycle post-fusion VAMP7. It will clearly require much further investigation to answer these questions about the precise timing and function of the Varp:Talpid3 interaction.

In a wider context, a known Talpid3-interacting protein, MACF1, has been shown to bind golginA4 [257] and it has been previously reported that golginA4, and another protein Kif5A, both bind to Varp [6]. However, neither golginA4 or Kif5A were detected in the Y2H screen presented in this thesis, nor, to date, could the proposed golginA4 interaction with the first set of ankyrin repeats in Varp be independently verified in directed Y2H experiments (Luzio and Owen labs – personal communication). Clearly, the link between Varp and Talpid3 remains to be fully explored, and there may yet be other linked proteins that have yet to be detected.

In further investigating the function of Varp, members of both the Luzio and Owen groups have discovered evidence that Varp binds to ERC-1 (ELKS/Rab6-interacting/CAST family member 1) (personal communication), a protein involved in the docking and fusion of secretory vesicles with the plasma membrane [258]. ERC-1 is known for its scaffolding role at the active zone; the active zone is a protein dense structure that is formed at a pre-synaptic membrane to enable rapid and precise exocytosis [259], as well as being implicated in Hedgehog signalling [260], melanosome biogenesis [261] and microtubule stabilisation [262]. It has been further demonstrated that ERC-1 is localised to centrioles in non-ciliated cells, and is retained at the basal body during primary cilium formation [260]. Although the role of ERC-1 in ciliogenesis is not yet fully understood, it is known that ERC-1 is required for proper docking and fusion of Rab6 and Rab8 double positive exocytotic vesicles at the specialised regions of the plasma membrane where active zones are formed [263]. Grigoriev et al also show that Rab8 is present at cilia and is also present on Rab6 positive exocytotic vesicles, whilst Rab6 is not present at cilia [263]. Interestingly, Rab6 has been demonstrated to be somehow involved in the trafficking of cilia-destined cargo, as the ciliary protein polycystin-1 co-immunoprecipitated Rab6, Rab8 and Rab11 in MDCKII cells [264]. More recently, it has been shown that the RhoGEF ARHGEF10 can localise Rab8 to Rab6-positive exocytotic vesicles [265, 266]. Whilst no binding has been detected between ERC-1 and Rab8, novel interactions between both ERC-1 and Rab8 with the protein MICAL3 have been demonstrated [263]. MICAL3 has subsequently been detected in a co-immunoprecipitation experiment with NINL, a microtubule organising protein which appears to control MICAL3 localisation to the ciliary base [267]. It is proposed by Bachmann-Gagescu et al that both NINL and CC2D2A act to dock Rab8 positive rhodopsin-containing vesicles at the ciliary base, where ERC-1 and MICAL3 subsequently act to remodel the docking complex, thus allowing fusion of the vesicle at the periciliary membrane [267].

How does one reconcile the fact that Varp binds to both Talpid3 and ERC-1? Both Talpid3 and ERC1 are present at the basal body and so perhaps this is an instance of ‘coincidence detection’ at the cilium base, making the delivery of vesicles containing Varp in their coats more efficient. Alternatively, it is possible that Talpid3 acts to sequester Rab8 positive vesicles near the mother centriole at a stage preceding plasma membrane insertion of the growing cilia, and at such a time that the vesicle components are required i.e. the transition zone is formed in the membrane and vesicles containing Varp are bound by ERC-1 in a process that enlarges the cilium membrane. It could be that vesicles containing Varp are required at an even earlier stage in some cells, that is in the generation of a ciliary vesicle from smaller, distal appendage vesicles. This is possible as Talpid3 is located at the centriole at this stage, where it binds to CP110.

It should be noted that the retromer regulating proteins, Eps15 homology domain containing proteins EHD1 and EHD3 are required for ciliogenesis also [268]. EHD1 and SNAP29 act to coalesce the distal appendage vesicles once CP110 is removed from the centriole [268]. In addition, the retromer core complex, comprising Vps29, Vps35 and Vps26, has been hypothesised to be involved in the recycling of the cilium protein polycystin-2 [269]. Clearly then there is an emerging interaction network that is building links between components of endosomal recycling pathways and both the ciliogenesis machinery and cargo proteins destined for the cilia. Future investigations of the role of Varp, Talpid3, ERC-1, VAMP7 in ciliogenesis should provide a better understanding of this process.

7.3 Further roles for Varp?

Clearly, there is a distinct possibility that Varp may have still more varied roles than are understood at present. By attempting to discover the function of the N-terminal region of Varp, this thesis has provided some insight into novel interaction partners for Varp.

It should be noted that it is well known that Yeast 2 Hybrid screening has limitations when detecting protein interactions. It is possible that by using a normalised human library for the Yeast 2 Hybrid screen that some interaction partners were missed due to low copy numbers of particular proteins from specialised cell types. For instance, as discussed, because Varp has a role in trafficking melanogenic enzymes to melanosomes, it may be interesting to carry out a Yeast 2 Hybrid screen with a melanocyte library. There is also a possibility that the N-terminus of Varp may have a function other than as a domain to interact with cellular proteins. Parts of the N-terminus could provide a structural feature required for proper folding of the entire Varp molecule or perhaps the N-terminus contains an as-yet unrecognised dimerization domain or motif.

The potential for the N-terminus of Varp to interact with other proteins has not gone unrecognised in the wider community. Indeed, although not yet published, Guillaume Jacquemet and Johanna Ivaska (University of Turku), in collaboration with both the Luzio and Owen groups, have demonstrated that Varp is present in the tips of filopodia and in focal adhesion complexes in human breast carcinoma U2OS cells, co-localising with Myosin X (MYO10) (personal communication). Furthermore, they have discovered that deleting the N-terminus of Varp (residues 2-136) prevents Varp localisation to the tip of filopodia, but does not have an impact on filopodia formation, or Myosin X localisation. The role of Varp in this process is therefore unclear and it also remains to be seen if retromer components are required in this process. Filopodium formation relies on actin polymerisation stimulated by nucleation factors, such as the Arp2/3 complex, and it should be noted that the retromer complex interacts with

Fam21 of the WASH complex, which is a nucleation promoting factor that activates the Arp2/3 complex [200]. Therefore, there is a distinct possibility that Varp localisation to filopodia is a result of its interaction with Vps29, which will undoubtedly be an area of future investigation.

7.4 Conclusion

It is clear that the work undertaken in this thesis has improved our understanding of the interaction between Varp and Vps29, and the biology of Varp more generally. The structure of the Varp:Vps29 complex shows that a zinc binding loop can be sufficient to mediate a protein-protein interaction with low micromolar affinity. Not only this, the minimal interaction interface allows both Varp and Vps29 to engage in further interactions with other partners, whilst still bound to each other. The presence of two zinc loops potentially allows Varp to form a bridge between two Vps29 molecules, and thus two instances of the retromer core complex. Structure-directed mutants have confirmed the roles of key interfacial residues, and using such mutants to disrupt the interaction will provide a good basis for probing the trafficking of retromer-dependent cargo such as GLUT1. Moreover, this thesis presents novel potential interaction partners for a previously unstudied domain of Varp, which warrant further investigation. Due to the variety of Varp-binding proteins, there is now evidence to suggest a role for Varp in many cellular processes including Hedgehog signalling, ciliogenesis, and filopodium formation, with further implications on the role of Varp in melanogenesis.

Undoubtedly, there are still many unanswered questions more generally about the role of Varp. Unpicking the newly identified Varp:protein interactions, and also characterising previously reported interactions, will be key to understanding many important processes within the cell.

Bibliography

1. Zhang, X., X. He, X.-Y. Fu, and Z. Chang, *Varp is a Rab21 guanine nucleotide exchange factor and regulates endosome dynamics.* Journal of Cell Science, 2006. **119**(6): p. 1053-1062.
2. Zhang, X., X. He, X.Y. Fu, and Z. Chang, *Varp is a Rab21 guanine nucleotide exchange factor and regulates endosome dynamics.* J Cell Sci, 2006. **119**(Pt 6): p. 1053-62.
3. Wang, F., H. Zhang, X. Zhang, Y. Wang, F. Ren, X. Zhang, Y. Zhai, and Z. Chang, *Varp interacts with Rab38 and functions as its potential effector.* Biochemical and Biophysical Research Communications, 2008. **372**(1): p. 162-167.
4. Tamura, K., N. Ohbayashi, Y. Maruta, E. Kanno, T. Itoh, and M. Fukuda, *Varp is a novel Rab32/38-binding protein that regulates Tyrp1 trafficking in melanocytes.* Molecular biology of the cell, 2009. **20**(12): p. 2900-2908.
5. Burgo, A., E. Sotirakis, M.-C. Simmler, A. Verraes, C. Chamot, J.C. Simpson, L. Lanzetti, V. Proux-Gillardeaux, and T. Galli, *Role of Varp, a Rab21 exchange factor and TI-VAMP/VAMP7 partner, in neurite growth.* EMBO reports, 2009. **10**(10): p. 1117-1124.
6. Burgo, A., V. Proux-Gillardeaux, E. Sotirakis, P. Bun, A. Casano, A. Verraes, R.K.H. Liem, E. Formstecher, M. Coppey-Moisan, and T. Galli, *A Molecular Network for the Transport of the TI-VAMP/VAMP7 Vesicles from Cell Center to Periphery.* Developmental Cell, 2012. **23**(1): p. 166-180.
7. Hesketh, G., I. Pérez-Dorado, L.P. Jackson, L. Wartosch, I.B. Schäfer, S.R. Gray, A.J. McCoy, O.B. Zeldin, E.F. Garman, M.E. Harbour, P.R. Evans, M.N.J. Seaman, J.P. Luzio, and D.J. Owen, *VARP is recruited on to endosomes by direct interaction with retromer, where together they function in export to the cell surface.* Developmental Cell, 2014. **29**(5): p. 591-606.
8. McGough, I.J., F. Steinberg, M. Gallon, A. Yatsu, N. Ohbayashi, K.J. Heesom, M. Fukuda, and P.J. Cullen, *Identification of molecular heterogeneity in SNX27-retromer-mediated endosome-to-plasma-membrane recycling.* J Cell Sci, 2014. **127**(Pt 22): p. 4940-53.
9. Yatsu, A., H. Shimada, N. Ohbayashi, and M. Fukuda, *Rab40C is a novel Varp-binding protein that promotes proteasomal degradation of Varp in melanocytes.* Biology open, 2015. **4**(3): p. 267-75.
10. Marubashi, S., N. Ohbayashi, and M. Fukuda, *A Varp-Binding Protein, RACK1, Regulates Dendrite Outgrowth through Stabilization of Varp Protein in Mouse Melanocytes.* J Invest Dermatol, 2016. **136**(8): p. 1672-80.
11. Rink, J., E. Ghigo, Y. Kalaidzidis, and M. Zerial, *Rab conversion as a mechanism of progression from early to late endosomes.* Cell, 2005. **122**(5): p. 735-49.
12. Vilarino-Guell, C., C. Wider, O.A. Ross, J.C. Dachsel, J.M. Kachergus, S.J. Lincoln, A.I. Soto-Ortolaza, S.A. Cobb, G.J. Wilhoite, J.A. Bacon, B. Behrouz, H.L. Melrose, E.

Hentati, A. Puschmann, D.M. Evans, E. Conibear, W.W. Wasserman, J.O. Aasly, P.R. Burkhard, R. Djaldetti, J. Ghika, F. Hentati, A. Krygowska-Wajs, T. Lynch, E. Melamed, A. Rajput, A.H. Rajput, A. Solida, R.M. Wu, R.J. Uitti, Z.K. Wszolek, F. Vingerhoets, and M.J. Farrer, *VPS35 mutations in Parkinson disease.* Am J Hum Genet, 2011. **89**(1): p. 162-7.

13. Zimprich, A., A. Benet-Pages, W. Struhal, E. Graf, S.H. Eck, M.N. Offman, D. Haubenberger, S. Spielberger, E.C. Schulte, P. Lichtner, S.C. Rossle, N. Klopp, E. Wolf, K. Seppi, W. Pirker, S. Presslauer, B. Mollenhauer, R. Katzenschlager, T. Foki, C. Hotzy, E. Reinthaler, A. Harutyunyan, R. Kralovics, A. Peters, F. Zimprich, T. Brucke, W. Poewe, E. Auff, C. Trenkwalder, B. Rost, G. Ransmayr, J. Winkelmann, T. Meitinger, and T.M. Strom, *A mutation in VPS35, encoding a subunit of the retromer complex, causes late-onset Parkinson disease.* Am J Hum Genet, 2011. **89**(1): p. 168-75.

14. Small, S.A. and G.A. Petsko, *Retromer in Alzheimer disease, Parkinson disease and other neurological disorders.* Nat Rev Neurosci, 2015. **16**(3): p. 126-32.

15. De Matteis, M.A. and A. Luini, *Mendelian disorders of membrane trafficking.* N Engl J Med, 2011. **365**(10): p. 927-38.

16. Bucci, C., O. Bakke, and C. Progida, *Charcot-Marie-Tooth disease and intracellular traffic.* Prog Neurobiol, 2012. **99**(3): p. 191-225.

17. Huizing, M., Y. Anikster, and W.A. Gahl, *Hermansky-Pudlak syndrome and Chediak-Higashi syndrome: disorders of vesicle formation and trafficking.* Thromb Haemost, 2001. **86**(1): p. 233-45.

18. Cai, H., K. Reinisch, and S. Ferro-Novick, *Coats, tethers, Rabs, and SNAREs work together to mediate the intracellular destination of a transport vesicle.* Dev Cell, 2007. **12**(5): p. 671-82.

19. Pearse, B.M., *Coated vesicles from pig brain: purification and biochemical characterization.* J Mol Biol, 1975. **97**(1): p. 93-8.

20. Robinson, M.S., *Forty Years of Clathrin-coated Vesicles.* Traffic, 2015. **16**(12): p. 1210-38.

21. Malhotra, V., T. Serafini, L. Orci, J.C. Shepherd, and J.E. Rothman, *Purification of a novel class of coated vesicles mediating biosynthetic protein transport through the Golgi stack.* Cell, 1989. **58**(2): p. 329-36.

22. Barlowe, C., L. Orci, T. Yeung, M. Hosobuchi, S. Hamamoto, N. Salama, M.F. Rexach, M. Ravazzola, M. Amherdt, and R. Schekman, *COPII: a membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum.* Cell, 1994. **77**(6): p. 895-907.

23. Seaman, M.N., J.M. McCaffery, and S.D. Emr, *A membrane coat complex essential for endosome-to-Golgi retrograde transport in yeast.* J Cell Biol, 1998. **142**(3): p. 665-81.

24. Dell'Angelica, E.C., C. Mullins, and J.S. Bonifacino, *AP-4, a novel protein complex related to clathrin adaptors.* J Biol Chem, 1999. **274**(11): p. 7278-85.

25. Robinson, M.S., *100-kD coated vesicle proteins: molecular heterogeneity and intracellular distribution studied with monoclonal antibodies.* J Cell Biol, 1987. **104**(4): p. 887-95.

26. Evans, P.R. and D.J. Owen, *Endocytosis and vesicle trafficking.* Curr Opin Struct Biol, 2002. **12**(6): p. 814-21.

27. He, G., S. Gupta, M. Yi, P. Michaely, H.H. Hobbs, and J.C. Cohen, *ARH is a modular adaptor protein that interacts with the LDL receptor, clathrin, and AP-2.* J Biol Chem, 2002. **277**(46): p. 44044-9.

28. Moore, C.A., S.K. Milano, and J.L. Benovic, *Regulation of receptor trafficking by GRKs and arrestins.* Annu Rev Physiol, 2007. **69**: p. 451-82.

29. Miller, S.E., D.A. Sahlender, S.C. Graham, S. Honing, M.S. Robinson, A.A. Peden, and D.J. Owen, *The molecular basis for the endocytosis of small R-SNAREs by the clathrin adaptor CALM.* Cell, 2011. **147**(5): p. 1118-31.

30. Pryor, P.R., L. Jackson, S.R. Gray, M.A. Edeling, A. Thompson, C.M. Sanderson, P.R. Evans, D.J. Owen, and J.P. Luzio, *Molecular Basis for the Sorting of the SNARE VAMP7 into Endocytic Clathrin-Coated Vesicles by the ArfGAP Hrb.* Cell, 2008. **134**(5): p. 817-827.

31. Mettlen, M., M. Stoeber, D. Loerke, C.N. Antonescu, G. Danuser, and S.L. Schmid, *Endocytic accessory proteins are functionally distinguished by their differential effects on the maturation of clathrin-coated pits.* Mol Biol Cell, 2009. **20**(14): p. 3251-60.

32. Slepnev, V.I. and P. De Camilli, *Accessory factors in clathrin-dependent synaptic vesicle endocytosis.* Nat Rev Neurosci, 2000. **1**(3): p. 161-72.

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

34. Antonny, B., C. Burd, P. De Camilli, E. Chen, O. Daumke, K. Faelber, M. Ford, V.A. Frolov, A. Frost, J.E. Hinshaw, T. Kirchhausen, M.M. Kozlov, M. Lenz, H.H. Low, H. McMahon, C. Merrifield, T.D. Pollard, P.J. Robinson, A. Roux, and S. Schmid, *Membrane fission by dynamin: what we know and what we need to know.* EMBO J, 2016. **35**(21): p. 2270-2284.

35. Sousa, R. and E.M. Lafer, *The role of molecular chaperones in clathrin mediated vesicular trafficking.* Front Mol Biosci, 2015. **2**: p. 26.

36. Henderson, D.M. and S.D. Conner, *A novel AAK1 splice variant functions at multiple steps of the endocytic pathway.* Mol Biol Cell, 2007. **18**(7): p. 2698-706.

37. Borner, G.H., R. Antrobus, J. Hirst, G.S. Bhumbra, P. Kozik, L.P. Jackson, D.A. Sahlender, and M.S. Robinson, *Multivariate proteomic profiling identifies novel accessory proteins of coated vesicles.* J Cell Biol, 2012. **197**(1): p. 141-60.

38. Lundmark, R. and S.R. Carlsson, *SNX9 - a prelude to vesicle release.* J Cell Sci, 2009. **122**(Pt 1): p. 5-11.

39. Lee, M.C., E.A. Miller, J. Goldberg, L. Orci, and R. Schekman, *Bi-directional protein transport between the ER and Golgi.* Annu Rev Cell Dev Biol, 2004. **20**: p. 87-123.

40. Huotari, J. and A. Helenius, *Endosome maturation.* EMBO J, 2011. **30**(17): p. 3481-500.

41. Seaman, M.N., *Recycle your receptors with retromer.* Trends Cell Biol, 2005. **15**(2): p. 68-75.

42. Arighi, C.N., L.M. Hartnell, R.C. Aguilar, C.R. Haft, and J.S. Bonifacino, *Role of the mammalian retromer in sorting of the cation-independent mannose 6-phosphate receptor.* J Cell Biol, 2004. **165**(1): p. 123-33.

43. Chavrier, P., R.G. Parton, H.P. Hauri, K. Simons, and M. Zerial, *Localization of low molecular weight GTP binding proteins to exocytic and endocytic compartments.* Cell, 1990. **62**(2): p. 317-29.

44. Zerial, M. and H. McBride, *Rab proteins as membrane organizers.* Nat Rev Mol Cell Biol, 2001. **2**(2): p. 107-17.

45. De Matteis, M.A. and A. Godi, *PI-loting membrane traffic.* Nat Cell Biol, 2004. **6**(6): p. 487-92.

46. Sasaki, T., A. Kikuchi, S. Araki, Y. Hata, M. Isomura, S. Kuroda, and Y. Takai, *Purification and characterization from bovine brain cytosol of a protein that inhibits the dissociation of GDP from and the subsequent binding of GTP to smg p25A, a ras p21-like GTP-binding protein.* J Biol Chem, 1990. **265**(4): p. 2333-7.

47. Araki, S., A. Kikuchi, Y. Hata, M. Isomura, and Y. Takai, *Regulation of reversible binding of smg p25A, a ras p21-like GTP-binding protein, to synaptic plasma membranes and vesicles by its specific regulatory protein, GDP dissociation inhibitor.* J Biol Chem, 1990. **265**(22): p. 13007-15.

48. Pylypenko, O., A. Rak, T. Durek, S. Kushnir, B.E. Dursina, N.H. Thomae, A.T. Constantinescu, L. Brunsveld, A. Watzke, H. Waldmann, R.S. Goody, and K. Alexandrov, *Structure of doubly prenylated Ypt1:GDI complex and the mechanism of GDI-mediated Rab recycling.* EMBO J, 2006. **25**(1): p. 13-23.

49. Soldati, T., A.D. Shapiro, A.B. Svejstrup, and S.R. Pfeffer, *Membrane targeting of the small GTPase Rab9 is accompanied by nucleotide exchange.* Nature, 1994. **369**(6475): p. 76-8.

50. Ullrich, O., H. Horiuchi, C. Bucci, and M. Zerial, *Membrane association of Rab5 mediated by GDP-dissociation inhibitor and accompanied by GDP/GTP exchange.* Nature, 1994. **368**(6467): p. 157-60.

51. Dirac-Svejstrup, A.B., T. Sumizawa, and S.R. Pfeffer, *Identification of a GDI displacement factor that releases endosomal Rab GTPases from Rab-GDI.* EMBO J, 1997. **16**(3): p. 465-72.

52. Sivars, U., D. Aivazian, and S.R. Pfeffer, *Yip3 catalyses the dissociation of endosomal Rab-GDI complexes.* Nature, 2003. **425**(6960): p. 856-9.

53. Gerondopoulos, A., L. Langemeyer, J.R. Liang, A. Linford, and F.A. Barr, *BLOC-3 mutated in Hermansky-Pudlak syndrome is a Rab32/38 guanine nucleotide exchange factor.* Curr Biol, 2012. **22**(22): p. 2135-9.

54. Blumer, J., J. Rey, L. Dehmelt, T. Mazel, Y.W. Wu, P. Bastiaens, R.S. Goody, and A. Itzen, *RabGEFs are a major determinant for specific Rab membrane targeting.* J Cell Biol, 2013. **200**(3): p. 287-300.

55. Cabrera, M. and C. Ungermann, *Guanine nucleotide exchange factors (GEFs) have a critical but not exclusive role in organelle localization of Rab GTPases.* J Biol Chem, 2013. **288**(40): p. 28704-12.

56. Muller, M.P. and R.S. Goody, *Molecular control of Rab activity by GEFs, GAPs and GDI.* Small GTPases, 2018. **9**(1-2): p. 5-21.

57. Lippe, R., M. Miaczynska, V. Rybin, A. Runge, and M. Zerial, *Functional synergy between Rab5 effector Rabaptin-5 and exchange factor Rabex-5 when physically associated in a complex.* Mol Biol Cell, 2001. **12**(7): p. 2219-28.

58. Vonderheit, A. and A. Helenius, *Rab7 associates with early endosomes to mediate sorting and transport of Semliki forest virus to late endosomes.* PLoS Biol, 2005. **3**(7): p. e233.

59. Martinez, O. and B. Goud, *Rab proteins.* Biochim Biophys Acta, 1998. **1404**(1-2): p. 101-12.

60. Esters, H., K. Alexandrov, A.T. Constantinescu, R.S. Goody, and A.J. Scheidig, *High-resolution crystal structure of S. cerevisiae Ypt51(DeltaC15)-GppNHp, a small GTP-binding protein involved in regulation of endocytosis.* J Mol Biol, 2000. **298**(1): p. 111-21.

61. Huber, S.K. and A.J. Scheidig, *High resolution crystal structures of human Rab4a in its active and inactive conformations.* FEBS Lett, 2005. **579**(13): p. 2821-9.

62. Fukuda, M., *TBC proteins: GAPs for mammalian small GTPase Rab?* Biosci Rep, 2011. **31**(3): p. 159-68.

63. Ortiz, D., M. Medkova, C. Walch-Solimena, and P. Novick, *Ypt32 recruits the Sec4p guanine nucleotide exchange factor, Sec2p, to secretory vesicles; evidence for a Rab cascade in yeast.* J Cell Biol, 2002. **157**(6): p. 1005-15.

64. Rivera-Molina, F.E. and P.J. Novick, *A Rab GAP cascade defines the boundary between two Rab GTPases on the secretory pathway.* Proc Natl Acad Sci U S A, 2009. **106**(34): p. 14408-13.

65. Suda, Y., K. Kurokawa, R. Hirata, and A. Nakano, *Rab GAP cascade regulates dynamics of Ypt6 in the Golgi traffic.* Proc Natl Acad Sci U S A, 2013. **110**(47): p. 18976-81.

66. Gillingham, A.K., R. Sinka, I.L. Torres, K.S. Lilley, and S. Munro, *Toward a comprehensive map of the effectors of rab GTPases.* Dev Cell, 2014. **31**(3): p. 358-373.

67. Stenmark, H., *Rab GTPases as coordinators of vesicle traffic.* Nature reviews. Molecular cell biology, 2009. **10**(8): p. 513-25.

68. Zhen, Y. and H. Stenmark, *Cellular functions of Rab GTPases at a glance.* J Cell Sci, 2015. **128**(17): p. 3171-6.

69. Wurmser, A.E., T.K. Sato, and S.D. Emr, *New component of the vacuolar class C-Vps complex couples nucleotide exchange on the Ypt7 GTPase to SNARE-dependent docking and fusion.* J Cell Biol, 2000. **151**(3): p. 551-62.

70. Brett, C.L., R.L. Plemel, B.T. Lobingier, M. Vignali, S. Fields, and A.J. Merz, *Efficient termination of vacuolar Rab GTPase signaling requires coordinated action by a GAP and a protein kinase.* J Cell Biol, 2008. **182**(6): p. 1141-51.

71. Brocker, C., A. Kuhlee, C. Gatsogiannis, H.J. Balderhaar, C. Honscher, S. Engelbrecht-Vandre, C. Ungermann, and S. Raunser, *Molecular architecture of the*

multisubunit homotypic fusion and vacuole protein sorting (HOPS) tethering complex. Proc Natl Acad Sci U S A, 2012. **109**(6): p. 1991-6.

72. Schonthaler, H.B., V.C. Fleisch, O. Biehlmaier, Y. Makhankov, O. Rinner, R. Bahadori, R. Geisler, H. Schwarz, S.C. Neuhauss, and R. Dahm, *The zebrafish mutant lbk/vam6 resembles human multisystemic disorders caused by aberrant trafficking of endosomal vesicles.* Development, 2008. **135**(2): p. 387-99.

73. Messler, S., S. Kropp, V. Episkopou, A. Felici, J. Wurthner, R. Lemke, M. Jerabek-Willemsen, R. Willecke, S. Scheu, K. Pfeffer, and J.U. Wurthner, *The TGF-beta signaling modulators TRAP1/TGFBRAP1 and VPS39/Vam6/TLP are essential for early embryonic development.* Immunobiology, 2011. **216**(3): p. 343-50.

74. Smits, P., A.D. Bolton, V. Funari, M. Hong, E.D. Boyden, L. Lu, D.K. Manning, N.D. Dwyer, J.L. Moran, M. Prysak, B. Merriman, S.F. Nelson, L. Bonafe, A. Superti-Furga, S. Ikegawa, D. Krakow, D.H. Cohn, T. Kirchhausen, M.L. Warman, and D.R. Beier, *Lethal skeletal dysplasia in mice and humans lacking the golgin GMAP-210.* N Engl J Med, 2010. **362**(3): p. 206-16.

75. Christoforidis, S., H.M. McBride, R.D. Burgoyne, and M. Zerial, *The Rab5 effector EEA1 is a core component of endosome docking.* Nature, 1999. **397**(6720): p. 621-5.

76. Simonsen, A., R. Lippe, S. Christoforidis, J.M. Gaullier, A. Brech, J. Callaghan, B.H. Toh, C. Murphy, M. Zerial, and H. Stenmark, *EEA1 links PI(3)K function to Rab5 regulation of endosome fusion.* Nature, 1998. **394**(6692): p. 494-8.

77. Gillingham, A.K. and S. Munro, *Finding the Golgi: Golgin Coiled-Coil Proteins Show the Way.* Trends Cell Biol, 2016. **26**(6): p. 399-408.

78. Wong, M. and S. Munro, *The specificity of vesicle traffic to the Golgi is encoded in the golgin coiled-coil proteins.* Science, 2014. **346**(6209): p. 1256898-1256898.

79. Shin, J.J.H., A.K. Gillingham, F. Begum, J. Chadwick, and S. Munro, *TBC1D23 is a bridging factor for endosomal vesicle capture by golgins at the trans-Golgi.* Nat Cell Biol, 2017. **19**(12): p. 1424-1432.

80. Navarro Negredo, P., J.R. Edgar, P.T. Manna, R. Antrobus, and M.S. Robinson, *The WDR11 complex facilitates the tethering of AP-1-derived vesicles.* Nat Commun, 2018. **9**(1): p. 596.

81. Cheung, P.Y. and S.R. Pfeffer, *Transport Vesicle Tethering at the Trans Golgi Network: Coiled Coil Proteins in Action.* Front Cell Dev Biol, 2016. **4**: p. 18.

82. Reddy, J.V., A.S. Burguete, K. Sridevi, I.G. Ganley, R.M. Nottingham, and S.R. Pfeffer, *A functional role for the GCC185 golgin in mannose 6-phosphate receptor recycling.* Mol Biol Cell, 2006. **17**(10): p. 4353-63.

83. Stenmark, H., R. Aasland, B.H. Toh, and A. D'Arrigo, *Endosomal localization of the autoantigen EEA1 is mediated by a zinc-binding FYVE finger.* J Biol Chem, 1996. **271**(39): p. 24048-54.

84. McBride, H.M., V. Rybin, C. Murphy, A. Giner, R. Teasdale, and M. Zerial, *Oligomeric complexes link Rab5 effectors with NSF and drive membrane fusion via interactions between EEA1 and syntaxin 13.* Cell, 1999. **98**(3): p. 377-86.

85. Brunet, S. and M. Sacher, *Are all multisubunit tethering complexes bona fide tethers?* Traffic, 2014. **15**(11): p. 1282-7.

86. Cai, H., S. Yu, S. Menon, Y. Cai, D. Lazarova, C. Fu, K. Reinisch, J.C. Hay, and S. Ferro-Novick, *TRAPPI tethers COPII vesicles by binding the coat subunit Sec23.* Nature, 2007. **445**(7130): p. 941-4.

87. Yamasaki, A., S. Menon, S. Yu, J. Barrowman, T. Meerloo, V. Oorschot, J. Klumperman, A. Satoh, and S. Ferro-Novick, *mTrs130 is a component of a mammalian TRAPPII complex, a Rab1 GEF that binds to COPI-coated vesicles.* Mol Biol Cell, 2009. **20**(19): p. 4205-15.

88. Lynch-Day, M.A., D. Bhandari, S. Menon, J. Huang, H. Cai, C.R. Bartholomew, J.H. Brumell, S. Ferro-Novick, and D.J. Klionsky, *Trs85 directs a Ypt1 GEF, TRAPPIII, to the phagophore to promote autophagy.* Proc Natl Acad Sci U S A, 2010. **107**(17): p. 7811-6.

89. Chia, P.Z. and P.A. Gleeson, *Membrane tethering.* F1000Prime Rep, 2014. **6**: p. 74.

90. Balderhaar, H.J. and C. Ungermann, *CORVET and HOPS tethering complexes - coordinators of endosome and lysosome fusion.* J Cell Sci, 2013. **126**(Pt 6): p. 1307-16.

91. Peplowska, K., D.F. Markgraf, C.W. Ostrowicz, G. Bange, and C. Ungermann, *The CORVET tethering complex interacts with the yeast Rab5 homolog Vps21 and is involved in endo-lysosomal biogenesis.* Dev Cell, 2007. **12**(5): p. 739-50.

92. Markgraf, D.F., F. Ahnert, H. Arlt, M. Mari, K. Peplowska, N. Epp, J. Griffith, F. Reggiori, and C. Ungermann, *The CORVET subunit Vps8 cooperates with the Rab5 homolog Vps21 to induce clustering of late endosomal compartments.* Mol Biol Cell, 2009. **20**(24): p. 5276-89.

93. Price, A., D. Seals, W. Wickner, and C. Ungermann, *The docking stage of yeast vacuole fusion requires the transfer of proteins from a cis-SNARE complex to a Rab/Ypt protein.* J Cell Biol, 2000. **148**(6): p. 1231-8.

94. Abenza, J.F., A. Galindo, A. Pantazopoulou, C. Gil, V. de los Rios, and M.A. Penalva, *Aspergillus RabB Rab5 integrates acquisition of degradative identity with the long distance movement of early endosomes.* Mol Biol Cell, 2010. **21**(15): p. 2756-69.

95. Ho, R. and C. Stroupe, *The HOPS/class C Vps complex tethers membranes by binding to one Rab GTPase in each apposed membrane.* Mol Biol Cell, 2015. **26**(14): p. 2655-63.

96. Baker, R.W., P.D. Jeffrey, M. Zick, B.P. Phillips, W.T. Wickner, and F.M. Hughson, *A direct role for the Sec1/Munc18-family protein Vps33 as a template for SNARE assembly.* Science, 2015. **349**(6252): p. 1111-4.

97. Lobingier, B.T. and A.J. Merz, *Sec1/Munc18 protein Vps33 binds to SNARE domains and the quaternary SNARE complex.* Mol Biol Cell, 2012. **23**(23): p. 4611-22.

98. Dubuke, M.L. and M. Munson, *The Secret Life of Tethers: The Role of Tethering Factors in SNARE Complex Regulation.* Front Cell Dev Biol, 2016. **4**: p. 42.

99. Whiteheart, S.W., M. Brunner, D.W. Wilson, M. Wiedmann, and J.E. Rothman, *Soluble N-ethylmaleimide-sensitive fusion attachment proteins (SNAPs) bind to a*

multi-SNAP receptor complex in Golgi membranes. J Biol Chem, 1992. **267**(17): p. 12239-43.

100. Wilson, D.W., S.W. Whiteheart, M. Wiedmann, M. Brunner, and J.E. Rothman, *A multisubunit particle implicated in membrane fusion.* J Cell Biol, 1992. **117**(3): p. 531-8.

101. Weidman, P.J., P. Melancon, M.R. Block, and J.E. Rothman, *Binding of an N-ethylmaleimide-sensitive fusion protein to Golgi membranes requires both a soluble protein(s) and an integral membrane receptor.* J Cell Biol, 1989. **108**(5): p. 1589-96.

102. Weber, T., B.V. Zemelman, J.A. McNew, B. Westermann, M. Gmachl, F. Parlati, T.H. Sollner, and J.E. Rothman, *SNAREpins: minimal machinery for membrane fusion.* Cell, 1998. **92**(6): p. 759-72.

103. Liu, T., W.C. Tucker, A. Bhalla, E.R. Chapman, and J.C. Weisshaar, *SNARE-driven, 25-millisecond vesicle fusion in vitro.* Biophys J, 2005. **89**(4): p. 2458-72.

104. Kadkova, A., J. Radecke, and J.B. Sorensen, *The SNAP-25 Protein Family.* Neuroscience, 2018.

105. Gonzalo, S. and M.E. Linder, *SNAP-25 palmitoylation and plasma membrane targeting require a functional secretory pathway.* Mol Biol Cell, 1998. **9**(3): p. 585-97.

106. Fukasawa, M., O. Varlamov, W.S. Eng, T.H. Sollner, and J.E. Rothman, *Localization and activity of the SNARE Ykt6 determined by its regulatory domain and palmitoylation.* Proc Natl Acad Sci U S A, 2004. **101**(14): p. 4815-20.

107. Fasshauer, D., R.B. Sutton, A.T. Brunger, and R. Jahn, *Conserved structural features of the synaptic fusion complex: SNARE proteins reclassified as Q- and R-SNAREs.* Proc Natl Acad Sci U S A, 1998. **95**(26): p. 15781-6.

108. van den Bogaart, G., M.G. Holt, G. Bunt, D. Riedel, F.S. Wouters, and R. Jahn, *One SNARE complex is sufficient for membrane fusion.* Nat Struct Mol Biol, 2010. **17**(3): p. 358-64.

109. Pobbati, A.V., A. Stein, and D. Fasshauer, *N- to C-terminal SNARE complex assembly promotes rapid membrane fusion.* Science, 2006. **313**(5787): p. 673-6.

110. Hong, W., *SNAREs and traffic.* Biochim Biophys Acta, 2005. **1744**(3): p. 493-517.

111. Daste, F., T. Galli, and D. Tareste, *Structure and function of longin SNAREs.* J Cell Sci, 2015. **128**(23): p. 4263-72.

112. Collins, B.M., A.J. McCoy, H.M. Kent, P.R. Evans, and D.J. Owen, *Molecular architecture and functional model of the endocytic AP2 complex.* Cell, 2002. **109**(4): p. 523-35.

113. Hu, S.H., C.F. Latham, C.L. Gee, D.E. James, and J.L. Martin, *Structure of the Munc18c/Syntaxin4 N-peptide complex defines universal features of the N-peptide binding mode of Sec1/Munc18 proteins.* Proc Natl Acad Sci U S A, 2007. **104**(21): p. 8773-8.

114. Burkhardt, P., D.A. Hattendorf, W.I. Weis, and D. Fasshauer, *Munc18a controls SNARE assembly through its interaction with the syntaxin N-peptide.* EMBO J, 2008. **27**(7): p. 923-33.

115. ter Beest, M.B., S.J. Chapin, D. Avrahami, and K.E. Mostov, *The role of syntaxins in the specificity of vesicle targeting in polarized epithelial cells.* Mol Biol Cell, 2005. **16**(12): p. 5784-92.

116. Misura, K.M., R.H. Scheller, and W.I. Weis, *Three-dimensional structure of the neuronal-Sec1-syntaxin 1a complex.* Nature, 2000. **404**(6776): p. 355-62.

117. Hackmann, Y., S.C. Graham, S. Ehl, S. Honing, K. Lehmberg, M. Arico, D.J. Owen, and G.M. Griffiths, *Syntaxin binding mechanism and disease-causing mutations in Munc18-2.* Proc Natl Acad Sci U S A, 2013. **110**(47): p. E4482-91.

118. Jiao, J., M. He, S.A. Port, R.W. Baker, Y. Xu, H. Qu, Y. Xiong, Y. Wang, H. Jin, T.J. Eisemann, F.M. Hughson, and Y. Zhang, *Munc18-1 catalyzes neuronal SNARE assembly by templating SNARE association.* Elife, 2018. **7**.

119. Sudhof, T.C. and J.E. Rothman, *Membrane fusion: grappling with SNARE and SM proteins.* Science, 2009. **323**(5913): p. 474-7.

120. Orr, A., H. Song, S.F. Rusin, A.N. Kettenbach, and W. Wickner, *HOPS catalyzes the interdependent assembly of each vacuolar SNARE into a SNARE complex.* Mol Biol Cell, 2017. **28**(7): p. 975-983.

121. D'Agostino, M., H.J. Risselada, A. Lurick, C. Ungermann, and A. Mayer, *A tethering complex drives the terminal stage of SNARE-dependent membrane fusion.* Nature, 2017. **551**(7682): p. 634-638.

122. Sollner, T., S.W. Whiteheart, M. Brunner, H. Erdjument-Bromage, S. Geromanos, P. Tempst, and J.E. Rothman, *SNAP receptors implicated in vesicle targeting and fusion.* Nature, 1993. **362**(6418): p. 318-24.

123. Littleton, J.T., R.J. Barnard, S.A. Titus, J. Slind, E.R. Chapman, and B. Ganetzky, *SNARE-complex disassembly by NSF follows synaptic-vesicle fusion.* Proc Natl Acad Sci U S A, 2001. **98**(21): p. 12233-8.

124. Ma, L., Y. Kang, J. Jiao, A.A. Rebane, H.K. Cha, Z. Xi, H. Qu, and Y. Zhang, *alpha-SNAP Enhances SNARE Zippering by Stabilizing the SNARE Four-Helix Bundle.* Cell Rep, 2016. **15**(3): p. 531-539.

125. Uchiyama, K., E. Jokitalo, F. Kano, M. Murata, X. Zhang, B. Canas, R. Newman, C. Rabouille, D. Pappin, P. Freemont, and H. Kondo, *VCIP135, a novel essential factor for p97/p47-mediated membrane fusion, is required for Golgi and ER assembly in vivo.* J Cell Biol, 2002. **159**(5): p. 855-66.

126. Kent, H.M., P.R. Evans, I.B. Schäfer, S.R. Gray, C.M. Sanderson, J.P. Luzio, A.A. Peden, and D.J. Owen, *Structural Basis of the Intracellular Sorting of the SNARE VAMP7 by the AP3 Adaptor Complex.* Developmental Cell, 2012. **22**(5): p. 979-988.

127. Martinez-Arca, S., P. Alberts, A. Zahraoui, D. Louvard, and T. Galli, *Role of tetanus neurotoxin insensitive vesicle-associated membrane protein (TI-VAMP) in vesicular transport mediating neurite outgrowth.* J Cell Biol, 2000. **149**(4): p. 889-900.

128. Chaineau, M., L. Danglot, and T. Galli, *Multiple roles of the vesicular-SNARE TI-VAMP in post-Golgi and endosomal trafficking.* FEBS Lett, 2009. **583**(23): p. 3817-26.

129. Schäfer, I.B., G.G. Hesketh, N.A. Bright, S.R. Gray, P.R. Pryor, P.R. Evans, J.P. Luzio, and D.J. Owen, *The binding of Varp to VAMP7 traps VAMP7 in a closed, fusogenically inactive conformation*. 2012.

130. Vivona, S., C.W. Liu, P. Strop, V. Rossi, F. Filippini, and A.T. Brunger, *The longin SNARE VAMP7/TI-VAMP adopts a closed conformation.* The Journal of biological chemistry, 2010. **285**(23): p. 17965-73.

131. Martinez-Arca, S., R. Rudge, M. Vacca, G. Raposo, J. Camonis, V. Proux-Gillardeaux, L. Daviet, E. Formstecher, A. Hamburger, F. Filippini, M. D'Esposito, and T. Galli, *A dual mechanism controlling the localization and function of exocytic v-SNAREs.* Proc Natl Acad Sci U S A, 2003. **100**(15): p. 9011-6.

132. Wassmer, T., N. Attar, M.V. Bujny, J. Oakley, C.J. Traer, and P.J. Cullen, *A loss-of-function screen reveals SNX5 and SNX6 as potential components of the mammalian retromer.* J Cell Sci, 2007. **120**(Pt 1): p. 45-54.

133. Swarbrick, J.D., D.J. Shaw, S. Chhabra, R. Ghai, E. Valkov, S.J. Norwood, M.N.J. Seaman, and B.M. Collins, *VPS29 is not an active metallo-phosphatase but is a rigid scaffold required for retromer interaction with accessory proteins.* PLoS ONE, 2011. **6**(5).

134. Harbour, M.E., S.Y. Breusegem, R. Antrobus, C. Freeman, E. Reid, and M.N. Seaman, *The cargo-selective retromer complex is a recruiting hub for protein complexes that regulate endosomal tubule dynamics.* J Cell Sci, 2010. **123**(Pt 21): p. 3703-17.

135. Bonifacino, J.S. and J.H. Hurley, *Retromer.* Curr Opin Cell Biol, 2008. **20**(4): p. 427-36.

136. Attar, N. and P.J. Cullen, *The retromer complex.* Adv Enzyme Regul, 2010. **50**(1): p. 216-36.

137. Hierro, A., A.L. Rojas, R. Rojas, N. Murthy, G. Effantin, A.V. Kajava, A.C. Steven, J.S. Bonifacino, and J.H. Hurley, *Functional architecture of the retromer cargo-recognition complex.* Nature, 2007. **449**(7165): p. 1063-7.

138. Pourcher, M., M. Santambrogio, N. Thazar, A.M. Thierry, I. Fobis-Loisy, C. Miege, Y. Jaillais, and T. Gaude, *Analyses of sorting nexins reveal distinct retromer-subcomplex functions in development and protein sorting in Arabidopsis thaliana.* Plant Cell, 2010. **22**(12): p. 3980-91.

139. Nisar, S., E. Kelly, P.J. Cullen, and S.J. Mundell, *Regulation of P2Y1 receptor traffic by sorting Nexin 1 is retromer independent.* Traffic, 2010. **11**(4): p. 508-19.

140. Wang, Y., Y. Zhou, K. Szabo, C.R. Haft, and J. Trejo, *Down-regulation of protease-activated receptor-1 is regulated by sorting nexin 1.* Mol Biol Cell, 2002. **13**(6): p. 1965-76.

141. Heydorn, A., B.P. Sondergaard, B. Ersboll, B. Holst, F.C. Nielsen, C.R. Haft, J. Whistler, and T.W. Schwartz, *A library of 7TM receptor C-terminal tails. Interactions with the proposed post-endocytic sorting proteins ERM-binding phosphoprotein 50 (EBP50), N-ethylmaleimide-sensitive factor (NSF), sorting nexin 1 (SNX1), and G protein-coupled receptor-associated sorting protein (GASP).* J Biol Chem, 2004. **279**(52): p. 54291-303.

142. Lauffer, B.E., C. Melero, P. Temkin, C. Lei, W. Hong, T. Kortemme, and M. von Zastrow, *SNX27 mediates PDZ-directed sorting from endosomes to the plasma membrane.* J Cell Biol, 2010. **190**(4): p. 565-74.

143. Harterink, M., F. Port, M.J. Lorenowicz, I.J. McGough, M. Silhankova, M.C. Betist, J.R.T. van Weering, R. van Heesbeen, T.C. Middelkoop, K. Basler, P.J. Cullen, and H.C. Korswagen, *A SNX3-dependent retromer pathway mediates retrograde transport of the Wnt sorting receptor Wntless and is required for Wnt secretion.* Nat Cell Biol, 2011. **13**(8): p. 914-923.

144. Gallon, M. and P.J. Cullen, *Retromer and sorting nexins in endosomal sorting.* Biochem Soc Trans, 2015. **43**(1): p. 33-47.

145. Kerr, M.C., J.S. Bennetts, F. Simpson, E.C. Thomas, C. Flegg, P.A. Gleeson, C. Wicking, and R.D. Teasdale, *A novel mammalian retromer component, Vps26B.* Traffic, 2005. **6**(11): p. 991-1001.

146. Haft, C.R., M. de la Luz Sierra, R. Bafford, M.A. Lesniak, V.A. Barr, and S.I. Taylor, *Human orthologs of yeast vacuolar protein sorting proteins Vps26, 29, and 35: assembly into multimeric complexes.* Mol Biol Cell, 2000. **11**(12): p. 4105-16.

147. Collins, B.M., C.F. Skinner, P.J. Watson, M.N.J. Seaman, and D.J. Owen, *Vps29 has a phosphoesterase fold that acts as a protein interaction scaffold for retromer assembly.* Nature structural & molecular biology, 2005. **12**(7): p. 594-602.

148. Lucas, M., D.C. Gershlick, A. Vidaurrazaga, A.L. Rojas, J.S. Bonifacino, and A. Hierro, *Structural Mechanism for Cargo Recognition by the Retromer Complex.* Cell, 2016. **167**(6): p. 1623-1635 e14.

149. Kovtun, O., N. Leneva, Y.S. Bykov, N. Ariotti, R.D. Teasdale, M. Schaffer, B.D. Engel, D.J. Owen, J.A.G. Briggs, and B.M. Collins, *Structure of the membrane-assembled retromer coat determined by cryo-electron tomography.* Nature, 2018. **561**(7724): p. 561-564.

150. Shi, H., R. Rojas, J.S. Bonifacino, and J.H. Hurley, *The retromer subunit Vps26 has an arrestin fold and binds Vps35 through its C-terminal domain.* Nat Struct Mol Biol, 2006. **13**(6): p. 540-8.

151. Collins, B.M., S.J. Norwood, M.C. Kerr, D. Mahony, M.N. Seaman, R.D. Teasdale, and D.J. Owen, *Structure of Vps26B and mapping of its interaction with the retromer protein complex.* Traffic, 2008. **9**(3): p. 366-79.

152. Gurevich, V.V. and E.V. Gurevich, *The structural basis of arrestin-mediated regulation of G-protein-coupled receptors.* Pharmacol Ther, 2006. **110**(3): p. 465-502.

153. Wang, D., M. Guo, Z. Liang, J. Fan, Z. Zhu, J. Zang, Z. Zhu, X. Li, M. Teng, L. Niu, Y. Dong, and P. Liu, *Crystal structure of human vacuolar protein sorting protein 29 reveals a phosphodiesterase/nuclease-like fold and two protein-protein interaction sites.* J Biol Chem, 2005. **280**(24): p. 22962-7.

154. Damen, E., E. Krieger, J.E. Nielsen, J. Eygensteyn, and J.E. van Leeuwen, *The human Vps29 retromer component is a metallo-phosphoesterase for a cation-independent mannose 6-phosphate receptor substrate peptide.* Biochem J, 2006. **398**(3): p. 399-409.

155. Nothwehr, S.F., S.A. Ha, and P. Bruinsma, *Sorting of yeast membrane proteins into an endosome-to-Golgi pathway involves direct interaction of their cytosolic domains with Vps35p.* J Cell Biol, 2000. **151**(2): p. 297-310.

156. McCormick, P.J., K. Dumaresq-Doiron, A.S. Pluviose, V. Pichette, G. Tosato, and S. Lefrancois, *Palmitoylation controls recycling in lysosomal sorting and trafficking.* Traffic, 2008. **9**(11): p. 1984-97.

157. Fjorback, A.W., M. Seaman, C. Gustafsen, A. Mehmedbasic, S. Gokool, C. Wu, D. Militz, V. Schmidt, P. Madsen, J.R. Nyengaard, T.E. Willnow, E.I. Christensen, W.B. Mobley, A. Nykjaer, and O.M. Andersen, *Retromer binds the FANSHY sorting motif in SorLA to regulate amyloid precursor protein sorting and processing.* J Neurosci, 2012. **32**(4): p. 1467-80.

158. Seaman, M.N., *Identification of a novel conserved sorting motif required for retromer-mediated endosome-to-TGN retrieval.* J Cell Sci, 2007. **120**(Pt 14): p. 2378-89.

159. Simonetti, B., C.M. Danson, K.J. Heesom, and P.J. Cullen, *Sequence-dependent cargo recognition by SNX-BARs mediates retromer-independent transport of CI-MPR.* J Cell Biol, 2017. **216**(11): p. 3695-3712.

160. Kvainickas, A., A. Jimenez-Orgaz, H. Nagele, Z. Hu, J. Dengjel, and F. Steinberg, *Cargo-selective SNX-BAR proteins mediate retromer trimer independent retrograde transport.* J Cell Biol, 2017. **216**(11): p. 3677-3693.

161. Cui, Y., J.M. Carosi, Z. Yang, N. Ariotti, M.C. Kerr, R.G. Parton, T.J. Sargeant, and R.D. Teasdale, *Retromer has a selective function in cargo sorting via endosome transport carriers.* J Cell Biol, 2019. **218**(2): p. 615-631.

162. Seaman, M.N., M.E. Harbour, D. Tattersall, E. Read, and N. Bright, *Membrane recruitment of the cargo-selective retromer subcomplex is catalysed by the small GTPase Rab7 and inhibited by the Rab-GAP TBC1D5.* J Cell Sci, 2009. **122**(Pt 14): p. 2371-82.

163. Yao, J., F. Yang, X. Sun, S. Wang, N. Gan, Q. Liu, D. Liu, X. Zhang, D. Niu, Y. Wei, C. Ma, Z.Q. Luo, Q. Sun, and D. Jia, *Mechanism of inhibition of retromer transport by the bacterial effector RidL.* Proc Natl Acad Sci U S A, 2018. **115**(7): p. E1446-E1454.

164. Harbour, M.E. and M.N.J. Seaman, *Evolutionary variations of VPS29, and their implications for the heteropentameric model of retromer.* Communicative & integrative biology, 2011. **4**(5): p. 619-22.

165. Seaman, M.N.J., A.S. Mukadam, and S.Y. Breusegem, *Inhibition of TBC1D5 activates Rab7a and can enhance the function of the retromer cargo-selective complex.* J Cell Sci, 2018.

166. Rojas, R., T. van Vlijmen, G.A. Mardones, Y. Prabhu, A.L. Rojas, S. Mohammed, A.J. Heck, G. Raposo, P. van der Sluijs, and J.S. Bonifacino, *Regulation of retromer recruitment to endosomes by sequential action of Rab5 and Rab7.* J Cell Biol, 2008. **183**(3): p. 513-26.

167. Liu, T.T., T.S. Gomez, B.K. Sackey, D.D. Billadeau, and C.G. Burd, *Rab GTPase regulation of retromer-mediated cargo export during endosome maturation.* Mol Biol Cell, 2012. **23**(13): p. 2505-15.

168. Priya, A., I.V. Kalaidzidis, Y. Kalaidzidis, D. Lambright, and S. Datta, *Molecular insights into Rab7-mediated endosomal recruitment of core retromer: deciphering the role of Vps26 and Vps35.* Traffic, 2015. **16**(1): p. 68-84.

169. Balderhaar, H.J., H. Arlt, C. Ostrowicz, C. Brocker, F. Sundermann, R. Brandt, M. Babst, and C. Ungermann, *The Rab GTPase Ypt7 is linked to retromer-mediated receptor recycling and fusion at the yeast late endosome.* J Cell Sci, 2010. **123**(Pt 23): p. 4085-94.

170. Derivery, E., C. Sousa, J.J. Gautier, B. Lombard, D. Loew, and A. Gautreau, *The Arp2/3 activator WASH controls the fission of endosomes through a large multiprotein complex.* Dev Cell, 2009. **17**(5): p. 712-23.

171. Derivery, E. and A. Gautreau, *Evolutionary conservation of the WASH complex, an actin polymerization machine involved in endosomal fission.* Commun Integr Biol, 2010. **3**(3): p. 227-30.

172. Harbour, M.E., S.Y. Breusegem, and M.N. Seaman, *Recruitment of the endosomal WASH complex is mediated by the extended 'tail' of Fam21 binding to the retromer protein Vps35.* Biochem J, 2012. **442**(1): p. 209-20.

173. Gokool, S., D. Tattersall, and M.N. Seaman, *EHD1 interacts with retromer to stabilize SNX1 tubules and facilitate endosome-to-Golgi retrieval.* Traffic, 2007. **8**(12): p. 1873-86.

174. Rojas, R., S. Kametaka, C.R. Haft, and J.S. Bonifacino, *Interchangeable but essential functions of SNX1 and SNX2 in the association of retromer with endosomes and the trafficking of mannose 6-phosphate receptors.* Mol Cell Biol, 2007. **27**(3): p. 1112-24.

175. Qualmann, B., D. Koch, and M.M. Kessels, *Let's go bananas: revisiting the endocytic BAR code.* EMBO J, 2011. **30**(17): p. 3501-15.

176. Nishimura, T., N. Morone, and S. Suetsugu, *Membrane re-modelling by BAR domain superfamily proteins via molecular and non-molecular factors.* Biochem Soc Trans, 2018. **46**(2): p. 379-389.

177. van Weering, J.R., P. Verkade, and P.J. Cullen, *SNX-BAR proteins in phosphoinositide-mediated, tubular-based endosomal sorting.* Semin Cell Dev Biol, 2010. **21**(4): p. 371-80.

178. Freeman, C.L., G. Hesketh, and M.N. Seaman, *RME-8 coordinates the activity of the WASH complex with the function of the retromer SNX dimer to control endosomal tubulation.* J Cell Sci, 2014. **127**(Pt 9): p. 2053-70.

179. Popoff, V., G.A. Mardones, S.K. Bai, V. Chambon, D. Tenza, P.V. Burgos, A. Shi, P. Benaroch, S. Urbe, C. Lamaze, B.D. Grant, G. Raposo, and L. Johannes, *Analysis of articulation between clathrin and retromer in retrograde sorting on early endosomes.* Traffic, 2009. **10**(12): p. 1868-80.

180. Griffin, C.T., J. Trejo, and T. Magnuson, *Genetic evidence for a mammalian retromer complex containing sorting nexins 1 and 2.* Proc Natl Acad Sci U S A, 2005. **102**(42): p. 15173-7.

181. Wassmer, T., N. Attar, M. Harterink, J.R. van Weering, C.J. Traer, J. Oakley, B. Goud, D.J. Stephens, P. Verkade, H.C. Korswagen, and P.J. Cullen, *The retromer coat*

complex coordinates endosomal sorting and dynein-mediated transport, with carrier recognition by the trans-Golgi network. Dev Cell, 2009. **17**(1): p. 110-22.

182. Harrison, M.S., C.S. Hung, T.T. Liu, R. Christiano, T.C. Walther, and C.G. Burd, *A mechanism for retromer endosomal coat complex assembly with cargo.* Proc Natl Acad Sci U S A, 2014. **111**(1): p. 267-72.

183. Steinberg, F., M. Gallon, M. Winfield, E.C. Thomas, A.J. Bell, K.J. Heesom, J.M. Tavare, and P.J. Cullen, *A global analysis of SNX27-retromer assembly and cargo specificity reveals a function in glucose and metal ion transport.* Nat Cell Biol, 2013. **15**(5): p. 461-71.

184. Gallon, M., T. Clairfeuille, F. Steinberg, C. Mas, R. Ghai, R.B. Sessions, R.D. Teasdale, B.M. Collins, and P.J. Cullen, *A unique PDZ domain and arrestin-like fold interaction reveals mechanistic details of endocytic recycling by SNX27-retromer.* Proc Natl Acad Sci U S A, 2014. **111**(35): p. E3604-13.

185. Temkin, P., B. Lauffer, S. Jager, P. Cimermancic, N.J. Krogan, and M. von Zastrow, *SNX27 mediates retromer tubule entry and endosome-to-plasma membrane trafficking of signalling receptors.* Nat Cell Biol, 2011. **13**(6): p. 715-21.

186. Lee, S., J. Chang, and C. Blackstone, *FAM21 directs SNX27-retromer cargoes to the plasma membrane by preventing transport to the Golgi apparatus.* Nat Commun, 2016. **7**: p. 10939.

187. Pfeffer, S.R., *A nexus for receptor recycling.* Nat Cell Biol, 2013. **15**(5): p. 446-8.

188. Seaman, M.N.J., *Retromer and the cation-independent mannose 6-phosphate receptor-Time for a trial separation?* Traffic, 2018. **19**(2): p. 150-152.

189. Hirst, J. and M.S. Robinson, *Clathrin and adaptors.* Biochim Biophys Acta, 1998. **1404**(1-2): p. 173-93.

190. Robinson, M.S., *Adaptable adaptors for coated vesicles.* Trends Cell Biol, 2004. **14**(4): p. 167-74.

191. Puertollano, R., R.C. Aguilar, I. Gorshkova, R.J. Crouch, and J.S. Bonifacino, *Sorting of mannose 6-phosphate receptors mediated by the GGAs.* Science, 2001. **292**(5522): p. 1712-6.

192. Zhu, Y., B. Doray, A. Poussu, V.P. Lehto, and S. Kornfeld, *Binding of GGA2 to the lysosomal enzyme sorting motif of the mannose 6-phosphate receptor.* Science, 2001. **292**(5522): p. 1716-8.

193. Doray, B., K. Bruns, P. Ghosh, and S. Kornfeld, *Interaction of the cation-dependent mannose 6-phosphate receptor with GGA proteins.* J Biol Chem, 2002. **277**(21): p. 18477-82.

194. Hinners, I. and S.A. Tooze, *Changing directions: clathrin-mediated transport between the Golgi and endosomes.* J Cell Sci, 2003. **116**(Pt 5): p. 763-71.

195. Meyer, C., D. Zizioli, S. Lausmann, E.L. Eskelinen, J. Hamann, P. Saftig, K. von Figura, and P. Schu, *mu1A-adaptin-deficient mice: lethality, loss of AP-1 binding and rerouting of mannose 6-phosphate receptors.* EMBO J, 2000. **19**(10): p. 2193-203.

196. Saint-Pol, A., B. Yelamos, M. Amessou, I.G. Mills, M. Dugast, D. Tenza, P. Schu, C. Antony, H.T. McMahon, C. Lamaze, and L. Johannes, *Clathrin adaptor epsinR is*

required for retrograde sorting on early endosomal membranes. Dev Cell, 2004. **6**(4): p. 525-38.

197. Jia, D., J.S. Zhang, F. Li, J. Wang, Z. Deng, M.A. White, D.G. Osborne, C. Phillips-Krawczak, T.S. Gomez, H. Li, A. Singla, E. Burstein, D.D. Billadeau, and M.K. Rosen, *Structural and mechanistic insights into regulation of the retromer coat by TBC1d5.* Nat Commun, 2016. **7**: p. 13305.

198. McNally, K.E., R. Faulkner, F. Steinberg, M. Gallon, R. Ghai, D. Pim, P. Langton, N. Pearson, C.M. Danson, H. Nagele, L.L. Morris, A. Singla, B.L. Overlee, K.J. Heesom, R. Sessions, L. Banks, B.M. Collins, I. Berger, D.D. Billadeau, E. Burstein, and P.J. Cullen, *Retriever is a multiprotein complex for retromer-independent endosomal cargo recycling.* Nat Cell Biol, 2017. **19**(10): p. 1214-1225.

199. Steinberg, F., K.J. Heesom, M.D. Bass, and P.J. Cullen, *SNX17 protects integrins from degradation by sorting between lysosomal and recycling pathways.* J Cell Biol, 2012. **197**(2): p. 219-30.

200. Jia, D., T.S. Gomez, D.D. Billadeau, and M.K. Rosen, *Multiple repeat elements within the FAM21 tail link the WASH actin regulatory complex to the retromer.* Molecular Biology of the Cell, 2012. **23**(12): p. 2352-2361.

201. Lee, J., C. Retamal, L. Cuitino, A. Caruano-Yzermans, J.E. Shin, P. van Kerkhof, M.P. Marzolo, and G. Bu, *Adaptor protein sorting nexin 17 regulates amyloid precursor protein trafficking and processing in the early endosomes.* J Biol Chem, 2008. **283**(17): p. 11501-8.

202. Small, S.A., K. Kent, A. Pierce, C. Leung, M.S. Kang, H. Okada, L. Honig, J.P. Vonsattel, and T.W. Kim, *Model-guided microarray implicates the retromer complex in Alzheimer's disease.* Ann Neurol, 2005. **58**(6): p. 909-19.

203. Vardarajan, B.N., S.Y. Bruesegem, M.E. Harbour, R. Inzelberg, R. Friedland, P. St George-Hyslop, M.N. Seaman, and L.A. Farrer, *Identification of Alzheimer disease-associated variants in genes that regulate retromer function.* Neurobiol Aging, 2012. **33**(9): p. 2231 e15-2231 e30.

204. Reitz, C., G. Tosto, B. Vardarajan, E. Rogaeva, M. Ghani, R.S. Rogers, C. Conrad, J.L. Haines, M.A. Pericak-Vance, M.D. Fallin, T. Foroud, L.A. Farrer, G.D. Schellenberg, P.S. George-Hyslop, R. Mayeux, and C. Alzheimer's Disease Genetics, *Independent and epistatic effects of variants in VPS10-d receptors on Alzheimer disease risk and processing of the amyloid precursor protein (APP).* Transl Psychiatry, 2013. **3**: p. e256.

205. Lane, R.F., S.M. Raines, J.W. Steele, M.E. Ehrlich, J.A. Lah, S.A. Small, R.E. Tanzi, A.D. Attie, and S. Gandy, *Diabetes-associated SorCS1 regulates Alzheimer's amyloid-beta metabolism: evidence for involvement of SorL1 and the retromer complex.* J Neurosci, 2010. **30**(39): p. 13110-5.

206. Burd, C.G., *Physiology and pathology of endosome-to-Golgi retrograde sorting.* Traffic, 2011. **12**(8): p. 948-55.

207. Brodin, L. and O. Shupliakov, *Retromer in Synaptic Function and Pathology.* Front Synaptic Neurosci, 2018. **10**: p. 37.

208. McMillan, K.J., H.C. Korswagen, and P.J. Cullen, *The emerging role of retromer in neuroprotection.* Curr Opin Cell Biol, 2017. **47**: p. 72-82.

209. Cui, Y., Z. Yang, and R.D. Teasdale, *The functional roles of retromer in Parkinson's disease.* FEBS Lett, 2018. **592**(7): p. 1096-1112.

210. Spoelgen, R., C.A. von Arnim, A.V. Thomas, I.D. Peltan, M. Koker, A. Deng, M.C. Irizarry, O.M. Andersen, T.E. Willnow, and B.T. Hyman, *Interaction of the cytosolic domains of sorLA/LR11 with the amyloid precursor protein (APP) and beta-secretase beta-site APP-cleaving enzyme.* J Neurosci, 2006. **26**(2): p. 418-28.

211. Ohbayashi, N., A. Yatsu, K. Tamura, and M. Fukuda, *The Rab21-GEF activity of Varp, but not its Rab32/38 effector function, is required for dendrite formation in melanocytes*. 2012. p. 669-678.

212. Fukuda, M., *Multiple Roles of VARP in Endosomal Trafficking: Rabs, Retromer Components and R-SNARE VAMP7 Meet on VARP.* Traffic, 2016. **17**(7): p. 709-19.

213. Mori, Y., T. Matsui, and M. Fukuda, *Rabex-5 protein regulates dendritic localization of small GTPase Rab17 and neurite morphogenesis in hippocampal neurons.* J Biol Chem, 2013. **288**(14): p. 9835-47.

214. Simpson, J.C., G. Griffiths, M. Wessling-Resnick, J.A. Fransen, H. Bennett, and A.T. Jones, *A role for the small GTPase Rab21 in the early endocytic pathway.* J Cell Sci, 2004. **117**(Pt 26): p. 6297-311.

215. Pellinen, T., A. Arjonen, K. Vuoriluoto, K. Kallio, J.A. Fransen, and J. Ivaska, *Small GTPase Rab21 regulates cell adhesion and controls endosomal traffic of beta1-integrins.* J Cell Biol, 2006. **173**(5): p. 767-80.

216. Tamura, K., N. Ohbayashi, K. Ishibashi, and M. Fukuda, *Structure-function analysis of VPS9-ankyrin-repeat protein (Varp) in the trafficking of tyrosinase-related protein 1 in melanocytes.* The Journal of biological chemistry, 2011. **286**(9): p. 7507-21.

217. Loftus, S.K., D.M. Larson, L.L. Baxter, A. Antonellis, Y. Chen, X. Wu, Y. Jiang, M. Bittner, J.A. Hammer, 3rd, and W.J. Pavan, *Mutation of melanosome protein RAB38 in chocolate mice.* Proc Natl Acad Sci U S A, 2002. **99**(7): p. 4471-6.

218. Dennis, M.K., C. Delevoye, A. Acosta-Ruiz, I. Hurbain, M. Romao, G.G. Hesketh, P.S. Goff, E.V. Sviderskaya, D.C. Bennett, J.P. Luzio, T. Galli, D.J. Owen, G. Raposo, and M.S. Marks, *BLOC-1 and BLOC-3 regulate VAMP7 cycling to and from melanosomes via distinct tubular transport carriers.* J Cell Biol, 2016. **214**(3): p. 293-308.

219. Bean, B.D.M., M. Davey, J. Snider, M. Jessulat, V. Deineko, M. Tinney, I. Stagljar, M. Babu, and E. Conibear, *Rab5-family guanine nucleotide exchange factors bind retromer and promote its recruitment to endosomes.* Molecular biology of the cell, 2015. **26**(6): p. 1119-28.

220. Krishna, S.S., I. Majumdar, and N.V. Grishin, *Structural classification of zinc fingers: survey and summary.* Nucleic Acids Res, 2003. **31**(2): p. 532-50.

221. Ravasi, T., T. Huber, M. Zavolan, A. Forrest, T. Gaasterland, S. Grimmond, D.A. Hume, R.G. Group, and G.S.L. Members, *Systematic characterization of the zinc-finger-containing proteins in the mouse transcriptome.* Genome Res, 2003. **13**(6B): p. 1430-42.

222. Andreini, C., I. Bertini, and G. Cavallaro, *Minimal functional sites allow a classification of zinc sites in proteins.* PLoS One, 2011. **6**(10): p. e26325.

223. Schwieters, C.D., J.J. Kuszewski, N. Tjandra, and G.M. Clore, *The Xplor-NIH NMR molecular structure determination package.* J Magn Reson, 2003. **160**(1): p. 65-73.

224. Eustermann, S., W.F. Wu, M.F. Langelier, J.C. Yang, L.E. Easton, A.A. Riccio, J.M. Pascal, and D. Neuhaus, *Structural Basis of Detection and Signaling of DNA Single-Strand Breaks by Human PARP-1.* Mol Cell, 2015. **60**(5): p. 742-754.

225. Williamson, M.P., *Using chemical shift perturbation to characterise ligand binding.* Progress in Nuclear Magnetic Resonance Spectroscopy, 2013. **73**: p. 1-16.

226. Wishart, D.S., C.G. Bigam, J. Yao, F. Abildgaard, H.J. Dyson, E. Oldfield, J.L. Markley, and B.D. Sykes, *1H, 13C and 15N chemical shift referencing in biomolecular NMR.* J Biomol NMR, 1995. **6**(2): p. 135-40.

227. Vranken, W.F., W. Boucher, T.J. Stevens, R.H. Fogh, A. Pajon, M. Llinas, E.L. Ulrich, J.L. Markley, J. Ionides, and E.D. Laue, *The CCPN data model for NMR spectroscopy: development of a software pipeline.* Proteins, 2005. **59**(4): p. 687-96.

228. Kuszewski, J., A.M. Gronenborn, and G.M. Clore, *Improving the quality of NMR and crystallographic protein structures by means of a conformational database potential derived from structure databases.* Protein Sci, 1996. **5**(6): p. 1067-80.

229. Hommel, U., T.S. Harvey, P.C. Driscoll, and I.D. Campbell, *Human epidermal growth factor. High resolution solution structure and comparison with human transforming growth factor alpha.* J Mol Biol, 1992. **227**(1): p. 271-82.

230. Diamond, R., *Coordinate-based cluster analysis.* Acta Crystallogr D Biol Crystallogr, 1995. **51**(Pt 2): p. 127-35.

231. Schrodinger, LLC, *The PyMOL Molecular Graphics System, Version 1.8*. 2015.

232. Fischer, M.J., *Amine coupling through EDC/NHS: a practical approach.* Methods Mol Biol, 2010. **627**: p. 55-73.

233. Fields, S. and O. Song, *A novel genetic system to detect protein-protein interactions.* Nature, 1989. **340**(6230): p. 245-6.

234. Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman, *Basic local alignment search tool.* J Mol Biol, 1990. **215**(3): p. 403-10.

235. States, D.J. and W. Gish, *Combined use of sequence similarity and codon bias for coding region identification.* J Comput Biol, 1994. **1**(1): p. 39-50.

236. Artimo, P., M. Jonnalagedda, K. Arnold, D. Baratin, G. Csardi, E. de Castro, S. Duvaud, V. Flegel, A. Fortier, E. Gasteiger, A. Grosdidier, C. Hernandez, V. Ioannidis, D. Kuznetsov, R. Liechti, S. Moretti, K. Mostaguir, N. Redaschi, G. Rossier, I. Xenarios, and H. Stockinger, *ExPASy: SIB bioinformatics resource portal.* Nucleic Acids Res, 2012. **40**(Web Server issue): p. W597-603.

237. Williamson, M.P., *Using chemical shift perturbation to characterise ligand binding.* Prog Nucl Magn Reson Spectrosc, 2013. **73**: p. 1-16.

238. Buchan, D.M., F; Nugent, TCO; Bryson, K; Jones, DT., *Scalable web services for the PSIPRED Protein Analaysis Workbench.* Nucleic Acids Research. p. W340-W348.

239. Wüthrich, K., *NMR of Proteins and Nucleic Acids*. 1986, New York, NY: Wiley.

240. Schubert, M., D. Labudde, H. Oschkinat, and P. Schmieder, *A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on 13C chemical shift statistics.* J Biomol NMR, 2002. **24**(2): p. 149-54.

241. Powers, R., D.S. Garrett, C.J. March, E.A. Frieden, A.M. Gronenborn, and G.M. Clore, *The high-resolution, three-dimensional solution structure of human interleukin-4 determined by multidimensional heteronuclear magnetic resonance spectroscopy.* Biochemistry, 1993. **32**(26): p. 6744-62.

242. Herman, E.K., M. Ali, M.C. Field, and J.B. Dacks, *Regulation of early endosomes across eukaryotes: Evolution and functional homology of Vps9 proteins.* Traffic, 2018. **19**(7): p. 546-563.

243. Bultema, J.J., A.L. Ambrosio, C.L. Burek, and S.M. Di Pietro, *BLOC-2, AP-3, and AP-1 proteins function in concert with Rab38 and Rab32 proteins to mediate protein trafficking to lysosome-related organelles.* J Biol Chem, 2012. **287**(23): p. 19550-63.

244. Kobayashi, T., S. Kim, Y.C. Lin, T. Inoue, and B.D. Dynlacht, *The CP110-interacting proteins Talpid3 and Cep290 play overlapping and distinct roles in cilia assembly.* J Cell Biol, 2014. **204**(2): p. 215-29.

245. Davey, M.G., I.R. Paton, Y. Yin, M. Schmidt, F.K. Bangs, D.R. Morrice, T.G. Smith, P. Buxton, D. Stamataki, M. Tanaka, A.E. Munsterberg, J. Briscoe, C. Tickle, and D.W. Burt, *The chicken talpid3 gene encodes a novel protein essential for Hedgehog signaling.* Genes Dev, 2006. **20**(10): p. 1365-77.

246. Yin, Y., F. Bangs, I.R. Paton, A. Prescott, J. James, M.G. Davey, P. Whitley, G. Genikhovich, U. Technau, D.W. Burt, and C. Tickle, *The Talpid3 gene (KIAA0586) encodes a centrosomal protein that is essential for primary cilia formation.* Development, 2009. **136**(4): p. 655-64.

247. Szalinski, C.M., A. Labilloy, J.R. Bruns, and O.A. Weisz, *VAMP7 modulates ciliary biogenesis in kidney cells.* PLoS One, 2014. **9**(1): p. e86425.

248. Kandachar, V., B.M. Tam, O.L. Moritz, and D. Deretic, *An interaction network between the SNARE VAMP7 and Rab GTPases within a ciliary membrane-targeting complex.* J Cell Sci, 2018. **131**(24).

249. Garcia, G., 3rd, D.R. Raleigh, and J.F. Reiter, *How the Ciliary Membrane Is Organized Inside-Out to Communicate Outside-In.* Curr Biol, 2018. **28**(8): p. R421-R434.

250. Yee, L.E. and J.F. Reiter, *Ciliary vesicle formation: a prelude to ciliogenesis.* Dev Cell, 2015. **32**(6): p. 665-6.

251. Garcia-Gonzalo, F.R. and J.F. Reiter, *Scoring a backstage pass: mechanisms of ciliogenesis and ciliary access.* J Cell Biol, 2012. **197**(6): p. 697-709.

252. Ojeda Naharros, I., F.B. Cristian, J. Zang, M. Gesemann, P.W. Ingham, S.C.F. Neuhauss, and R. Bachmann-Gagescu, *The ciliopathy protein TALPID3/KIAA0586 acts upstream of Rab8 activation in zebrafish photoreceptor outer segment formation and maintenance.* Sci Rep, 2018. **8**(1): p. 2211.

253. Nachury, M.V., E.S. Seeley, and H. Jin, *Trafficking to the ciliary membrane: how to get across the periciliary diffusion barrier?* Annu Rev Cell Dev Biol, 2010. **26**: p. 59-87.

254. Milenkovic, L., M.P. Scott, and R. Rohatgi, *Lateral transport of Smoothened from the plasma membrane to the membrane of the cilium.* J Cell Biol, 2009. **187**(3): p. 365-74.

255. Moritz, O.L., B.M. Tam, L.L. Hurd, J. Peranen, D. Deretic, and D.S. Papermaster, *Mutant rab8 Impairs docking and fusion of rhodopsin-bearing post-Golgi membranes and causes cell death of transgenic Xenopus rods.* Mol Biol Cell, 2001. **12**(8): p. 2341-51.

256. Wang, L., M. Failler, W. Fu, and B.D. Dynlacht, *A distal centriolar protein network controls organelle maturation and asymmetry.* Nat Commun, 2018. **9**(1): p. 3938.

257. Kakinuma, T., H. Ichikawa, Y. Tsukada, T. Nakamura, and B.H. Toh, *Interaction between p230 and MACF1 is associated with transport of a glycosyl phosphatidyl inositol-anchored protein from the Golgi to the cell periphery.* Exp Cell Res, 2004. **298**(2): p. 388-98.

258. Fourriere, L., A. Kasri, N. Gareil, S. Bardin, H. Bousquet, D. Pereira, F. Perez, B. Goud, G. Boncompain, and S. Miserey-Lenkei, *RAB6 and microtubules restrict protein secretion to focal adhesions.* J Cell Biol, 2019.

259. Held, R.G. and P.S. Kaeser, *ELKS active zone proteins as multitasking scaffolds for secretion.* Open Biol, 2018. **8**(2).

260. Airik, R., M. Schueler, M. Airik, J. Cho, K.A. Ulanowicz, J.D. Porath, T.W. Hurd, S. Bekker-Jensen, J.M. Schroder, J.S. Andersen, and F. Hildebrandt, *SDCCAG8 Interacts with RAB Effector Proteins RABEP2 and ERC1 and Is Required for Hedgehog Signaling.* PLoS One, 2016. **11**(5): p. e0156081.

261. Patwardhan, A., S. Bardin, S. Miserey-Lenkei, L. Larue, B. Goud, G. Raposo, and C. Delevoye, *Routing of the RAB6 secretory pathway towards the lysosome related organelle of melanocytes.* Nat Commun, 2017. **8**: p. 15835.

262. Lansbergen, G., I. Grigoriev, Y. Mimori-Kiyosue, T. Ohtsuka, S. Higa, I. Kitajima, J. Demmers, N. Galjart, A.B. Houtsmuller, F. Grosveld, and A. Akhmanova, *CLASPs attach microtubule plus ends to the cell cortex through a complex with LL5beta.* Dev Cell, 2006. **11**(1): p. 21-32.

263. Grigoriev, I., K.L. Yu, E. Martinez-Sanchez, A. Serra-Marques, I. Smal, E. Meijering, J. Demmers, J. Peranen, R.J. Pasterkamp, P. van der Sluijs, C.C. Hoogenraad, and A. Akhmanova, *Rab6, Rab8, and MICAL3 cooperate in controlling docking and fusion of exocytotic carriers.* Curr Biol, 2011. **21**(11): p. 967-74.

264. Ward, H.H., U. Brown-Glaberman, J. Wang, Y. Morita, S.L. Alper, E.J. Bedrick, V.H. Gattone, 2nd, D. Deretic, and A. Wandinger-Ness, *A conserved signal and GTPase complex are required for the ciliary transport of polycystin-1.* Mol Biol Cell, 2011. **22**(18): p. 3289-305.

265. Shibata, S., T. Kawanai, T. Hara, A. Yamamoto, T. Chaya, Y. Tokuhara, C. Tsuji, M. Sakai, T. Tachibana, and S. Inagaki, *ARHGEF10 directs the localization of Rab8 to Rab6-positive executive vesicles.* J Cell Sci, 2016. **129**(19): p. 3620-3634.

266. Shibata, S., Y. Teshima, K. Niimi, and S. Inagaki, *Involvement of ARHGEF10, GEF for RhoA, in Rab6/Rab8-mediating membrane traffic.* Small GTPases, 2019. **10**(3): p. 169-177.

267. Bachmann-Gagescu, R., M. Dona, L. Hetterschijt, E. Tonnaer, T. Peters, E. de Vrieze, D.A. Mans, S.E. van Beersum, I.G. Phelps, H.H. Arts, J.E. Keunen, M. Ueffing, R. Roepman, K. Boldt, D. Doherty, C.B. Moens, S.C. Neuhauss, H. Kremer, and E. van Wijk, *The Ciliopathy Protein CC2D2A Associates with NINL and Functions in RAB8-MICAL3-Regulated Vesicle Trafficking.* PLoS Genet, 2015. **11**(10): p. e1005575.

268. Lu, Q., C. Insinna, C. Ott, J. Stauffer, P.A. Pintado, J. Rahajeng, U. Baxa, V. Walia, A. Cuenca, Y.S. Hwang, I.O. Daar, S. Lopes, J. Lippincott-Schwartz, P.K. Jackson, S. Caplan, and C.J. Westlake, *Early steps in primary cilium assembly require EHD1/EHD3-dependent ciliary vesicle formation.* Nat Cell Biol, 2015. **17**(4): p. 531.

269. Tilley, F.C., M. Gallon, C. Luo, C.M. Danson, J. Zhou, and P.J. Cullen, *Retromer associates with the cytoplasmic amino-terminus of polycystin-2.* J Cell Sci, 2018. **131**(11).

Appendix One – Buffers, Solutions and Abbreviations

Abbreviations/Initialisms

Biological systems

DNA – Deoxyribonucleic acid

GAP – GTPase-activating protein

GDP – Guanosine diphosphate

GEF – Guanine nucleotide exchange factor

GTP – Guanosine triphosphate

PDZ – Post synaptic density protein, Drosophila disc large tumour suppressor and zonula occuludens-1 protein. (from first three proteins found the share the domain)

SNARE – SNAP Receptor

SNAP - Soluble NSF (N-ethylmaleimide-sensitive factor) Attachment Protein

SNX-PX – Sorting nexin with Phox homology domain

SNX-BAR – Sorting nexin with Bin/Amphiphysin/Rvs domain

SNX-FERM – Sorting nexin with Four-point-one, Ezrin, Radixin, moesin domain

SNX-PXA-RGS-PXC – Sorting nexin with PX-associated domain A, regulator of G protein signalling domain, PX-associated domain C

SNX-MIT – Sorting nexin with microtubule interacting and transport domain

Materials and Chemicals

AEBSF – 4-(2-Aminoethyl)benzenesulfonnyl fluoride

βME – Beta mercaptoethanol

DDO – Double dropout medium

DDO/X/A - Double dropout medium with X-alpha galactosidase and aureobasidin A

DMSO – Dimethyl Sulfoxide

EDC - *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride

ETDA – Ethylenediaminetetraacetic acid

HPLC – High performance liquid chromatography

IPTG – Isopropyl β–D-1-thiogalactopyranoside

LiAc – Lithium Acetate

MES – 2-(N-Morpholino)ethanesulfonic acid

NHS - N-hydroxysuccinimide

PCR – Polymerase chain reaction

PEG – Polyethylene glycol

PMSF – Phenyl methane sulfonyl fluoride

QDO - Quadruple dropout medium

QDO/X/A - Quadruple dropout medium with X-alpha galactosidase and aureobasidin A

QELS – Quasi-Elastic light scattering

SDO – Single dropout medium

SDO/X/A - Single dropout medium with X-alpha galactosidase and aureobasidin A

SDS PAGE – Sodium docecyl sulphate polyacrylamide gel electrophoresis

SEC-MALS – Size exclusion chromatography with multiangle light scattering

SPR – surface plasmon resonance

SOC – Super optimal broth with catabolite repression

TE – Tris/EDTA solution

TSP - Sodium-3,3,3-trimethylsilylpropionate

TY/TYE – Tryptone, Yeast extract media

YPD – Yeast extract – peptone – dextrose media

NMR Experiments

COSY – Correlated Spectroscopy

CT-HSQC – Constant Time HSQC

DQF COSY – Double Quantum Filtered COSY

HSQC – Heteronuclear Single Quantum Correlation

NOESY – Nuclear Overhauser Effect Spectroscopy

TOCSY – Total Correlation Spectroscopy

NMR notation

$d_{\alpha N}$ – distance between backbone carbon alpha proton and backbone amide proton.

d_{NN} – distance between backbone amide proton and backbone amide.

Recipes

M9 Minimal Media Recipe

Component	Concentration
Na_2HPO_4	50mM
KH_2PO_4	25mM
NaCl	10mM
Thiamine	150μM
$MgSO_4$	4mM
$CaCl_2$	0.2mM
$^{15}NH_4Cl$ / NH_4Cl	18.7mM
^{13}C6 Glucose / Glucose	11.1mM

Appendix Two – NMR experiment tables

Experiment	Instrument	Solvent	Temp.
2D Datasets (H_2O):			
^{15}N HSQC	Av-600	H_2O	25°C
^{13}C HSQC aliphatic	Av-600	H_2O	25°C
^{13}C HSQC aromatic	Av-600	H_2O	25°C
^{13}C CT HSQC aliphatic	Av-600	H_2O	25°C
^{13}C CT HSQC aromatic	Av-600	H_2O	25°C
3D Datasets (H_2O):			
CBCANH	Av-600	H_2O	25°C
CBCA(CO)NH	Av-600	H_2O	25°C
HNCA	Av-600	H_2O	25°C
HN(CO)CA	Av-600	H_2O	25°C
HBHANH	Av-600	H_2O	25°C
HBHA(CO)NH	Av-600	H_2O	25°C
[^{1}H-^{13}C-^{1}H] HCCH-COSY	Av-600	H_2O	25°C
[^{1}H-^{13}C-^{1}H] HCCH-TOCSY	Av-600	H_2O	25°C
[^{13}C-^{13}C-^{1}H] HCCH-TOCSY	Av-600	H_2O	25°C
Used for restraints [a]			
^{15}N NOESY-HSQC (τ_m = 150 ms)	Av-600	H_2O	25°C
^{13}C NOESY-HSQC aliphatic (τ_m = 150 ms)	Av-600	H_2O	25°C
2D Datasets (^2H_2O):			

DQ correlation	Av3-800	^2H_2O	25°C
3D Datasets (^2H_2O):			
Used for restraints [a]			
^{13}C NOESY-HSQC aromatic (τ_m = 150 ms)	Av3-800	^2H_2O	25°C

Table A2.1: NMR Datasets acquired for ^{15}N,^{13}C-labelled free Vps29.

[a] Although structures were not calculated for free Vps29, analysis of cross-peaks in these spectra was helpful to confirm assignments for corresponding cross-peaks in spectra of samples of the complex.

Experiment	**Instrument**	**Solvent**	**Temp.**
2D Datasets (H_2O):			
^{15}N HSQC	Av3-800	H_2O	25°C
^{13}C HSQC Full width in ^{13}C	Av3-800	H_2O	25°C
^{13}C CT HSQC aliphatic	Av3-800	H_2O	25°C
^{13}C CT HSQC aromatic	Av3-800	H_2O	25°C
Used for restraints [a]			
[^1H-^1H] NOESY ($\tau_m = 150$ ms)	Av3-800	H_2O	25°C
3D Datasets (H_2O):			
CBCANH	Av-600	H_2O	25°C
CBCA(CO)NH	Av-600	H_2O	25°C
HBHANH	Av-600	H_2O	25°C
HBHA(CO)NH	Av-600	H_2O	25°C
[^1H-^{13}C-^1H] HCCH-COSY	Av-600	H_2O	25°C
[^1H-^{13}C-^1H] HCCH-TOCSY	Av-600	H_2O	25°C
[^{13}C-^{13}C-^1H] HCCH-TOCSY	Av-600	H_2O	25°C
Used for restraints [a]			
HACAHB	Av-600	H_2O	25°C
HNHB	Av-600	H_2O	25°C
^{15}N NOESY-HSQC ($\tau_m = 50$ ms)	Av-600	H_2O	25°C
^{15}N NOESY-HSQC ($\tau_m = 150$ ms)	Av3-800	H_2O	25°C
^{13}C NOESY-HSQC aliphatic ($\tau_m = 150$ ms)	Av3-800	H_2O	25°C
^{13}C NOESY-HSQC aromatic ($\tau_m = 150$ ms)	Av3-800	^2H_2O	25°C

Table A2.2: NMR Datasets acquired for ^{15}N,^{13}C-labelled free Varp 692-746.

[a] Although structures were not calculated for free Varp 692-746, analysis of cross-peaks in these spectra was helpful to confirm assignments for corresponding cross-peaks in spectra of samples of the complex.

Experiment	**Instrument**	**Solvent**	**Temp.**
2D Datasets (H_2O):			
^{15}N HSQC	Av-600	H_2O	25°C
^{13}C HSQC aliphatic	Av-600	H_2O	25°C
^{13}C HSQC aromatic	Av-600	H_2O	25°C
^{13}C CT HSQC aliphatic	Av-600	H_2O	25°C
Used for restraints			
[^{1}H-^{1}H] filtered NOESY (τ_m = 150 ms) [a]	Av-600	H_2O	25°C
3D Datasets (H_2O):			
HNCA	Av-600	H_2O	25°C
HN(CO)CA	Av-600	H_2O	25°C
CBCANH	Av-600	H_2O	25°C
CBCA(CO)NH	Av-600	H_2O	25°C
HBHA(CO)NH	Av-600	H_2O	25°C
[^{1}H-^{13}C-^{1}H] HCCH-TOCSY	Av-600	H_2O	25°C
[^{13}C-^{13}C-^{1}H] HCCH-TOCSY	Av-600	H_2O	25°C
Used for restraints			
^{15}N NOESY-HSQC (τ_m = 150 ms)	Av-600	H_2O	25°C
^{13}C NOESY-HSQC aliphatic (τ_m = 150 ms)	Av-600	H_2O	25°C
^{13}C NOESY-HSQC aromatic (τ_m = 150 ms)	Av-600	H_2O	25°C
2D Datasets (^2H_2O):			
^{13}C HSQC aliphatic	Av3-800	^2H_2O	25°C
^{13}C HSQC aromatic	Av3-800	^2H_2O	25°C

^{13}C CT HSQC aliphatic	Av3-800	$^{2}H_{2}O$	25°C
^{13}C CT HSQC aromatic	Av3-800	$^{2}H_{2}O$	25°C
Used for restraints			
Filtered [^{1}H-^{1}H] NOESY (τ_m = 70 ms) [a]	Av3-800	$^{2}H_{2}O$	25°C
Filtered [^{1}H-^{1}H] NOESY (τ_m = 150 ms) [a]	Av3-800	$^{2}H_{2}O$	25°C
3D Datasets ($^{2}H_{2}O$):			
[^{13}C-^{13}C-^{1}H] HCCH-TOCSY	Av-600	$^{2}H_{2}O$	25°C
Used for restraints			
^{13}C NOESY-HSQC aliphatic (τ_m = 50 ms)	Av3-800	$^{2}H_{2}O$	25°C
^{13}C NOESY-HSQC aliphatic (τ_m = 120 ms)	Av3-800	$^{2}H_{2}O$	25°C
^{13}C NOESY-HSQC aromatic (τ_m = 50 ms)	Av3-800	$^{2}H_{2}O$	25°C
^{13}C NOESY-HSQC aromatic (τ_m = 120 ms)	Av3-800	$^{2}H_{2}O$	25°C
Filtered ^{13}C NOESY-HSQC aliphatic (τ_m = 50 ms) [b]	Av3-800	$^{2}H_{2}O$	25°C
Filtered ^{13}C NOESY-HSQC aliphatic (τ_m = 120 ms) [b]	Av3-800	$^{2}H_{2}O$	25°C
Filtered ^{13}C NOESY-HSQC aromatic (τ_m = 50 ms) [b]	Av3-800	$^{2}H_{2}O$	25°C
Filtered ^{13}C NOESY-HSQC aromatic (τ_m = 120 ms) [b]	Av3-800	$^{2}H_{2}O$	25°C

Table A2.3: NMR Datasets acquired for 1:1 complex of ^{15}N,^{13}C-labelled Vps29 and natural abundance Varp 692-746.

[a] These experiments had filter elements set to reject protons coupled to ^{13}C or ^{15}N in F_1 and to accept protons coupled to ^{13}C or ^{15}N in F_2.

[b] These experiments had filter elements set to reject protons coupled to ^{13}C or ^{15}N in F_1.

Experiment	**Instrument**	**Solvent**	**Temp.**
2D Datasets (H_2O):			
^{15}N HSQC	Av-600	H_2O	2°C
^{15}N HSQC	Av-600	H_2O	10°C
^{15}N HSQC	Av-600	H_2O	25°C
^{13}C HSQC Full-width in ^{13}C	Av-600	H_2O	25°C
3D Datasets (H_2O):			
CBCA(CO)NH	Av-600	H_2O	25°C
HNCA	Av-600	H_2O	25°C
HN(CO)CA	Av-600	H_2O	25°C
HBHA(CO)NH	Av-600	H_2O	25°C
HNHA	Av-600	H_2O	25°C
[^1H-^{13}C-^1H] HCCH-TOCSY	Av-600	H_2O	25°C
[^{13}C-^{13}C-^1H] HCCH-TOCSY	Av-600	H_2O	25°C
Used for restraints			
^{15}N NOESY-HSQC ($\tau_m = 50$ ms)	Av-600	H_2O	2°C
^{15}N NOESY-HSQC ($\tau_m = 50$ ms)	Av-600	H_2O	10°C
^{15}N NOESY-HSQC ($\tau_m = 70$ ms)	Av-600	H_2O	25°C
^{15}N NOESY-HSQC ($\tau_m = 150$ ms)	Av3-800	H_2O	25°C
^{13}C NOESY-HSQC aliphatic ($\tau_m = 70$ ms)	Av-600	H_2O	25°C
2D Datasets (^2H_2O):			
^{13}C HSQC aliphatic	Av3-800	^2H_2O	25°C
^{13}C HSQC aromatic	Av3-800	^2H_2O	25°C
3D Datasets (^2H_2O):			

[^{1}H-^{13}C-^{1}H] HCCH-COSY	Av-600	$^{2}H_{2}O$	25°C
[^{13}C-^{13}C-^{1}H] HCCH-TOCSY	Av-600	$^{2}H_{2}O$	25°C
Used for restraints			
^{13}C NOESY-HSQC aliphatic (τ_m = 50 ms)	Av-600	$^{2}H_{2}O$	25°C
^{13}C NOESY-HSQC aliphatic (τ_m = 70 ms)	Av3-800	$^{2}H_{2}O$	25°C
^{13}C NOESY-HSQC aliphatic (τ_m = 120 ms)	Av3-800	$^{2}H_{2}O$	25°C
^{13}C NOESY-HSQC aromatic (τ_m = 70 ms)	Av-600	$^{2}H_{2}O$	25°C

Table A2.4: NMR Datasets acquired for 1:1 complex of natural abundance Vps29 and ^{15}N,^{13}C-labelled Varp 692-746.

Appendix Three – Protein Resonance Assignments

Assignments for Bound Znk2

	H	N	N δ	N ε	H α	H β	H γ	H δ	H ε	C α	C β	C γ	C δ	C ε
688 Pro					4.41	2.26	1.96	3.51		63.10	32.28	27.03	49.61	
						1.89								
689 Leu	8.50	122.14			4.29	1.55	1.59	0.88		55.41	42.00	26.96	24.75	
						1.62		0.83					23.39	
690 Gly										45.18				
690 Gly	8.38	110.16			3.89									
					3.98									
691 Ser	8.14	115.66												
691 Ser					4.51	3.87				58.16	64.02			
						3.81								
692 Thr	8.30	115.47			4.36	4.25	1.15			61.80	69.72	21.48		
693 Glu	8.38	122.38			4.22	2.01	2.20			56.85	30.08	36.24		
						1.87	2.02							
694 Glu	8.26	120.99			4.19	1.83				56.64	30.49	36.30		
						1.97								
695 Asp	8.27	121.25			4.54	2.53				54.22	41.11			
						2.67								
696 Leu	8.09	122.54												
696 Leu					4.29	1.58	1.56	0.86		55.07	42.45	26.86	23.38	
						1.52		0.80					24.86	
697 Glu	8.31	121.76			4.18	1.98				56.56	30.37			
						1.88								
698 Asp	8.30	121.65			4.52	2.65				54.13	41.12			
						2.51								
699 Ala	8.15	124.65			4.25	1.33				52.38	19.52			
700 Glu	8.36	120.09			4.19	1.88					30.33			
						1.99								
700 Glu										56.49				
701 Asp	8.33	121.35			4.59	2.57				54.22	41.02			
						2.65								
702 Thr	8.05	114.73			4.26	4.14	1.14			62.02	69.74	21.51		
703 Val	8.10	122.65			4.07	2.03	0.87			62.48	32.60	20.20		
							0.86					20.96		
704 Ser	8.35	119.79			4.38	3.77				58.05	63.74			
705 Ala	8.26	126.43			4.21	1.30				52.28	19.23			
706 Ala	8.09	123.08												
706 Ala					4.21	1.28				52.04	19.31			
707 Asp	8.18	121.35			4.75	2.39				52.58	40.62			
						2.63								
708 Pro					4.29	2.15	1.94	3.71		62.86	32.06	27.20	50.50	
						1.77		3.59						
709 Glu	8.30	120.71			4.08	1.75				56.25	30.12	36.30		
710 Phe					5.04	3.23				55.86	40.54			
						2.96								
711 Cys	8.29	123.85			4.18	2.91				61.98	31.54			
						2.65								
712 His		112.43												
712 His	7.11							7.19	7.36				120.47	138.16
713 Pro					3.92	2.28	1.43			65.98				
						1.75	1.55							
714 Leu					4.53	1.96	1.55	0.71		52.02		27.13	20.32	
						1.60		0.54					25.50	
715 Cys	6.75	123.73			3.35	2.96				62.05	30.79			
						1.75								
716 Gln	8.40	126.68		111.37	4.44	1.65	2.37		7.74	54.57				
						1.26			6.51					
717 Cys					4.74									
717 Cys	9.70	127.14			4.73	2.87				58.95	27.33			
						3.45								
718 Pro						2.37	1.82	3.88		65.39	32.10	27.91	50.41	
						2.48	2.05	3.60						
718 Pro					4.21									
719 Lys	8.68	118.10												
719 Lys					4.12	1.76				58.35	33.16	24.87	29.40	41.92
720 Cys	7.82	123.06			4.00	2.82				61.89	31.75			
						2.36								
721 Ala	8.37	122.78												
721 Ala					4.39	1.15				54.39	17.83			
722 Pro					4.29			3.45		64.47			50.60	
723 Ala	7.90	121.97												
723 Ala					4.19	1.41				53.26	19.03			
724 Gln	8.18	118.55		111.37	4.07	1.79	2.16		7.47	56.84	29.93	36.24		
						1.83	2.23		6.52					

	H	N	N δ	N ε	H α	H β	H γ	H δ	H ε	C α	C β	C γ	C δ	C ε
725 Lys	7.93	119.87			4.49									
726 Arg					4.12	1.68	1.54	3.13		57.13	30.70	27.21	43.34	
						1.77	1.59							
727 Leu	7.99	122.16												
727 Leu											42.31			
728 Ala	7.95	123.95												
728 Ala					4.22					52.36				
729 Lys	8.09	120.56			4.24	1.73	1.31	1.64	2.92	56.04	32.99	24.66	29.07	42.14
						1.67	1.43	1.60						
730 Val	8.09	122.98			4.34	2.02	0.92			59.82	32.54	20.20		
							0.88					20.95		
731 Pro					4.33	2.24	1.90	3.79		63.04	32.10	27.33	50.94	
						1.86	1.98	3.61						
732 Ala	8.38	124.56			4.23	1.37				52.63	19.06			
733 Ser	8.20	114.31								58.32	63.80			
734 Gly										45.22				
734 Gly	8.31	110.45												
735 Leu	8.05	121.20								55.15	42.35			
736 Gly										45.21				
736 Gly	8.38	109.56			3.88									
737 Val	7.82	118.53			4.06	2.01	0.83			62.15	32.57	20.17		
							0.86					21.03		
738 Asn	8.49	122.27	112.92		4.69	2.77		6.82		53.14	38.84			
						2.66		7.50						
739 Val	8.07	120.47			4.13	2.06	0.85			62.36	32.61	21.15		
							0.87					20.14		
740 Thr	8.20	117.20			4.34	4.16	1.14			61.64	69.87	21.50		
741 Ser	8.23	117.81			4.39	3.83				58.21	63.79			
						3.78								
742 Gln	8.43	122.26		112.35	4.29	2.05			6.75	56.02	29.24	33.67		
						1.88			7.42					
743 Asp	8.23	120.82			4.53	2.62				54.36	41.17			
						2.66								
744 Gly										45.36				
744 Gly	8.24	109.53			3.81									
					3.91									
745 Ser	8.10	115.60			4.33	3.71				58.48	63.86			
746 Ser	8.08	117.49			4.38	3.67				58.22	63.91			
747 Trp														
747 Trp	7.63	127.70		128.53	4.46	3.27			9.94	58.49	29.97			
						3.10								

Assignment for Free Znk2:

	H	N	N δ	N ε	H α	H β	H γ	H δ	H ε	H ζ	H η	C α	C β	C γ	C δ	C ε	C ζ
688 Pro					4.40	2.25	1.95	3.51				63.08	32.25	27.02	49.65		
						1.88		4.05									
689 Leu																	
689 Leu	8.50				4.27	1.62		0.87				55.40	42.06	26.98	23.42		
						1.54		0.82							24.84		
690 Gly	8.37	110.11			3.88							45.26					
					3.98												
691 Ser	8.14	115.59			4.49	3.86						58.21	64.06				
						3.79											
692 Thr	8.29	115.48			4.34	4.25	1.14					61.96	69.75	21.55			
693 Glu	8.38	122.39			4.21	2.00	1.99					56.91	30.11	36.41			
						1.85	2.20										
694 Glu	8.25	120.98			4.17	1.84	2.16					56.65	30.36	36.30			
						1.98	2.20										
695 Asp	8.26	121.22				2.65						54.42	41.04				
						2.52											
695 Asp					4.52												
696 Leu	8.08	122.58			4.27	1.57		0.79				55.05	42.35	26.90	24.91		
						1.51		0.85							23.39		
697 Glu	8.31	121.81			4.17	1.96	2.20					56.69	30.33	36.21			
						1.87	2.17										
698 Asp	8.30	121.68			4.51	2.62						54.23	41.09				
						2.51											
699 Ala	8.14	124.64			4.25	1.32						52.48	19.55				
700 Glu	8.35	120.13			4.17	1.97	2.15					56.66	30.39	36.25			
						1.88	2.20										
701 Asp	8.32	121.41			4.58	2.56						54.29	41.05				
						2.64											
702 Thr	8.04	114.73			4.25	4.14	1.11					61.99	69.79	21.52			
703 Val					4.06	2.02								21.17			
														20.68			
703 Val	8.09	122.73					0.87					62.46	32.65				
							0.85										
704 Ser	8.33	119.82			4.35	3.76						58.23	63.85				
705 Ala	8.23	126.43			4.23	1.29						52.26	19.23				
706 Ala	8.08	123.25			4.21	1.26						52.05	19.38				
707 Asp	8.17	121.43			4.72	2.59						52.58	40.69				
						2.37											
708 Pro					4.24	1.57	1.88	3.58				62.93	31.95	27.18	50.54		
						2.04	1.81	3.71									
709 Glu	8.23	120.35			4.03	1.79	2.07					56.15	30.12	36.20			
						1.70	1.96										
710 Phe	7.90	120.69			5.11	3.24		7.11	7.24	7.03		56.32	40.61		131.78	131.22	129.28
						2.66											
711 Cys	8.12	124.05			3.99	3.03						62.30	32.00				
						2.76											
712 His	7.68	114.76			4.67	3.17		7.03	7.50			54.92	29.75		120.02	138.85	
						3.24											
713 Pro					4.16	1.79	1.94	3.86				66.07	32.30	27.88	50.63		
						2.37	2.07	3.46									
714 Leu	8.33	115.90			4.37	1.70	1.54	0.88				53.60	41.10	27.46	25.16		
						1.77		0.78							22.67		
715 Cys	7.14	124.26			3.51	2.35						62.60	31.62				
						2.99											
716 Gln									6.31								
716 Gln	8.32	125.78		110.56	4.43	2.37	2.29		7.30			54.31	28.83	32.98			
						1.61	2.25										
718 Pro					4.22	2.37	2.08	3.87				65.53	32.07	28.07	50.15		
						1.82	1.94	3.59									
719 Lys	8.62	118.05			4.13	1.80	1.34	1.58	2.89			58.44	33.42	25.00	29.21	41.96	
						1.75	1.45										
720 Cys	7.84	123.31			4.06	2.45						62.34	32.27				
						2.88											
721 Ala	8.65	122.12			4.22	1.22						54.10	16.75				
722 Pro					4.28	2.26	1.98	3.50				64.98	31.29	28.08	50.62		
						1.80	1.91	3.40									
723 Ala	7.75	121.05			4.20	1.43						53.40	19.04				
724 Gln	8.14	118.44			4.23	1.85	2.21					57.18	28.55	32.55			
						1.96	2.28										
724 Gln				111.37					6.50								
									7.18								
725 Lys	7.97	120.21			4.12	1.80	1.44	1.60	2.91			57.29	32.54	24.88	29.09	42.13	
						1.73	1.35										
726 Arg	7.87	120.17			4.18	1.79	1.62	3.15				56.93	30.49	27.38	43.37		
						1.76	1.53										
727 Leu	7.95	121.76			4.23	1.64	1.52	0.83				55.25	42.34	26.96	25.00		
						1.55		0.88							23.37		
728 Ala	7.92	123.72										52.52	19.08				
728 Ala					4.21	1.32											
729 Lys					4.24	1.67								24.77			
						1.74											

	H	N	N δ	N ε	H α	H β	H γ	H δ	H ε	H ζ	H η	C α	C β	C γ	C δ	C ε	C ζ	C η
729 Lys	8.05	120.17					1.37 1.30	1.60	3.11 2.91			56.09	32.97		28.98	42.10		
730 Val	8.05	122.88			4.33	2.00	0.91 0.87					59.86	32.54	20.31 20.96				
731 Pro					4.32	2.22 1.85	1.95 1.90	3.79 3.59				63.01	32.07	27.33	51.00			
732 Ala	8.37	124.56			4.22	1.34						52.74	19.07					
733 Ser	8.19	114.32			4.34	3.83 3.78						58.55	63.89					
734 Gly	8.31	110.35			3.78 3.90							45.31						
735 Leu	8.04	121.21			4.27	1.57 1.52		0.79 0.85				55.20	42.42	26.93	24.84 23.25			
736 Gly	8.38	109.59			3.97 3.86							45.34						
737 Val	7.82	118.54			4.04	1.99	0.82					62.22	32.72	20.22 20.99				
738 Asn	8.48	122.27			4.69	2.65 2.75						53.17	38.80					
738 Asn			112.91					6.82 7.50										
739 Val	8.06	120.42			4.12	2.04	0.84					62.37	32.70	20.27 21.18				
740 Thr	8.19	117.26			4.34	4.15	1.11					61.71	69.85	21.58				
741 Ser	8.23	117.82			4.38	3.82 3.76						58.30	63.85					
742 Gln	8.42	122.21		112.34	4.28	1.89 2.06	2.27		7.42 6.74			56.02	29.28	33.73				
743 Asp	8.22	120.84			4.52	2.63 2.51						54.36	41.14					
744 Gly	8.24	109.51			3.79 3.90							45.44						
745 Ser	8.09	115.58			4.33	3.69						58.48	63.91					
746 Ser	8.08	117.44			4.36	3.67 3.82						58.29	64.00					
747 Trp		127.67			4.46	3.26 3.09			7.58 9.94	7.06 7.40	7.13	58.46	30.00			121.22	121.80 114.44	124.37
747 Trp								7.11							126.92			

Assignments for bound Vps29:

	H	N	N δ	N ε	H α	H β	H γ	H δ	H ε	H ζ	H η	C α	C β	C γ	C δ	C ε	C ζ
-7 Pro					4.33	2.15 1.72	1.92	3.76 3.65				63.39	31.95	27.22	50.65		
-6 Glu	8.39	120.88			4.10	1.84 1.75	2.14 2.04					56.73	30.13	36.19			
-5 Phe	8.12	120.48			4.56	3.16 2.95		7.18				57.84	39.63		131.80		
-5 Phe									7.24							131.40	
-4 Gly	8.35	110.21			3.97 3.91							45.42					
-3 Thr	8.01	113.32			4.31	4.17						61.91	69.89	21.61			
-2 Arg	8.38	122.43						3.03 3.15				56.24		26.81	43.31		
-2 Arg					4.32	1.85 1.74							30.73				
-1 Asp	8.25	120.72			4.54	2.57 2.69						54.53	41.21				
0 Arg	7.85	121.11			4.94	1.65		3.03 3.06				55.48	32.31	26.24	44.01		
1 Met	9.04	124.67			4.86	2.08 1.95	2.50 2.57					54.80	34.83	32.31			
2 Leu					5.52	2.04 1.36	1.73	0.85 0.87				53.48	43.76	27.03	23.77 24.95 23.61		
3 Val	9.37	124.23			5.02	2.06	0.69 0.83					59.87	34.55	21.68 21.65			
4 Leu	8.75	130.12			4.93	1.36 1.24		0.67 0.51				54.53	44.71	27.30	26.04 24.85		
5 Val	9.63	128.60			5.38	2.04	0.87 0.90					60.58	33.21	21.91 21.66			
6 Leu	8.26	123.37			5.53	1.68 1.97		0.57 0.70				53.54	44.79		24.81 24.39		
7 Gly	8.09	103.48			5.43 3.46							45.90					
8 Asp	9.39	116.64			4.25							56.72	36.19				
9 Leu	7.19	118.77			4.17	1.39	1.49	0.84				57.27	39.84	29.49	25.18		
10 His	8.28	109.88			3.94	2.68 2.92						57.89	28.23				
11 Ile	7.99	119.80			5.17	1.56	0.61 1.27	0.55				57.80	40.22	27.62 15.83	13.69		
12 Pro					5.48	2.18	1.89					64.09	33.68	30.51			
13 His	7.87	123.22			4.26	3.10 2.97						59.82	31.35 33.66				
14 Arg	8.50	112.47			4.63							56.76	31.23		43.26		
15 Cys	7.48	116.06											30.55				
16 Asn					4.69	2.71						54.56	39.97				
17 Ser	7.34	109.82			4.34	3.47 3.67						57.10	64.74				
18 Leu	8.45	122.80				1.36 1.27		0.66 0.73				55.26	42.34		25.58 25.62		
18 Leu					4.28												
20 Ala					3.85	1.41						56.21	18.12				
21 Lys	9.05	114.58			3.91	1.65	1.56 1.35		2.98			59.84	32.59	27.10	29.33	43.37	
22 Phe	7.10	113.86			4.76	2.90		6.88	7.03			56.15	36.57		129.67	130.63	
22 Phe										6.51							130.65
23 Lys	8.01	119.20			3.77	1.82 1.66	1.20 1.34	1.51				60.33	31.85	26.03	29.20		
24 Lys	7.56	114.55											32.82	24.70	28.98	41.44	
24 Lys					3.94	1.37 1.77	1.45	1.64				57.95					
25 Leu	6.84	114.77			4.19	1.89 1.40		0.95 0.94				55.75	43.79		26.50 22.73		
26 Leu	7.10	120.20			3.93	1.67		0.07 0.07				53.62	40.05	25.05	24.06 22.86		
27 Val	6.75	118.51			4.48	1.98	0.75 0.84					57.86	32.76	21.16 18.55			
28 Pro					4.29	1.76 2.41	1.92	3.76 3.65				63.42	32.60	27.34	50.88		
29 Gly	8.24	108.50			4.07 3.68							46.08					
30 Lys	8.30	119.50			4.40	1.56	1.26		2.86 2.87			56.60	35.32	24.86		41.76	
31 Ile	8.52	119.79			3.78	1.88	1.61 1.16					61.02	39.36	26.80			
31 Ile							0.77	0.70						20.26	13.97		
32 Gln				113.27					6.36 7.48								
32 Gln	8.09	123.46					2.45 2.19					59.02	31.18	36.35			
32 Gln					4.29	2.19 1.64											
33 His	7.47	114.44			5.39	2.84						53.86					

	H	N	N δ	N ε	H α	H β	H γ	H δ	H ε	H ζ	H η	C α	C β	C γ	C δ	C ε	C ζ
34 Ile					4.99	1.54		0.67									
34 Ile	8.79	121.19					1.00					60.34	40.84	16.86	15.69		
							0.64							27.49			
							0.40										
35 Leu	9.37	128.62			4.83	1.74		0.61				53.68	42.36	27.62	24.76		
						1.25		0.62							26.41		
36 Cys	9.03	122.72			6.08	2.39						54.72	28.93				
						2.68											
37 Thr	8.46	117.09				4.01	1.22					63.20	69.11	21.64			
38 Gly	7.49	104.14			3.44							43.56					
39 Asn	10.64	117.64															
39 Asn					4.24	3.03						52.91	34.68				
40 Leu					4.04		1.68	0.94					41.23	28.17	24.14		
								0.78									
40 Leu	10.13	124.77										59.32			26.67		
41 Cys	8.02	116.17			4.97	2.48						60.09	31.50				
						3.67											
42 Thr	8.45	111.95			5.07	3.70	1.16					59.55	72.74	21.59			
43 Lys	8.67	120.49			4.05	1.93	1.58	1.70	2.99			58.88	32.68	24.59	29.21	42.08	
						1.79			2.93								
44 Glu	8.31	118.25															
44 Glu					4.11	1.94	2.24					60.36	29.50	36.84			
						1.84											
45 Ser	7.38	113.71			4.47	3.96						62.21	63.42				
						3.58											
46 Tyr									6.64							117.60	
46 Tyr	7.24	124.72			3.87	2.96		7.01				62.03	37.60		133.31		
						3.31											
47 Asp	8.62	118.40			4.02	2.52						57.29	39.43				
						2.67											
48 Tyr									6.39							117.56	
48 Tyr	7.28	118.84			4.32	2.96		6.83				60.50	37.55		132.60		
						3.23											
49 Leu	7.81	121.55			3.31	1.00	2.16	0.66				58.31	40.42	26.17	22.12		
						1.86		0.71							25.56		
50 Lys	7.61	114.77			4.05	1.61	0.98	1.48	2.82			57.00	30.95	24.39	28.34	41.70	
						1.54	1.11										
51 Thr	7.39	109.08			4.07	4.05	1.27					63.59	69.75	21.45			
52 Leu	7.22	120.05			3.85	1.80		0.26				55.88	43.36		25.39		
								0.17							21.42		
53 Ala					4.34	0.85						51.27	22.48				
53 Ala	7.19	117.52															
54 Gly	8.02	102.97			3.63							45.93					
					3.96												
55 Asp	6.87	119.24			4.81	2.50						54.00	42.07				
56 Val	7.97	124.56			4.31	1.69	0.53					61.69	34.01	21.37			
							0.50							21.60			
57 His	9.03	128.69			4.45	3.06						55.64	33.29				
58 Ile	7.99	122.85										59.44	41.28	28.04	15.24		
														17.09			
58 Ile					4.83	0.82	1.07	0.58									
							0.14										
							0.67										
59 Val	6.80	116.00			4.95	1.35	0.55					58.26	34.13	22.23			
							0.59							17.63			
60 Arg	7.73	115.84			4.09	1.66		3.12				57.75	32.02	27.14	43.33		
						1.63		3.04									
61 Gly	10.59	118.17										63.99					
62 Asp	9.01	115.63			3.74	1.98						54.94	36.40				
						2.04											
63 Phe	9.42	117.37			4.46	3.41						58.35	41.38				
						2.24											
64 Asp	6.72	120.42										57.57	42.40				
65 Glu					4.13	1.99	2.34					57.01	30.02	36.46			
						2.05											
65 Glu	8.27	120.38															
66 Asn			111.37					6.62									
66 Asn	7.92	117.39						6.83				57.09	37.45				
												52.98					
66 Asn					4.38	2.40							37.56				
						2.50											
67 Leu	7.85	121.88			4.06	1.29	1.48	0.77				54.62	41.19	27.06	25.13		
						1.38		0.67				52.99			22.51		
68 Asn			113.24					6.77									
								7.36									
68 Asn	8.18	115.81			4.50	2.66						53.38	38.22				
						2.51											
69 Tyr					4.86	2.99			6.51							118.29	
						2.28											
69 Tyr	6.91	120.34						6.82				54.61	36.44		131.94		
70 Pro					4.60	2.41	2.14					62.29	32.02	27.31	49.78		
						2.19	2.06										

	H	N	N δ	N ε	H α	H β	H γ	H δ	H ε	H ζ	H η	C α	C β	C γ	C δ	C ε	C ζ	C η
71 Glu	8.96	119.60			3.98	2.05	2.40 2.30					59.48	30.55	37.23				
72 Gln	7.47	109.67			5.24	2.04 1.98	2.20 2.17					53.59	31.07	31.88				
73 Lys	8.38	119.08			4.38	1.43	1.26		3.15			55.96	37.64	25.33	29.47	43.32		
74 Val	8.02	122.02			4.78	1.69	0.58 0.61					61.19	33.43	21.07 21.20 21.07				
75 Val	9.29	126.18			4.21	1.47	0.16 0.20					59.80	34.86	19.74 19.22				
76 Thr	8.37	121.62			5.03	3.85	0.95					61.66	69.48	21.09				
77 Val	8.67	127.63			3.87	1.66	0.58 0.22					61.23	33.34 69.64	19.52 21.91				
78 Gly	8.77	115.10										46.55						
79 Gln	9.18	124.80			4.02	1.35 1.86	2.00 2.08					55.97	28.54	34.02				
80 Phe									7.26							131.70		
80 Phe	8.30	118.24			4.47	3.10 2.54		7.11				58.75	41.48		131.17			
81 Lys	9.36	123.37			4.67	1.77 1.51	1.28 1.20		2.82 2.88			55.95	34.23	24.86	29.34	41.89		
82 Ile	9.33	126.64			4.98	1.56	0.57	0.50				60.28	40.18	17.81 27.33	13.48			
83 Gly	9.50	113.48			3.05 4.96							44.32						
84 Leu	8.57	124.84			5.25	0.94 1.51	1.34	0.52 0.55				54.08	47.23	26.75	25.62 24.08			
85 Ile	8.30	121.57			4.59	1.88	0.79	0.73				61.35	41.90	28.45 14.58	14.56			
86 His	9.52	123.95																
86 His					4.73							62.14						
87 Gly	7.90	100.53										45.15						
88 His					3.47	2.92 2.87						58.40	28.26					
89 Gln	9.02	120.01										56.58						
89 Gln					4.13													
90 Val	6.99	123.65			3.79	2.27	0.86 0.73					62.43	31.79	21.05				
91 Ile	7.82	123.97			4.27	1.60	0.65 1.29 1.02	0.70				56.50	41.19	26.67 16.53	12.12			
92 Pro					4.48	2.23 1.89	1.76	3.31 3.43				62.80	34.00	25.49	49.79			
93 Trp									7.57							120.98		
93 Trp	7.82	119.86		129.68				7.81	10.20	7.59 7.05	7.20	57.21			128.09		114.49 122.17	124.67
94 Gly					3.45 3.88							45.54						
95 Asp	7.41	119.62			4.19	2.44 2.64						55.01	43.71					
96 Met	8.54	127.79			3.78	1.68	1.96					58.85						
97 Ala	8.20	120.75			3.95	1.35						55.15	17.52					
98 Ser	7.91	115.51										62.06						
99 Leu	8.16	124.06			4.19	0.73 1.03		0.76 0.65				58.25	41.20		22.41 25.17			
100 Ala	8.38	122.26			4.79	1.38						54.44	17.51					
101 Leu	7.47	119.10			4.03	1.76 1.59	1.67	0.86 0.80				58.08	41.44	26.81	24.70 23.75			
102 Leu	7.35	120.60			3.68	1.82 0.94	1.50	0.59 0.71				57.40	41.22	26.75	23.10 25.86			
103 Gln	8.36	119.98			3.58	2.14 2.28						60.58	40.85	33.32				
104 Arg	7.56	118.04			3.92	1.86 1.79		3.15 3.22				58.77	30.29	27.14	43.75			
105 Gln	8.05	120.19			3.91	2.02 1.81	2.09 2.33					58.64	28.69	33.69				
106 Phe	8.43	116.37			4.49	3.11 2.86						56.37	39.72					
107 Asp	8.27	119.97			4.22	2.24 3.14						54.92	39.29					
108 Val	6.86	105.51			4.46	2.42	0.77 0.70					58.51	33.38	18.72 23.84				
109 Asp	8.71	119.50										57.37						
110 Ile	7.28	113.77			5.10	1.49	0.56	0.49				58.10	42.44	27.25 16.19	13.75			
111 Leu	8.56	129.00			4.71			0.87 0.68				53.12	44.41		23.53 27.79			
112 Ile	8.93	129.84			5.30	1.77	0.54 1.41	0.60				59.15	38.51	27.89 18.36	15.70			
113 Ser	8.51	119.81											66.38					
113 Ser					5.65	3.49 3.97						56.81						

	H	N	N δ	N ε	H α	H β	H γ	H δ	H ε	H ζ	H η	C α	C β	C γ	C δ	C ε	C ζ
114 Gly	7.17	104.67										46.54					
115 His	9.62	125.56															
116 Thr	8.16	109.77			3.47	4.09	1.08					62.78	69.01	21.90			
117 His	9.10	113.47			3.84	3.39						56.21	26.31				
117 His						3.54			8.27							139.25	
118 Lys	7.49	119.46			4.59	1.56	1.29	1.56	2.85			55.03	34.20	24.25	28.89	42.04	
						1.41	1.18										
119 Phe	8.31	125.16										57.45	38.67			130.56	
119 Phe					4.60	2.95		7.05	6.96						132.49		
						2.47											
120 Glu	8.28	125.88										55.67					
120 Glu					3.61	1.87	2.25						34.17	35.97			
						1.60	1.98										
121 Ala	8.07	122.56			5.04	1.17						51.43	20.72				
122 Phe									6.68							130.95	
122 Phe	8.11	117.77			4.83	2.97		6.67		6.58		56.24	40.56		132.05		129.39
						3.16											
123 Glu	8.84	121.04			5.59	2.30	2.32					54.10	32.58	36.86			
						2.20	2.46										
124 His	9.30	124.95										56.90	33.00				
125 Glu					3.65	1.99	1.42					57.48	27.37	35.75			
						1.66	1.65										
126 Asn	9.08	108.35			4.71	3.19		7.59				55.50	38.04				
						3.04											
126 Asn			111.82					6.94									
127 Lys	8.06	121.42										53.57	33.28				
128 Phe									6.97							129.17	
128 Phe	7.90	123.09			4.97			7.08				55.46			128.88		
129 Tyr									6.52							116.66	
129 Tyr	7.99	125.90			5.08	3.10		6.65				56.55	40.43		133.53		
						2.39											
130 Ile					4.36	1.23	0.62	0.71				58.81	42.94	27.53	15.37		
							0.37							17.39			
131 Asn	8.77	125.89			6.18	2.56						48.58	41.14				
						2.91											
132 Pro					4.46												
133 Gly	8.58	107.12										44.25					
134 Ser	10.08	117.94			5.31	3.58						55.49	63.59				
						3.07											
135 Ala	7.92	133.29			5.09	1.50						54.81	19.16				
136 Thr	7.45	97.46			4.66	4.28	1.07					59.49	69.65	23.60			
137 Gly	7.79	110.10			3.30							46.56					
					3.87												
138 Ala	7.69	122.73			3.95	0.98						46.63	19.72				
												52.25					
139 Tyr								7.01	6.66						134.00	117.45	
139 Tyr	7.18	117.80			4.44	2.97						56.98	37.96				
						3.13											
140 Asn	8.08	116.94			4.16	1.61						51.25	38.91				
						2.59											
140 Asn			109.83					6.55									
								5.91									
141 Ala	8.01	118.60			4.07	1.39						54.25	18.57				
142 Leu	7.74	116.19			4.37	1.54		0.77				55.36	43.93	27.00	24.94		
						1.63		0.86							22.99		
143 Glu	8.18	119.66			4.54	1.90	2.07					55.79	32.64	36.08			
							2.18										
144 Thr	8.05	111.16			4.47	4.21	1.05					61.20	69.85	21.40			
145 Asn	8.78	123.21			4.82	2.81						53.02	37.47				
						2.71											
145 Asn			112.82					6.80									
								7.60									
146 Ile	7.88	123.77			4.17	1.66	1.05	0.75				58.29	38.26	26.36	10.07		
														18.66			
147 Ile					4.48												
147 Ile												58.22					
147 Ile	8.68	130.23				1.89	0.95	0.87					39.99	27.59	12.74		
														17.23			
148 Pro					4.37	1.77	1.93	3.58				62.98	31.94	27.30	50.51		
149 Ser	8.45	117.10			5.96	3.84						57.56	67.53				
						3.71											
150 Phe	8.36	113.36			4.96	2.97		6.34	5.61			57.57	39.94		131.85	129.77	
						3.35											
150 Phe										5.67							128.11
151 Val	8.87	118.24			5.35							59.72					
151 Val						1.93	0.95						37.20	22.57			
														23.29			
152 Leu	9.70	127.44			5.56	2.15	1.67	0.93				53.03	47.96	27.50	25.36		
						1.60		0.92							26.58		
153 Met	9.83	123.76															
153 Met					5.04	1.64						55.34	31.40				
						2.33											

	H	N	N δ	N ε	H α	H β	H γ	H δ	H ε	H ζ	H η	C α	C β	C γ	C δ	C ε
154 Asp	9.50	125.59			5.19	2.32						52.92	42.57			
						2.19										
155 Ile	8.94	126.60			4.24	1.89	0.96	0.56				62.55	38.51	17.84	12.84	
							1.49							27.84		
156 Gln	8.63	128.00			4.67	2.05	2.15					55.02	30.65	36.19		
						1.87										
157 Ala	8.92	127.57			3.90	1.52						54.68	17.04			
158 Ser	7.93	114.24			4.41	4.00						58.32	63.62			
						3.86										
159 Thr	8.20	117.61			4.88	4.01	0.98					62.03	71.48	21.69		
160 Val	9.20	124.80			4.38	1.55	0.53					60.74	33.88	21.57		
							0.13							21.48		
161 Val	8.67	129.69			4.16	2.12	0.37					62.49	31.87	21.16		
							0.81							22.05		
162 Thr													69.62			
162 Thr	9.13	126.07			4.50	4.01						63.31				
163 Tyr	9.53	125.60			5.08	3.12		6.96				56.83	40.51		133.19	
						2.39										
163 Tyr									6.76							118.18
164 Val	8.90	123.81			4.16	1.31	-0.16					60.53	32.98	20.07		
							0.25							18.10		
165 Tyr								6.17							133.10	
165 Tyr	8.32	127.70			4.37				5.99			56.71	39.17			116.68
166 Gln	8.58	117.60			5.42							53.89	34.06	35.35		
167 Leu	8.08	124.58			4.84	1.56	0.54	0.53				53.82	42.73	27.55	26.50	
						0.72		0.38							24.01	
168 Ile	8.45	127.87			4.19	1.72	0.82	0.75				59.69	38.50	27.18	12.53	
							1.04							17.13		
							1.31									
169 Gly	8.77	118.33			3.60							47.25				
					3.87											
170 Asp	8.75	123.37			4.52	2.39						54.44	42.03			
						2.73										
171 Asp	7.71	118.38			4.85	2.54						53.24	44.24			
172 Val	8.43	121.48			4.60	2.04	0.85					62.04	32.01	21.63		
							0.96							22.75		
173 Lys	9.32	131.17			4.50	1.66	1.36	1.54				54.44	34.26	24.85	28.98	42.09
							1.25									
174 Val	8.31	124.75			4.97	1.52	0.01					60.17	33.74	21.42		
							0.25							20.99		
175 Glu	8.19	127.01			4.66	1.54	2.08					54.37	33.19	36.08		
						1.86										
176 Arg	8.51	122.16			5.14			2.70				54.42	33.46	26.32	43.48	
177 Ile	9.09	126.46					0.98									
177 Ile					4.31	1.59	1.12	0.96				60.87	42.80	18.05	15.23	
														27.88		
178 Glu	8.77	126.90			4.88	1.79	2.01					55.62	32.68	37.26		
							1.89									
179 Tyr	8.53	125.14			4.43	1.61			6.41			56.58	40.35			117.24
179 Tyr								6.53							132.50	
181 Lys	7.96	122.58			4.06	1.63	1.29	1.13	2.65			56.41	33.57	24.62	29.01	41.87
						1.55		1.02								
182 Ser	8.33	126.06			4.17	3.76						60.14	64.99			

Assignments for Free Vps29

	H	N	N ε	H α	H β	H γ	H δ	H ε	H ζ	H η	C α	C β	C γ	C δ	C ε	C ζ
-6 Glu	8.39	120.95		4.10	1.76						56.65	30.16				
-5 Phe	8.13	120.63		4.57	2.94		7.18	7.25			57.76	39.53		131.81	131.34	
					3.14											
-4 Gly	8.34	110.09		3.91							45.44					
-3 Thr	8.01	113.29		4.30	4.21	1.15					61.94	69.75	21.63			
-2 Arg	8.36	122.35		4.31	1.74	1.61	3.14				56.41	30.58	27.10	43.51		
					1.86	1.58										
-1 Asp	8.26	120.46		4.53	2.66						54.55	41.14				
					2.60											
0 Arg	7.81	120.44		4.93	1.74	1.47	3.03				55.28	32.19	26.55	43.75		
					1.64	1.60	3.06									
1 Met	8.98	124.09		4.82	2.10	2.55					54.95	34.64	32.25			
					1.93	2.49										
2 Leu	8.91	123.80		5.46	1.69	1.53	0.80				53.57	43.99				
							0.87									
3 Val	9.38	125.04		4.98	2.06	0.70					59.87	35.12	21.82			
						0.85							21.78			
4 Leu	8.57	129.46		5.03	1.85	1.42	0.55				54.22	44.62	28.07	24.42		
							0.72							26.21		
5 Val	9.59	128.56		5.38	2.02	0.89					60.53	33.24	21.80			
						0.86							21.81			
6 Leu	8.26	123.39		5.53	1.74	1.67	0.71				53.55	44.67	26.98			
					1.99		0.61									
7 Gly	8.09	103.37		5.44							45.82					
				3.48												
8 Asp	9.39	116.56		4.33							56.61	41.24				
9 Leu	7.20	118.79		4.17	1.37	1.49	0.84				57.36	39.94	29.53	25.21		
							0.84							25.21		
10 His	8.28	109.93		3.94	2.78						57.80	28.61				
					2.71											
11 Ile	7.99	119.80		5.16	1.62	0.61	0.47				57.77	40.34	15.71	13.63		
12 Pro				5.48	2.19		2.98				64.08	33.59	23.71	50.19		
					1.99		3.44									
13 His	7.87	123.32		4.28	2.98						59.58	31.41				
					3.10											
14 Arg	8.50	112.47		4.63							56.63	31.20				
15 Cys	7.49	116.15									56.46					
16 Asn				4.69	2.73						54.60	40.03				
					2.73											
17 Ser	7.35	109.79		4.35	3.48						57.03	64.76				
					3.67											
18 Leu	8.46	122.86		4.69	2.71						55.07	42.47				
					2.71											
20 Ala				3.83	1.40						56.07	18.12				
21 Lys	8.99	114.41		3.90	1.44	1.34					59.76	32.89	27.31			
					1.67	1.55										
22 Phe	7.17	113.90		4.82	2.87						55.88	36.58				
					3.38											
23 Lys	7.92	119.22		3.78	1.69	1.20	1.52	2.68			60.41	31.85	26.15	29.32	41.69	
					1.83	1.36	1.45	2.68								
24 Lys	7.28	113.73		3.95	1.75	1.47	1.62	2.91			57.77	32.69	24.92	29.05	42.13	
					1.74	1.42	1.70	2.89								
25 Leu	7.00	116.38		4.10	1.28		0.89				55.50	43.34		22.76		
					1.83		0.82							23.02		
26 Leu	7.02	121.38		4.09	1.68	0.81	0.09				53.31	40.17	25.25	24.13		
					0.87									22.93		
27 Val	7.29	120.58		3.77	1.66						58.79	33.12				
28 Pro				4.31	2.40	1.69	3.64				63.68	32.02	27.32	50.72		
					1.80	1.91	3.74									
29 Gly	8.37	109.79		3.64							45.82					
				4.14												
30 Lys	8.19	119.24		4.35	1.71	1.32	1.53	2.82			56.18	34.58	24.89	28.90	42.12	
					1.70	1.28	1.53	2.82			56.35					
31 Ile	8.21	119.24		3.83	1.66	0.83	0.69				60.83	39.53	28.13	14.02		
						1.43							20.20			
						0.74										
32 Gln	8.38	123.70		4.27	2.15	2.20					58.80	31.11	36.32			
					1.64	2.43										
33 His	7.39	114.09		5.37							53.84	34.34				
34 Ile	8.74	121.20		5.01	1.59	0.65	0.70				60.24	41.01	16.79	15.42		
						1.03							27.53			
						0.41										
35 Leu	9.36	128.68		4.83	1.76	1.42	0.62				53.77	42.34	27.74	24.85		
					1.26		0.61							26.34		
36 Cys	9.03	122.73		6.10	2.70						54.75	29.13				
					2.48											

	H	N	N ε	H α	H β	H γ	H δ	H ε	H ζ	H η	C α	C β	C γ	C δ	C ε	C ζ
37 Thr	8.43	117.03		4.47	4.01						62.97	69.37				
38 Gly	7.49	104.14		3.44							43.78					
39 Asn	10.62	117.56		4.24	3.05 3.05						52.84	34.63				
40 Leu	10.13	124.84		5.31	1.77 1.78	1.42	0.53				59.24	38.73	27.89			
41 Cys	8.02	116.15		5.00	3.69 2.49						60.05	31.47				
42 Thr	8.45	111.89		5.07	4.69						59.52	72.72				
43 Lys	8.67	120.49		4.05	1.83 1.94	1.59 1.43	1.70	3.00			58.90	32.83	24.69	29.23	42.16	
44 Glu	8.31	118.29		4.15	1.94 1.85						60.33	29.49				
45 Ser	7.38	113.70		4.47							62.21	63.42				
46 Tyr	7.25	124.70		3.88	2.95 3.29		7.00	6.64			62.06	37.58		133.33	117.67	
47 Asp	8.67	118.42		4.03	2.68 2.51						57.21	39.35				
48 Tyr	7.28	118.84		4.33	2.97 3.24		6.85	6.39			60.49	37.57		132.64	117.61	
49 Leu	7.81	121.54		3.32	1.02 1.86		0.68 0.71				58.30	40.31		22.09 25.73		
50 Lys	7.63	114.87		4.07	1.57						56.90	31.09				
51 Thr	7.41	109.34		4.06	4.07						63.60	69.70				
52 Leu	7.22	120.32		3.86	0.83 1.81	1.07	0.19 0.27				55.97	43.13		21.48 21.48 25.40		
53 Ala	7.19	117.44		4.34	0.86						51.21	22.40				
54 Gly	8.00	102.94		3.63 3.96							45.87					
55 Asp	6.95	119.52			2.47						53.89	42.08				
56 Val	7.91	124.61		4.25	1.71						61.81	33.97				
57 His	9.03	128.81		4.43	3.07 2.43		6.54				55.72	33.29				
58 Ile	7.98	122.87		4.87	0.83	1.09 0.15 0.68	0.59				59.39	41.35	17.13 28.11	15.42		
59 Val	6.79	116.02		4.96	1.39						58.19	34.17				
60 Arg	7.73	115.85									57.73	32.13				
61 Gly	10.60	118.24									63.92					
62 Asp	9.02	115.74		3.74	2.01 2.01						54.91	36.48				
63 Phe	9.42	117.33		4.48	3.42 2.24						58.50	41.36				
64 Asp	6.72	120.41			2.51 2.64						57.72	42.16				
65 Glu	8.26	120.45		4.13	2.00 2.00						56.95	30.00				
66 Asn	7.91	117.34		4.38	2.41 2.51						52.98	37.49				
67 Leu	7.85	121.89		4.06	1.37 1.37						54.60	41.18				
68 Asn	8.17	115.70		4.50	2.51 2.66						53.32	38.22				
69 Tyr	6.90	120.39					6.85	6.50			54.59	36.38		131.92	118.37	
70 Pro				4.61	2.19 2.40						62.17	32.05				
71 Glu	8.96	119.62		3.99	2.06 2.13	2.41 2.31					59.45	30.54	37.20			
72 Gln	7.47	109.70		5.23	1.73 2.18						53.57	30.91				
73 Lys	8.37	119.10		4.39	1.44 1.44						56.03	37.66				
74 Val	8.02	122.07		4.78	1.69	0.59 0.61					61.13	33.54	21.18 21.05			
75 Val	9.30	126.20		4.21	1.46	0.14 0.15					59.73	34.88	19.92 19.39			
76 Thr	8.37	121.59		5.01	3.85	0.96					61.64	69.42	21.10			
77 Val	8.68	127.68		3.87	1.65	0.56 0.14					61.26	33.47	21.85 19.73			
78 Gly	8.76	115.05		3.55 3.96							46.40					
79 Gln	9.18	124.94		4.03	1.34 1.94 1.84	2.01 2.09					56.04	28.56	34.03			
80 Phe	8.31	118.28		4.45	2.53 3.09						58.41	41.35				

	H	N	N ε	H α	H β	H γ	H δ	H ε	H ζ	H η	C α	C β	C γ	C δ	C ε	C ζ	C η
81 Lys	9.36	123.33		4.67	1.77 1.53						55.94	34.17					
82 Ile	9.33	126.66		4.98	1.56	0.59	0.48				60.16	40.07	17.67	13.36			
83 Gly	9.51	113.54		3.11 4.98							44.32						
84 Leu	8.58	124.85		5.28	0.96 1.50	1.33	0.53 0.54				54.07	47.29	26.75	25.67 24.17			
85 Ile	8.30	121.36			1.90	0.79 0.73					61.40	41.94	14.59 28.53				
86 His	9.54	124.09		4.75							61.84	29.38					
87 Gly	7.90	100.49		3.23 4.47							44.96						
88 His				3.46	2.00 2.91						58.34	28.34					
89 Gln	9.02	120.09		4.47 4.13							56.52	28.59					
90 Val	6.98	123.68		3.81	2.28						62.33	31.87					
91 Ile	7.82	123.97		4.26	1.59	0.64 1.01 1.29	0.69				56.44	40.93	16.63 26.67	12.27			
92 Pro				4.45													
93 Trp		129.72	129.72	4.24	3.18 3.06		7.85	7.56 10.20	7.59 7.04	7.18				128.20	120.92	114.51 122.20	124.66
93 Trp	7.82	119.84									57.10						
94 Gly				3.46 3.91							45.41						
95 Asp	7.42	119.70		4.19							54.93	43.92					
96 Met	8.54	127.78		3.76	1.67 1.67						58.83	32.99					
97 Ala	8.20	120.80		3.95	1.36						55.11	17.53					
98 Ser	7.91	115.52									62.07	63.17					
99 Leu	8.17	124.07		4.20							58.18	41.18					
100 Ala	8.38	122.26		4.82	1.36						54.29	17.56					
101 Leu	7.48	119.10		4.02	1.59 1.77	1.68	0.87 0.83				57.97	41.45	26.82	23.73 24.77			
102 Leu	7.34	120.62		3.68	1.82 0.93	0.70	0.59				57.45	41.18	26.01	23.13			
103 Gln	8.35	120.02		3.58							60.57	27.75	33.46				
104 Arg	7.56	117.99		3.92	1.84 1.75	1.86 1.65	3.24 3.15				58.64	30.22	27.26	43.65			
105 Gln	8.04	120.16		3.90	1.81 2.03	2.11 2.33					58.58	28.66	33.66				
106 Phe	8.43	116.43		4.50	3.14 2.87						56.29	39.83					
107 Asp	8.27	120.00		4.22	3.15 2.24						54.79	39.35					
108 Val	6.86	105.47		4.44							58.51	33.54					
109 Asp	8.71	119.47									57.18	45.32					
110 Ile	7.28	113.77		5.11	1.52	0.55	0.46				58.09	42.42	26.73 16.20	13.61			
111 Leu	8.56	129.01									53.17	44.60					
112 Ile	8.94	129.89		5.07							59.02	33.94					
113 Ser	8.49	119.75		5.66	3.50 3.47						56.76	66.36					
114 Gly	7.18	104.61		3.36 4.87							46.44						
115 His	9.64	125.54									61.26	30.13					
116 Thr	8.17	109.77		3.46	4.09	1.07					62.78	69.13	21.92				
117 His	9.10	113.45		3.87	3.38 3.54						56.18	26.26					
118 Lys	7.50	119.48		4.59	1.56 1.41						54.96	34.21					
119 Phe							7.05							132.50			
119 Phe	8.28	124.96		4.59	2.96 2.47						57.26	38.66					
120 Glu	8.30	125.38		3.69	1.61 1.86	2.00 2.00					55.58	34.11					
121 Ala	8.13	122.69		5.05	1.17						51.41	20.79					
122 Phe	8.14	117.92		4.86	3.21 2.96						56.15	40.47					
123 Glu	8.83	120.99		5.60	2.20 2.20	2.29 2.47					54.02	32.59	36.90				
124 His	9.30	124.96									56.79	33.39					
125 Glu				3.65	1.64	1.96					57.35	27.46					
126 Asn	9.08	108.43		4.71	3.03 3.19						55.46	38.02					

	H	N	N ε	H α	H β	H γ	H δ	H ε	H ζ	H η	C α	C β	C γ	C δ	C ε	C ζ
127 Lys	8.06	121.40									53.66	33.40				
128 Phe				4.94							55.42	41.56				
129 Tyr	7.99	125.97		5.58			6.67	6.51			56.44			133.59	116.63	
130 Ile											58.79					
131 Asn	8.76	125.92		6.19							48.47					
132 Pro				4.42							63.51	32.92				
133 Gly	8.60	107.24		3.16							44.15					
				3.98												
134 Ser	10.09	117.92		5.30	3.06						55.41	63.56				
					3.59											
135 Ala	7.94	133.22		5.11	1.52						54.70	19.32				
136 Thr	7.45	97.40		4.28	4.66	1.08					59.48	69.64				
137 Gly	7.79	110.05		3.30							46.60					
				3.87												
138 Ala	7.69	122.76		3.96	0.97						52.16	19.78				
139 Tyr	7.18	117.90		4.43	3.34		7.03	6.66			56.98	37.99		134.04	117.43	
					3.01											
140 Asn	8.08	116.96		4.17	2.57						51.23	38.97				
					1.61											
141 Ala	8.01	118.59		4.08	1.38						54.20	18.67				
142 Leu	7.74	116.17		4.36	1.62	1.49	0.78				55.37	43.93	26.95	24.87		
					1.54		0.77							22.98		
143 Glu	8.18	119.64		4.53	1.97	2.18					55.69	32.63				
					1.88	2.09										
144 Thr	8.05	111.22		4.48	4.22	1.07					61.20	69.77	21.95			
145 Asn	8.79	123.12		4.81	2.71						53.03	37.45				
					2.81											
146 Ile	7.88	123.74		4.17	1.66	1.04	0.74				58.21	38.26	18.69	10.08		
						1.59							26.46			
						1.20										
147 Ile	8.68	130.16		4.48	1.92	1.25	0.88				58.27	39.93	27.59	12.84		
						0.97							17.33			
						1.62										
148 Pro				4.95	2.15						62.79	32.25				
					1.80											
149 Ser	8.45	116.78		5.96	3.70						57.56	67.36				
					3.87											
150 Phe	8.31	113.66		4.92	2.95		6.31	5.80			57.37	40.09		131.77	129.80	
					3.32											
151 Val	8.79	118.79		5.35	1.99	0.92					59.73	37.31	23.08			
													23.05			
152 Leu	9.40	126.79		5.52	1.75	1.68	0.82				53.43	46.47	27.12	24.83		
					1.67		0.88							24.83		
														24.84		
153 Met	9.80	121.92		5.11	1.52						54.85	36.07				
					1.52											
154 Asp	9.47	125.54		5.16	3.24						52.82	42.61				
					2.35											
155 Ile	8.94	126.84		4.26	1.87	0.94	0.54				62.36	38.66	28.13	12.82		
						1.48							17.92			
						0.75										
156 Gln	8.62	128.03		4.64	1.87						54.98	30.59				
					2.03											
157 Ala	8.92	127.77		3.88	1.52						54.66	17.05				
158 Ser	7.92	114.24		4.37	3.86						58.29	63.59				
					4.00											
159 Thr	8.20	117.42		4.91	4.00	0.98					62.08	71.42	21.87			
160 Val	9.21	124.99		4.38	1.56	0.50					60.64	33.91	21.70			
						0.13							21.70			
161 Val	8.65	129.50		4.10	2.05	0.80					62.41	32.06	21.88			
						0.33							21.92			
162 Thr	9.06	125.56		4.46	4.49	0.98					62.92	69.21	21.71			
163 Tyr	9.56	125.58		5.10	3.07			6.74			56.39	40.36			118.09	
					2.44											
164 Val	8.91	123.39		4.21	1.33	-0.15					60.29	32.77	18.30			
						0.26							20.16			
165 Tyr	8.34	128.25		4.44	1.56						56.95	39.60				
					1.07											
166 Gln	8.62	117.55		5.44	1.78						53.74	34.35				
					1.77											
167 Leu	8.09	124.40		4.84	1.58		0.53				53.71	42.93		26.49		
					0.77		0.40							24.06		
168 Ile	8.50	127.89		4.19	1.73	1.05	0.76				59.70	38.33	17.15	12.68		
						0.82							27.29			
						1.32										

	H	N	Nε	Hα	Hβ	Hγ	Hδ	Hε	Hζ	Hη	Cα	Cβ	Cγ	Cδ	Cε	Cζ
169 Gly	8.76	118.27		3.59							47.20					
				3.89												
170 Asp	8.77	123.98		4.56	2.39						54.35	42.11				
					2.71											
171 Asp	7.74	118.60		4.85	2.55						53.21	44.11				
172 Val	8.42	121.56		4.69	2.04						61.96	32.17				
173 Lys	9.32	131.04		4.56	1.66	1.38	1.58	2.90			54.58	34.62	24.88	28.94	42.20	
						1.38	1.58	2.90								
174 Val	8.31	124.44		4.99	1.54	0.42					60.18	34.19	21.65			
						0.28							20.70			
175 Glu	8.22	127.48		4.57	1.84	2.01					54.47	32.88	35.90			
						2.10										
176 Arg	8.45	123.36		5.13	1.45	0.87	2.97				54.42	33.27	26.66	43.47		
					1.32	0.99	2.97									
177 Ile	9.20	127.81		4.33	1.59	0.99	0.96				60.95	42.73	18.48	15.43		
						1.45							28.15			
						1.17										
178 Glu	8.79	127.10		4.90	1.77	1.90					55.38	32.64	37.22			
						2.02										
179 Tyr	8.54	125.14		4.43	3.18		6.53	6.41			56.51	40.51		132.47	117.26	
					2.96											
180 Lys				4.72	1.51	1.13	1.46	2.79				34.81	24.44	29.14	42.09	
					1.37	1.21	1.46	2.79								
180 Lys	7.73	127.19									53.97					
181 Lys	7.97	122.64			1.56						56.34	33.25				
					1.56											

Appendix Four – Structure Statistics

Structural restraints

NOE-derived distance restraints	**Vps29**	**Varp**
Intraresidue	33	39
Sequential	22	40
Medium $(2 \le \vert i-j \vert \le 4)$	24	17
Long $(\vert i-j \vert > 4)$	40	2
Total	119	98
Intermolecular	49	

Dihedral angle restraints		
χ^1	none	11

Template non-crystallographic symmetry restraints for Vps29 (to pdb 2R17)

Backbone	
Strong[a]	residues 7-19, 60-148
Medium[b]	residues 1-6, 20-59, 149-181
Sidechain	
Medium[b]	residues 1, 3-6, 20-24, 26, 28-29, 31-59, 149-151, 153 155-160, 162, 164, 166-171, 173, 175-181
Weak[c]	residues 2, 25, 27, 30, 152, 154, 161, 163, 165, 172, 174

Statistics for accepted structures

Number of accepted structures	25

Mean XPLOR-NIH energy terms (kcal.mol^{-1} ± S.D.)	
E(total)	-1698.0 ± 22.2
E(van der Waals)	178.8 ± 2.1
E(NCS)[d]	156.8 ± 13.3
E(distance restraints)	37.2 ± 2.4

Restraint violations (average number per structure)

distance > 0.2Å	4.0 ± 1.2
dihedral angles > 5°	0.0 ± 0.0

RMS deviations from the ideal geometry used within XPLOR-NIH

Bond lengths	0.0024 Å
Bond angles	0.68°
Improper angles	0.51°

Ramachandran statistics

	Vps29 1-181	**Varp 710-721**
Most favoured	89.4%	36.4%
Additionally allowed	9.7%	62.8%
Generously allowed	0.2%	0.4%
Disallowed	0.6%	0.4%

Average atomic RMS deviations
from the average structure (± S.D.)

	Vps29 1-181	**Varp 710-721**
(N, Cα, C' atoms)	0.07 ± 0.02 Å	0.23 ± 0.08 Å
(All heavy atoms)	0.14 ± 0.03 Å	0.49 ± 0.07 Å

Table A3.1: Structural statistics for the final set of 25 accepted structures.

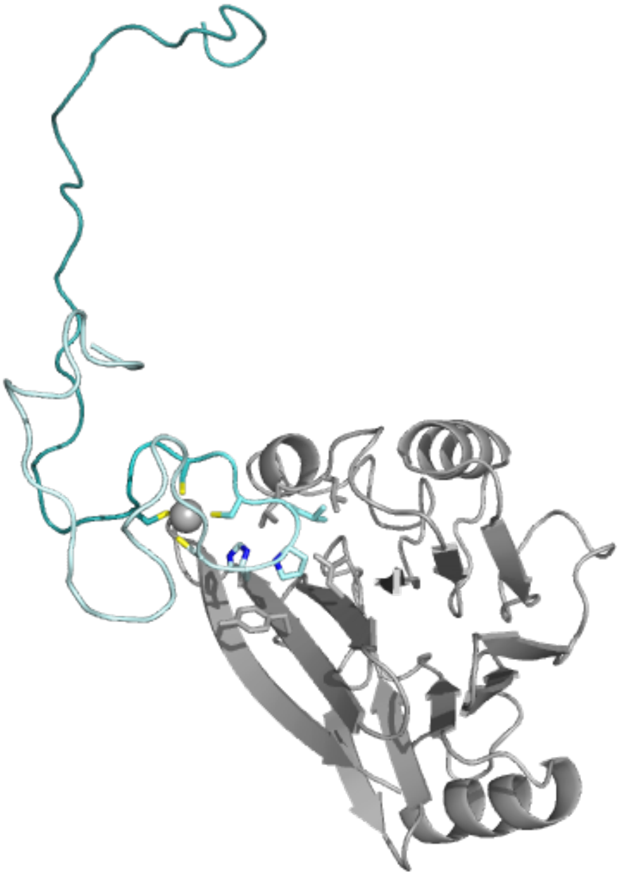
a) Force constant 100.0 kcal.mol^{-1}.

b) Force constant 2.0 kcal.mol^{-1}.

c) Force constant 0.1 kcal.mol^{-1}.

d) NCS energy terms are reported for the structures prior to addition of the unstructured tails of Vps29; no NCS terms are active during the final stage of the calculation when these tails are added.

Appendix Five – Additional Structure Image



A. Figure 5. 1: An image depicting the lowest energy structure from an ensemble of Varp Znk2: Vps29 complex structures. Here, the disordered 'tails' are shown. Varp is coloured from pale cyan at the N-terminus, to deep teal at the C-terminus. Vps29 is shown in grey.