Structural Basis for Rab5 Activation and Effector Specificity in Endosome Tethering: A Dissertation

A Dissertation Presented

By

Eric Lee Merithew

Submitted to the Faculty of the

University of Massachusetts Graduate School of Biomedical Sciences, Worcester

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

April 20th, 2004

Biomedical Sciences

COPYRIGHT

Portions of this thesis appeared in:

Merithew, E., Hatherly, S., Dumas, J.J., Lawe, D.C., Heller-Harrison, R. and Lambright, D.G. (2001) Structural Plasticity of an Invariant Hydrophobic Triad in the Switch Regions of Rab GTPases is a Determinant of Effector Recognition. *J. Biol. Chem.* **276**, 13982-13988.

Dumas, J.J., Merithew, E., Sudharshan, E., Rajamani, D., Hayes, S., Lawe, D.C., Corvera, S. and Lambright, D.G. (2001) Multivalent Endosome Targeting by Homodimeric EEA1. *Mol. Cell* **8**, 947-958.

Lawe, D.C., Chawla, A., Merithew, E., Dumas, J.J., Carrington, W., Fogarty, K., Lifshitz, L., Tuft, R., Lambright, D.G., and Corvera, S. (2002) Sequential Roles for Phosphatidylinositol 3-Phosphate and Rab5 in Tethering and Fusion of Early Endosomes via their Interaction with EEA1. *J. Biol. Chem.* **277**, 8611-8617.

Merithew, E., Stone, C., Eathiraj, S., and Lambright, D.G. (2003) Determinants of Rab5 Interaction with the N Terminus of Early Endosome Antigen 1. *J. Biol. Chem.* **278**, 8494-8500.

Zhu, Z., Delprato, A., Merithew, E. and Lambright, D.G. (2001) Determinants of the Broad Recognition of Exocytic Rab GTPases by Mss4. *Biochemistry* **40**, 15699-15706.

Delprato, A., Merithew, E. and Lambright, D.G. (2004) Structure, Exchange Determinants, and Family-wide Rab Specificity of the Tandem Helical Bundle and Vps9 Domains of Rabex-5. (submitted).

STRUCTURAL BASIS FOR RAB5 ACTIVATION AND EFFECTOR SPECIFICITY IN ENDOSOME TETHERING

A Dissertation Presented

By

Eric Lee Merithew

Approved as to style and content by:

_____
Silvia Corvera, MD, Chair of Committee

_____
Roger Davis, Ph.D., Member of Committee

_____
Kai Lin, Ph.D., Member of Committee

_____
Mary Munson, Ph.D., Member of Committee

_____
A. Andrew Bohm, Ph.D., Member of Committee

_____
David G. Lambright, Ph.D., Dissertation Mentor

_____
Anthony Carruthers, Ph.D.,
Dean of the Graduate School of Biomedical Sciences

Interdisciplinary Graduate Program

April 20th, 2004

ACKNOWLEDGEMENTS

Many people have been a part of my graduate work, as friends, teachers, and colleagues. My mentor David Lambright, first and foremost, has been all of these. The best advisor and teacher I could have wished for, the unmatched dedication to his research, his students and discovery of the elusive "finger of God" is inspiring. Thank you for knowing when to push and when to be patient. Time after time, your effortless grasp of science at its most fundamental level helped me in the struggle for understanding and appreciation.

In my lab, I was fortunate to be surrounded by knowledgeable and friendly people who helped me daily. Special thanks go to Anna Delprato, John Dumas and Craig Stone, without whom this work would never have been completed. My other labmates, including Tom Cronin, Jonathan DiNitto, Sudharshan Eathiraj, Scott Hatherly, Susan Lietzke, Ashwini Mishra, XiaoJing Pan, Chris Ritacco and Zhongyuan Zhu were each a great help in their own way. Many thanks also to Deirdre Lawe and Kim Crowley for their vital collaborative support.

I would like to thank my committee, Silvia Corvera, Roger Davis, Kai Lin and Mary Munson for their insight and critical evaluation of my research. I would also like to thank Andrew Bohm for the time and effort given to the completion of this thesis.

Finally, I would like to thank those closest to me, whose presence helped make the completion of my graduate work possible. Life would not have been the same without my classmates and good friends Corey Smith, Susan Hayes and Leanne Wilson-

Fritch. I will never think about frogs, heated seats and mayonnaise the same way again. I would like to thank my family, and especially my parents, for their absolute confidence in me. The knowledge that they will always be there to push me up that slippery slope makes it just that much less steep. Most of all, my wife Jennifer O'Neil, thank you for your love, patience and understanding, and the assurance that life is just beginning when I step out of the laboratory.

ABSTRACT

As critical regulators of vesicular trafficking, Rab proteins comprise the largest GTPase family, with thirty-eight functionally distinct members and another twenty isoforms in the human genome. Activated Rab GTPases interact with effector proteins involved in vesicle formation, transport, tethering, docking and fusion. The specificity of Rab interactions with effectors and regulatory factors plays a central role with respect to the fidelity of membrane trafficking. Rab recognition determinants and the mechanisms underlying interactions with structurally diverse regulatory factors and effectors are complex and poorly understood. Using Rab5 mediated endocytic transport as a model system, the work described in this thesis provides insight into the structural basis underlying the interaction of effectors and regulatory factors with Rab GTPases. In addition, structural and biochemical approaches have been used to define how specific Rab5 interacting proteins function in the endocytic and recycling pathways. These results establish novel structural and functional concepts that can be tested using family wide analyses of Rab GTPase recognition determinants and regulatory roles in the cell.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Å	Angstrom
ALS2	Amyotrophic Lateral Sclerosis Type 2
Arl	ARF-like
Arf	ADP-ribosylation factor
BCR-ABL	Breakpoint Cluster Region-Ableson Kinase
BT	bud tip
CCV	clathrin coated vesicle
CDR	complementary determining region
CFTR	cystic fibrosis transmembrane conductance regulator
CG	cis-Golgi
COG	conserved oligomeric Golgi
COP	coatomer protein
DH	Dbl homology domain
Dss4	Dominant suppressor of Sec4
EE	early endosome
EEA1	early endosome autoantigen 1
EGFR	epidermal growth factor receptor
EM	electron microscopy
Eps	Epsin
ER	endoplasmic reticulum
ERK	extracellular signal-regulated kinase
ESCRT	endosomal sorting complex required for transport
FYVE	Fab1, YotB, Vac1p and EEA1
GAP	GTPase activating protein
GARP	Golgi-associated retrograde protein
GDF	GDI dissociation factor
GDP	guanosine diphosphate
GEF	guanine nucleotide exchange factor
GGA	Golgi-associated γ-adaptin related ARF binding protein
GM130	Golgi matrix of 130kDa
GPCR	G-protein coupled receptor
GppNHp	guanosine-5'-[β,γ-imido] triphosphate
GTP	guanosine triphosphate
Grb2	growth factor receptor-bound protein 2
Gyp	GAP of Ypt
HB	helix bundle
HMM	Hidden Markov Model
HOPS	homotypic fusion and vacuole protein sorting
Hrs	hepatocyte receptor tyrosine kinase substrate

IHASP	Infantile Ascending Hereditary Spastic Paralysis
Ins(1,3)P_2	inositol-(1,3)-bisphosphate
Ins(1,4,5)P_3	inositol-(1,4,5)-trisphosphate
IPTG	isopropyl-1-thio-β-D-glactopyranoside
IQ	calmodulin binding motif
JNK3	jun-kinase 3
LE	late endosome
LIC-1	light intermediate chain 1
MAD	multiwavelength anomalous diffraction
mant-GDP	2'-(3')-bis-O-(N-methylanthraniloyl)-GDP
MAPK	mitogen activated protein kinase
MEK	MAPK/ERK kinase
MPR	Mannose-6-Phosphate Receptor
Mss4	Mammalian suppressor of Sec4
MTOC	microtubule organizing center
MVB	multivesicular body
NGF	nerve growth factor
NSF	N-ethylmaleimide-sensitive factor
PH	plextrin homology domain
PI3K	phosphatidylinositol-3-kinase
PM	plasma membrane
PRA	prenylated Rab acceptor
PtdIns(3)P	phosphatidylinositol-(3)-phosphate
PtdIns(4,5)P_2	phosphatidylinositol-(4,5)-bisphosphate
PtdIns(3,4,5)P_3	phosphatidylinositol-(3,4,5)-trisphosphate
PX	phagocyte oxidase homology
Rab	Ras like in brain
Rab11-FIP2	Rab11 family interacting protein 2
Rabex-5	Rabaptin-5 associated exchange factor for Rab5
RabF	Rab family
RabGDIα	Rab GDP dissociation inhibitor α
RabSF	Rab subfamily
RBD	Ras or Rab binding domain
REP	Rab escort protein
RILP	Rab-interacting lysosomal protein
RIN	Ras Interaction/Interference
RN-tre	related to the N-terminus of Tre
RTK	receptor tyrosine kinase
SARA	Smad anchor for receptor activation
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SNAP	soluble NSF attachment protein
SNARE	SNAP receptor
SNP	single nucleotide polymorphism

SOD1	superoxide dismutase 1
Sos	Son of sevenless
SPR	surface plasmon resonance
STK	serine/threonine kinase
t-SNARE	target SNARE
TGFβR	transforming growth factor-β receptor
TGN	trans-Golgi network
TIP47	tail interacting protein of 47 kDa
TRAPP	transport protein particle
TSG101	tumor susceptibility gene 101
UIM	ubiquitin interaction motif
v-SNARE	vesicle SNARE
Vps	vacuolar protein sorting
WT	wild type
Yip	Ypt interacting protein
Ypt	yeast protein transport

CHAPTER I

INTRODUCTION

The compartmentalization of the eukaryotic cell requires a complex transport system to distribute lipids and membrane proteins to subcellular locations. The flow of material between donor and acceptor membranes is mediated by vesicles containing various cargo. Transport pathways can be classified into two basic types of membrane trafficking. Exocytosis facilitates the transport of *de novo* proteins from the endoplasmic reticulum (ER) through the Golgi to the plasma membrane. Endocytosis represents the internalization of membrane from the cell surface through endosomes to the lysosome, and is often associated with receptor recycling and degradation (1). In addition, a host of specialized transport steps, including anterograde and retrograde transport as well as sorting between organelles, target cargo molecules to sites of action or recycle transport factors to donor compartments. These processes allow organelles to maintain a protein content and physical structure necessary for normal biological function (2,3). The cellular machinery required to maintain membrane integrity may also play a fundamental role in the formation of subcellular compartments by assembling the protein and lipid components that define them (4).

In each membrane transport event, protein-protein and protein-lipid interactions promote the budding of a segment of membrane from the donor compartment to form a transport vesicle containing cellular cargo. The vesicles are shuttled, often along

cytoskeletal elements, to specific cellular locations where tethering factors promote association of vesicles with acceptor membranes. Docking allows increased association of vesicles with acceptor membranes and further recruitment of cellular factors that promote fusion (2,5,6). The transition from the docking to the fusion step involves mixing of the lipid bilayers to form a single fused membrane (7,8).

Enrichment of sorted cargo proteins in part of a pre-existing membrane promotes the recruitment of adaptor proteins as well as GTPases of the Arf, Sar, and Arl families to define sites of vesicle formation. Coat scaffolds are then recruited to the adaptors, where assembly causes membrane curvature and vesicle bud formation. The coatomer coat complexes (COPI and COPII) facilitate ER to Golgi transport, whereas clathrin coat formation promotes vesicle budding at the plasma membrane, endosomes and the trans-Golgi network (TGN). The GTPase dynamin is necessary for scission of vesicles from donor membranes. Immediately after vesicle budding, scaffold and adaptor protein dissociation results in coat disassembly. Following vesicle formation, long range cellular transport occurs along either the actin or microtubule cytoskeleton. The GTPases of the Rac/Rho family aid in cytoskeletal rearrangements to promote cellular transport. Vesicle lipids or cargo proteins interact with motor proteins, either directly or through adaptor proteins, to propel transport along the cytoskeleton. Along microtubules, kinesins and dyneins promote minus and plus end directed transport, respectively, while myosins coordinate movement along actin.

Vesicle tethering at the target membrane by extended coiled-coil proteins or multisubunit complexes promotes recruitment of docking factors necessary for fusion.

Integral membrane proteins known as soluble N-ethylmaleimide-sensitive factor (NSF) attachment protein (SNAP) receptors (SNAREs) are activated or "primed" by alleviation of inhibitory N-terminal domains by Sec1-like proteins. Vesicle SNARE (v-SNARE), target SNARE (t-SNARE), and SNAP-25 complex formation brings membranes within close proximity (9,10). ATP dependent dissociation of SNARE complexes by NSF proteins completes the membrane fusion event (7,8).

Rab GTPases

Vesicle transport has evolved the directionality and specificity necessary to distribute membrane proteins and lipids between organelles or subdomains (5). The small monomeric GTPases of the Rab family have been implicated in regulation of each transport step and play a substantial role in establishing and maintaining subcellular domains (2,5,6,11). This family of proteins also provides a basis for coupling cellular signals to the membrane transport machinery. As the largest GTPase family, Rab proteins regulate cellular transport, organelle biogenesis as well as protein sorting, and have a localization consistent with their role as regulators of specific intercellular transport steps (2,5,11-14).

Like other GTPases, Rab proteins have been evolutionarily designed as poor enzymes, with relatively slow intrinsic rates of nucleotide exchange and GTP hydrolysis that can be accelerated by accessory factors (15-18). This cycle of GTP turnover promotes Rab regulation by "switching" between two conformational states, an inactive (GDP bound) form and an active (GTP bound) form (19,20). In the active form,

localization to the cytoplasmic face of membranes and subsequent recruitment of effector proteins is central to the role of Rab GTPases in the regulation of membrane transport (Figure 1). Modulation of the GTPase cycle by Rab accessory proteins allows control of trafficking by external signaling pathways (11,17).

Rab homologues in the yeast *S. cerevisiae*, Sec4 and Ypt1, were first identified as Ras-like proteins involved in transport and secretion. Deletion of these proteins is lethal whereas temperature-sensitive mutations fail to complete vesicle secretion and accumulate secretory vesicles following transfer to the restrictive temperature (21,22). Complete analysis of the yeast genome identifies 11 Rab-like GTPases defined by conservation of Rab/Ypt family motifs (12). Disruption of these proteins results in perturbed protein sorting, vesicle transport and organelle morphology (12,23,24). Ypt1, Ypt31/32, and Sec4 function in the exocytic pathway whereas Ypt51 (Vps21), Ypt52, and Ypt53 function in the endocytic pathway. Ypt7 functions in vesicular transport to the vacuole and Ypt6 functions in intra-Golgi and late endosome (LE) to Golgi transport (2). The role of Ypt10, which targets to the Golgi and is non-essential, is poorly understood (25). Ypt11 appears to associate with the class V myosin Myo2p and may be involved in transport of mitochondria to newly forming buds (26). In budding yeast, Ypt proteins involved in exocytosis are essential for viability whereas those that function in endocytosis are not (2). Table 1 summarizes the localization, function and known accessory factors of the Ypt proteins in *S. cerevisiae*.

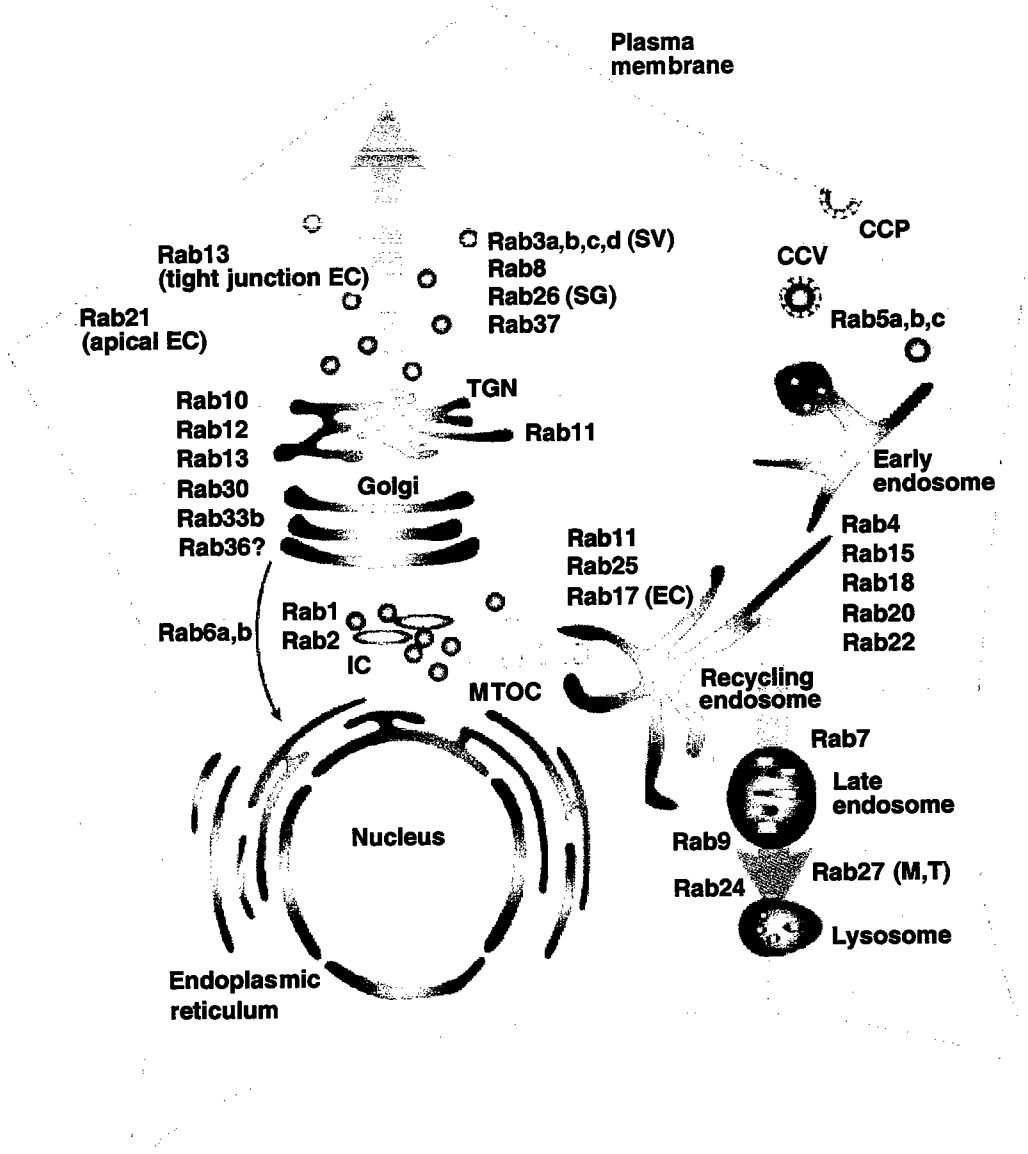


Figure 1. Intracellular localization of Rab proteins. Diagram of the localization of some of the Rab proteins in mammalian cells. Rab GTPases function at various steps in either the exocytic (blue) or endocytic (green) pathways. CCV, clathrin coated vesicle; CCP, clathrin coated pit; EC, epithelial cells; IC, ER/Golgi intermediate compartment; M, melanosomes; MTOC, microtubule organizing center; SG, secretory granules; SV, synaptic vesicles; T, T-cell granules; TGN, trans-Golgi network. Taken from Zerial & McBride, 2001 (11).

Table 1: Ypt Protein Localization and Function

Protein	Localization	GEF	GAP (27)	Function
Ypt1	ER, CG	TRAPP (28,29)		ER to Golgi transport, TRAPP complex (30-32)
Sec4	PM	Sec2 (33), Dss4 (34)	Gyp1-4p	Golgi to PM transport, Exocyst (22,35)
Ypt6	LE, CG	Ric1/Rgp1 (36)	Gyp2,3,6p	LE to Golgi and intra-Golgi transport (37,38)
Ypt7	LE, VA	Vps39 (39)	Gyp4,7p	LE to vacuole transport (40,41)
Ypt10	CG (25)			
Ypt11	BT			Movement of mitochondria to bud (26)
Ypt31	TGN	TRAPP (28)		Budding from the TGN (42,43)
Ypt32	TGN	TRAPP (28)		Budding from the TGN (42,43)
Ypt51	PM, EE	Vps9 (44,45)	Gyp1,3p	PM to endosome transport (46)
Ypt52	PM, EE	Vps9 (44,45)	Gyp1,3p	PM to endosome transport (46)
Ypt53	PM, EE	Vps9 (44,45)	Gyp1,3p	PM to endosome transport (46)

Rab family members were first isolated in mammals from rat brain cDNA libraries on the basis of similarity to Sec4 and Ypt1 (47). Over 60 Rab family members have been identified with 38 functionally distinct Rab GTPases and the remainder representing isoforms resulting from gene duplication or mRNA splicing (14,48,49). The dramatic expansion of the Rab family compared to yeast likely reflects the increased regulation and tissue specific expression often associated with the more complex mammalian systems.

Some Rab family members are involved in many trafficking events. For example, Rab5 serves as a master regulator of endosome trafficking and is essential for plasma membrane to endosome transport, early endosome (EE) homo- and heterotypic fusion and functions in EE to LE transport (50-53). Other Rab proteins play a more specific role, as is evident from either cellular function or tissue distribution. Whereas Rab13 has a specialized function in tight junction formation, Rab17 is only expressed in polarized

epithelial cells (54-56). Additionally, many mammalian Rab GTPases with overlapping cellular distributions coordinate discrete steps along the same pathways (57,58). Table 2 summarizes Rab protein localization and function in mammalian systems.

Table 2: Rab Protein Localization and Function (2,3,11,48,75)

Protein	Localization	Function	Isoforms
Rab1	ER, CG	ER to Golgi and intra-Golgi transport	A,B
Rab2	ER, CG	Golgi to ER retrograde transport	A,B
Rab3	SV, SG, PM	synaptic vesicle and chromaffin granule secretion	A,B,C,D
Rab4	EE, RE, PM	sorting/recycling in early endosomes	A,B
Rab5	EE, CCV, PM	CCV to EE transport, EE motility	A,B,C
Rab6	ER, CG, TGN	retrograde Golgi to ER and intra-Golgi transport	A,B,C
Rab7	LE, LS	EE to LE and LE to lysosome transport	
Rab8	TGN, SV, SG, PM	TGN to PM transport and regulated secretion	A,B
Rab9	TGN, LE	LE to Golgi transport	A,B
Rab10	TGN		
Rab11	RE, TGN, PM	endosomal recycling, PM to Golgi transport	A,B
Rab12	CG, SG		
Rab13	ETJ	tight junction formation	
Rab14	CG, TGN, EE	EE to Golgi recycling	
Rab15	EE, RE, PM	EE to PM recycling	
Rab17	EE, PM	transcytosis	
Rab18	EE, RE, PM	EE to PM recycling	
Rab19			
Rab20	EE		
Rab21	TGN, PM	TGN to PM polarized secretion	
Rab22	EE, PM	interaction between EE and biosynthetic pathways	A,B,C
Rab23	EE, PM	PM to EE transport, neural patterning	
Rab24	ER, CG, LE		
Rab25	EE, RE, PM	apical recycling and transcytosis	
Rab26	SG	regulated secretion	
Rab27	MS	granule secretion, melanosome transport	A,B
Rab28			
Rab29			
Rab30	CG		
Rab32	MS	melanosome transport	
Rab33	CG	intra-Golgi transport	A,B
Rab34	LE, LS	regulation of lysosome morphology	
Rab35			
Rab36			
Rab37	SG	regulated secretion	
Rab38	MS	melanosome transport	
Rab39	CG	promotes endocytosis	A,B
Rab41			

ER= endoplasmic reticulum, PM= plasma membrane, TGN= trans Golgi network, SV= secretory vesicle, SG= secretory granule, LE= late endosome, EE= early endosome, MS= melanosome, CCV= clathrin coated vesicle, ETJ= epithelial tight junctions, CG= cis-Golgi, RE= recycling endosome, PG= phagosome, LS=lysosome

The Ypt proteins localize and appear to regulate transport in the same cellular compartments as their closest mammalian homologues. Within the exocytic subfamily, Ypt1, Sec4 and Ypt31/32 are functionally related to Rab1, Rab8 and Rab11 respectively, which is consistent with comparisons of sequence similarity. In the endocytic group, Ypt51/52/53, Ypt6 and Ypt7 are equivalent to Rab5, Rab6 and Rab7 respectively although the vacuolar sorting pathway in yeast lacks much of the complexity of the endosome system in mammals (46,59,60). In many cases, Rab specific accessory and effector proteins are also evolutionarily conserved, such as the yeast Ypt51 interacting protein Vac1p and the mammalian Rab5 effector Rabenosyn-5 (61). Additional evidence for a functionally conserved core of membrane trafficking proteins is provided by protein substitution experiments using orthologs. The mouse Rab1A can functionally substitute for Ypt1 in *S. cerevisiae* and Rab5 can replace Ypt5 in *S. pombe* (62,63). At present, the only Rab deletions in mammalian systems are Rab3 deficient mice. Rab3A deficient mice have mild phenotypic effects; however, deletion of all Rab3 isoforms (Rab3A, Rab3B, Rab3C and Rab3D) is postnatal lethal (64,65).

The common localization for evolutionarily conserved Rab homologues is a key point in understanding their function. In many cases, Rab localization is one of the defining features of an organelle or vesicular compartment and there appear to be discrete membrane domains within these compartments where each Rab functions (4,57).

Mislocalization of individual Rab family members often leads to aberrant trafficking events or morphological changes in the compartment with which it is associated. The C-terminal hypervariable region of Rab family members, which is doubly prenylated *in vivo*, contributes to targeting of Rab GTPases to membranes (66,67). Rab5 and Rab2 chimeric proteins with C-terminal sequences substituted for the equivalent region of the late endosomal Rab7 localize to late endosomes (68). However, relocalization of chimeric Rab proteins based on the variable C-terminal region does not confer function and is not always accurate, consistent with a requirement for additional determinants for targeting and function (67,69,70). The targeting of Rab GTPases by selective membrane factors, coupled with the subsequent recruitment of effector proteins, is hypothesized to promote a steady-state membrane environment that defines a compartment or organelle. For example, the Rab9 effector TIP47 has an increased affinity for the Mannose-6-Phosphate Receptor (MPR) when complexed with Rab9, creating an enriched environment for MPR recycling (71,72).

The requirement for Rab proteins in the regulation of membrane trafficking suggests a potential role in disease (73-75). Mutations in Rab GTPases and other genes encoding proteins essential for specific transport steps have been identified. Griscelli syndrome, a disease that causes depigmentation and immune deficiency, is due to mutations in Rab27A (76-78). Loss of the Rab accessory factor, Rab escort protein (REP), causes choroideremia, a genetic disease resulting in retinal degeneration (79). Mutation in another Rab accessory factor, Rab GDP dissociation inhibitor α (RabGDIα) is associated with X-linked mental retardation (80,81). Diseases or infections that utilize

or alter trafficking pathways may also function through manipulation of Rab GTPase pathways. A number of diseases are caused by aberrant transport of proteins, such as GLUT4 transporters in non-insulin dependent diabetes mellitus, the cystic fibrosis transmembrane conductance regulator (CFTR) in cystic fibrosis, or β-amyloid in Alzheimer's disease (2,82-84). In addition, phagosomes containing the pathogens *M. tuberculosis* or *L. donovani* block recruitment of Rab7, delaying phagosome maturation and lysosomal degradation, thereby allowing for propagation of infection (85,86).

Rab Family Structure and Phylogenetics

The crystal structure of Rab3A revealed an overall fold similar to other Ras-like GTPases, with a six-stranded beta sheet surrounded by five alpha helices (Figure 2) (18,87). GTPases typically conserve five guanine-nucleotide binding motifs that underlie a common mode of nucleotide binding with some differences in nucleotide contacts that appear to account for observed variations in affinity and intrinsic rates of GTP hydrolysis (Figure 3) (18,88,89). Conversion from the GDP to GTP bound state is accompanied by the ordering of the flexible "switch regions" (switch I: α1/β2 loop; switch II: β3/α2 loop, α2 helix and α2/β4 loop) into an activate conformation (20). The comparison of GTP and GDP bound structures of Sec4p demonstrates that the conformational change is localized to these regions (88).

Crystallographic studies of Rab GTPases have also identified structural motifs and modes of effector interaction that are distinct from other GTPase families. A glycine insertion in the switch I region of Rab proteins facilitates a more extensive and highly

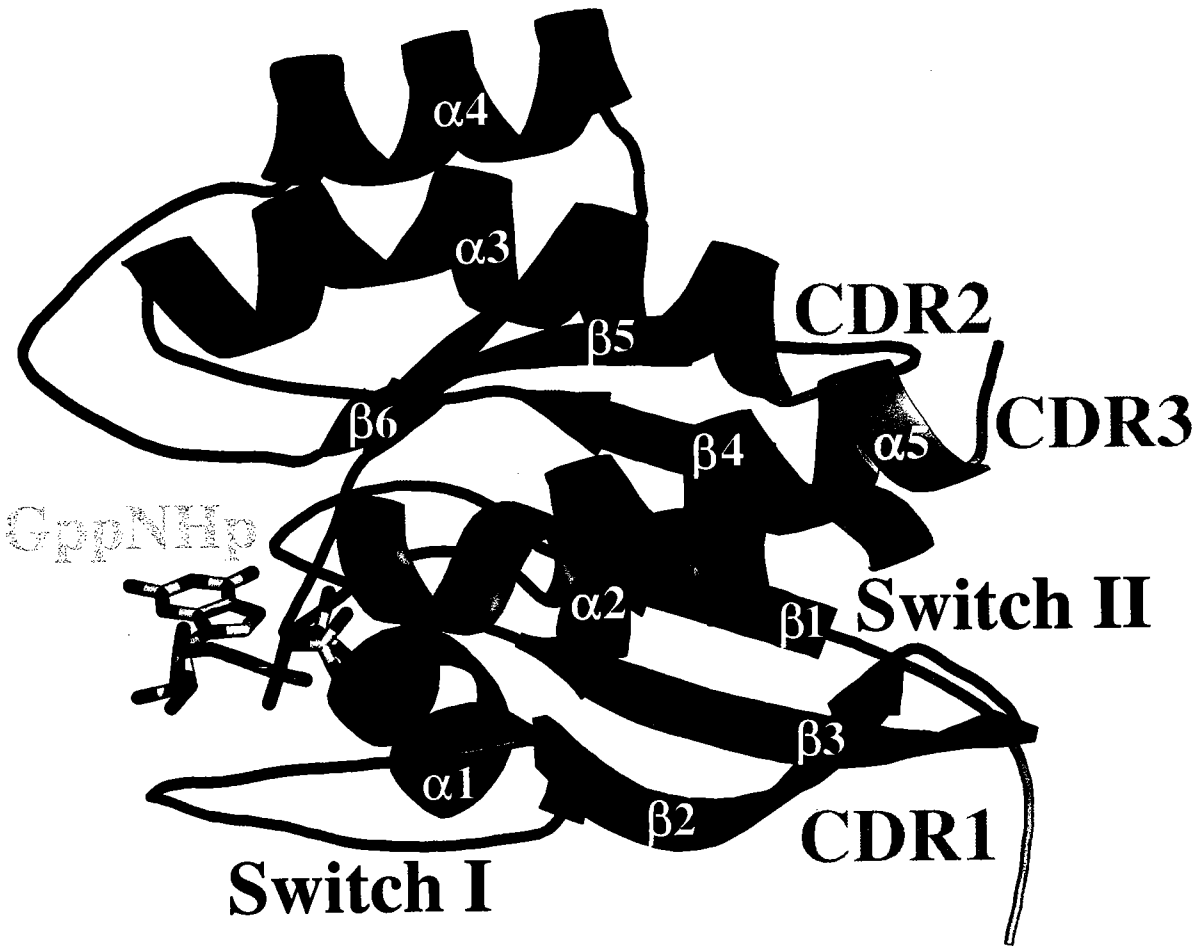


Figure 2. Overall structure of Rab3A. Ribbon representation of the Rab3A structure (from Dumas et al., 1999) (18) with the putative switch regions highlighted in green, the regions corresponding to the complementary determining regions (CDRs) (see text and Figure 5) in light blue and the nucleotide in yellow.

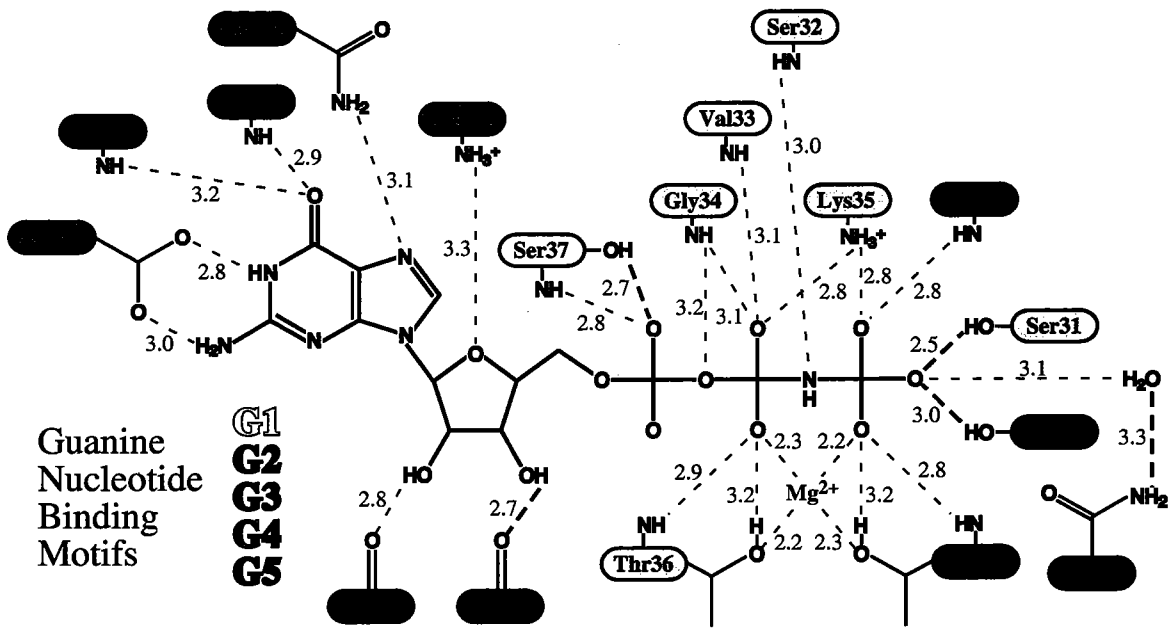


Figure 3. Nucleotide binding of Rab3A. Schematic of contacts between Rab3A and the GTP analog GppNHp from the Rab3A structure (Dumas et al., 1999) (18). Colors denote contributions from the five guanine nucleotide binding motifs (G1-G5).

ordered hydrophobic interface between the switch regions compared with other GTPases (Figure 4) (18). Rab GTPases have flexible N- and C-terminal hypervariable regions that were either excluded or disordered in the known Rab GTPase structures. The structure of the complex between a constitutively active mutant of Rab3A and the effector, Rabphilin-3A, revealed an interaction epitope that extends from the switch interface to an adjacent pocket formed by three hypervariable complementary determining regions (CDRs) corresponding to the N- and C-termini of the GTPase domain and the α3/β5 loop (Figure 5) (90).

A number of phylogenetic studies have evaluated Rab protein families from multiple organisms (14,60). Members of the Rab GTPase family are defined by five Rab family specific sequences (RabF motifs) located either in the putative switch I and II regions or in the adjacent β3 and β4 strands. Four additional subfamily-specific regions (RabSF motifs) define eight subfamilies within the Rab family. The RabSF motifs include all three Rab CDRs as well as a region corresponding to the α1 helix and most of the α1/β2 loop (Figure 6) (48,60). Overall, the N- and C-terminal extensions of Rab proteins are the most variable regions whereas those involved in nucleotide binding are highly conserved (68,87,91).

A key question concerns the molecular and structural mechanisms by which Rab GTPases generate highly selective recognition of diverse effectors while retaining broad specificity for particular accessory factors. Phylogenetic, structural, biochemical and genetic studies of chimeric and mutant Rab proteins have implicated the hypervariable CDR regions as partial but not exclusive determinants of functional specificity (69,92).

Figure 4. The switch interface of Rab3A in the active conformation. Like other GTPases, the conformation of the switch regions is coupled to the presence of GTP. In Rab GTPases, the active conformation is further stabilized by extensive hydrophobic interactions involving highly conserved residues in the switch regions. This interface is facilitated by a conserved glycine insertion (G56) in the switch I region. Rab3A structure from Dumas et al., 1999 (18).

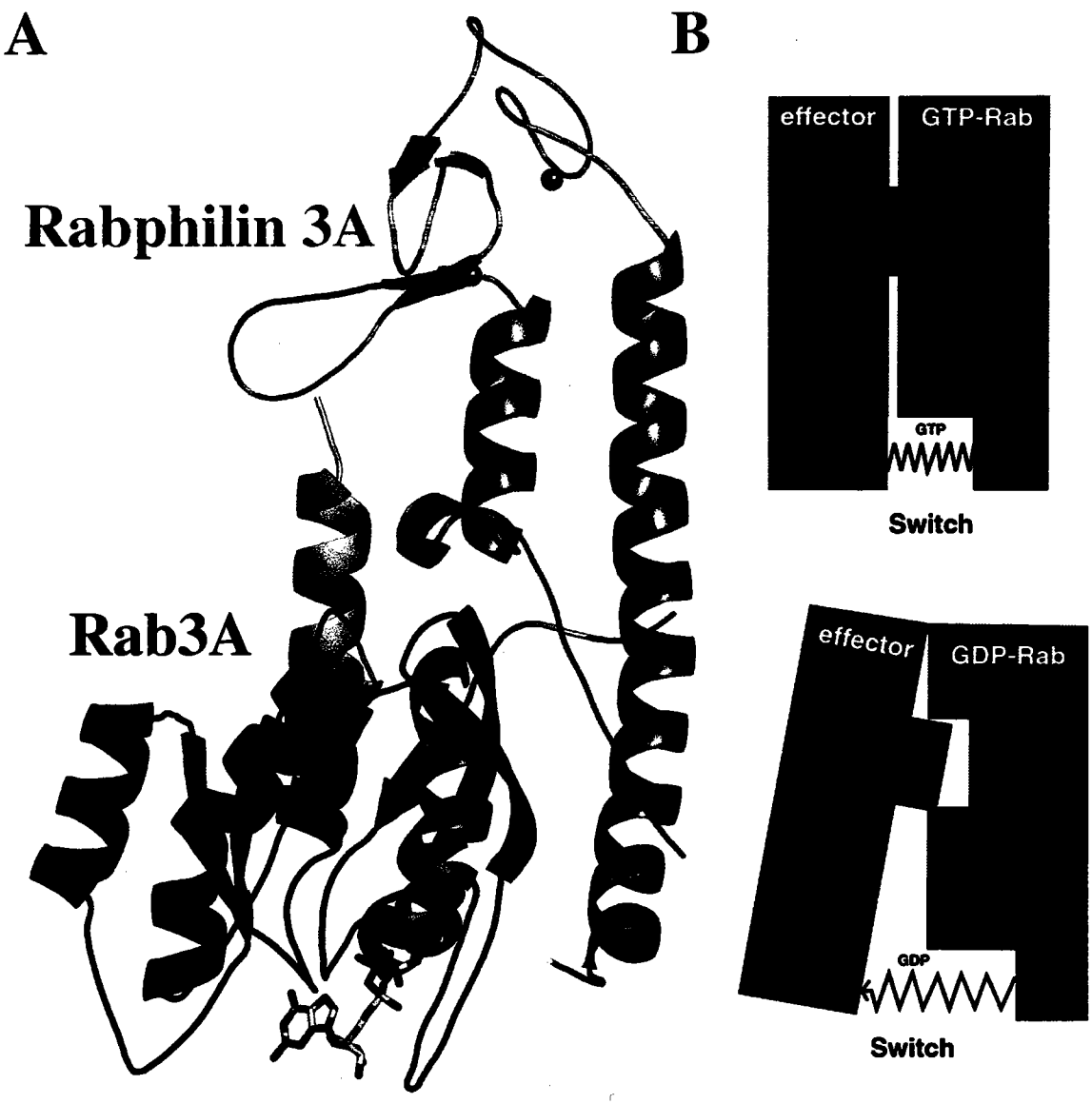


Figure 5. The Rab3A/Raphilin-3A structure. (A) The interaction epitope extends from the relatively conserved switch interface to an adjacent pocket formed by three hypervariable CDRs corresponding to the N- and C- termini of the GTPase domain and the α3/β5 loop. The orientation of Rab3A is rotated 90° relative to that in Figure 2, Raphilin-3A is shown in orange with coordinated Zn^{2+} ions shown as gray spheres **(B)** Hypothetic Rab switch model in which binding specificity is determined by the CDRs. GTP hydrolysis leads to a conformational transition in the switch I and II regions, resulting in dissociation of the Rab3A/Rabphilin-3A complex. Figure adapted from Ostermeier & Brunger, 1999 (90).

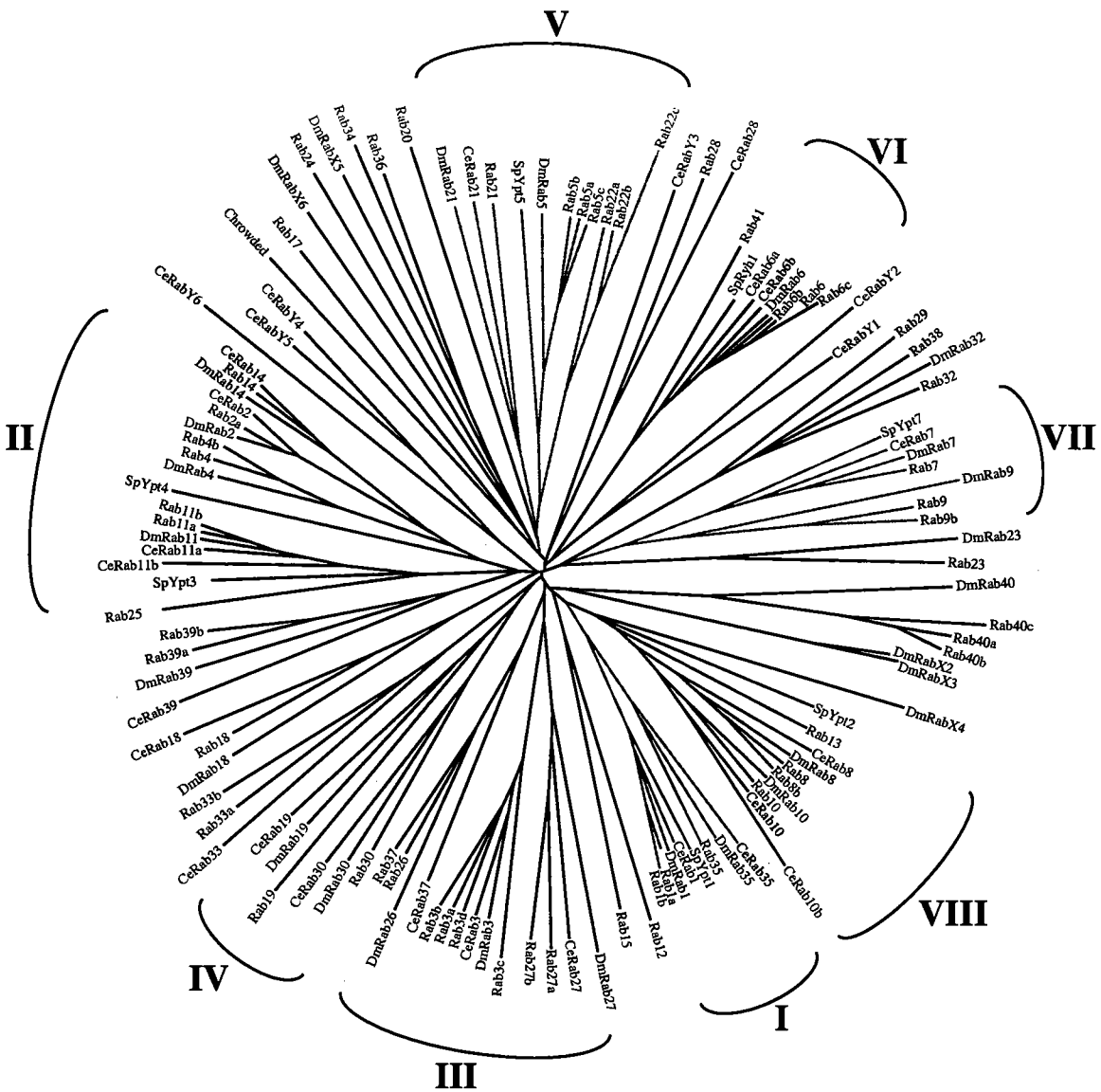


Figure 6. Neighbor-joining dendogram of Rab family members. The roman numerals (I - VIII) and colored branches represent Rab subfamilies. Rab proteins with similar functions group together irrespective of organism suggesting that many are related or direct orthologs. While it does serve as a guideline, phylogenetic relationships do not account for all of the observed functional relationships between Rab GTPases. Adapted from Pereira-Leal & Seabra, 2001 (60).

Furthermore, interactions involving hypervariable regions can not explain the ability of RabGDI to recognize all Rab GTPases, yet still discriminate against other GTPase families (93-96). Likewise, a number of regulatory factors and effectors are known to interact with two or more Rab GTPases. These observations imply the existence of specificity determinants that are common to Rab GTPases or Rab subfamilies, but not other GTPase families.

The Rab GTPase Cycle

As previously described, Rab GTPases cycle between GDP and GTP bound forms. This conversion allows Rab proteins to function as regulatory factors capable of integrating signaling with membrane trafficking. The Rab GTPase cycle involves a series of steps, mediated by accessory factors, as illustrated in Figure 7. Rab GDP dissociation inhibitor (RabGDI) sequesters inactive Rab proteins in the cytoplasm and promotes recycling to donor membranes. Interaction with GDI dissociation factors (GDFs) accelerates association of prenylated Rab proteins with membranes, where they are activated by guanine nucleotide exchange factors (GEFs). Effector recruitment persists until GTP hydrolysis occurs or GTPase activating proteins (GAPs) accelerate the slow intrinsic hydrolysis rate, returning the Rab protein to the inactive state.

Cytoplasmic Sequestration and Membrane Association. Following synthesis, Rab proteins associate with Rab Escort Protein (REP) in the cytoplasm. This complex is recognized by Rab geranylgeranyl transferase, which transfers two 20 carbon geranyl moieties to the conserved cysteine motif present in the C-terminal tails of most Rab

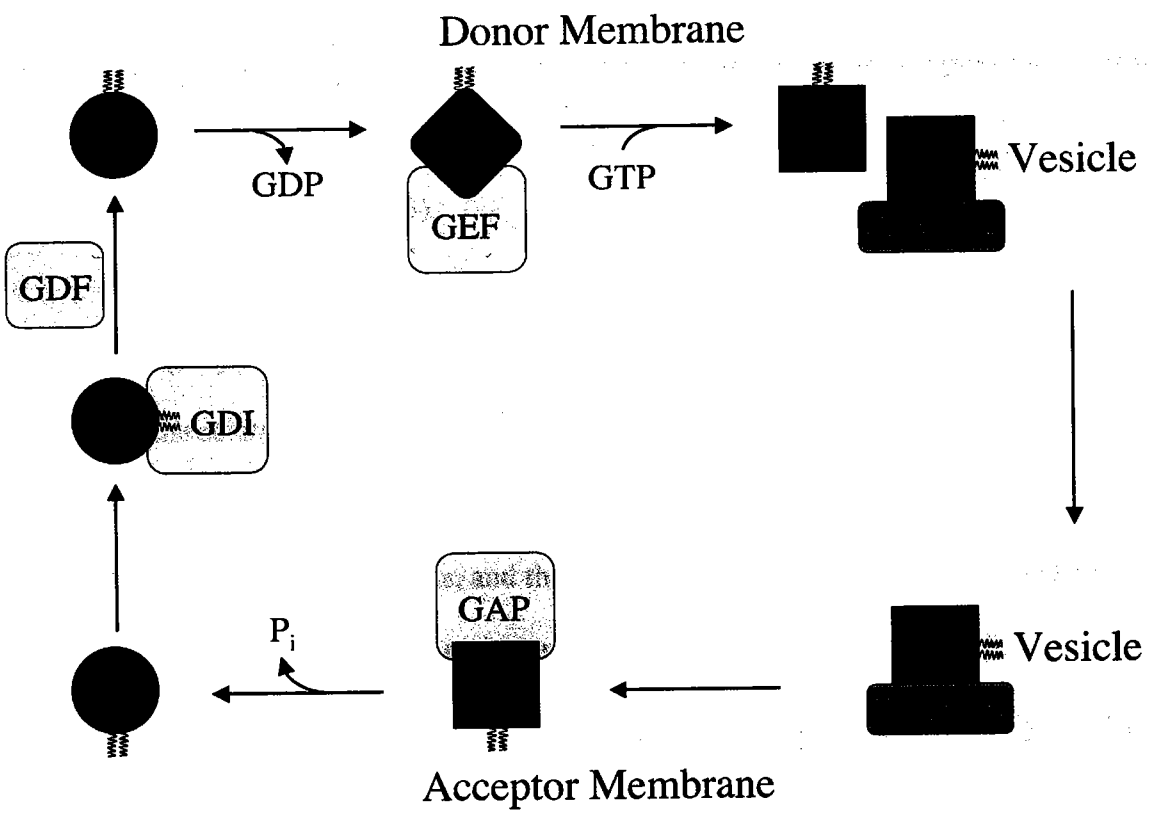


Figure 7. Regulators of the Ypt/Rab GTPase nucleotide cycle. Ypt/Rab GTPases cycle between GTP and GDP bound forms. Accessory factors which regulate nucleotide exchange or GTP hydrolysis control the activation of Ypt/Rab proteins. The activated Ypt/Rab GTPases recruit effector proteins which coordinate distinct membrane transport steps.

GTPases. REP also maintains the solubility of Rab proteins until targeted to the appropriate membrane for activation (97-99). Within subsequent rounds of the GTPase cycle, Rab proteins may dissociate from target membranes and return to donor membranes through association with RabGDI. RabGDI binds the GDP form of Rab proteins and also associates with the prenyl groups to maintain Rab solubility in the cytoplasm (93,100-103). Sec19, the RabGDI in budding yeast, is essential for viability and prenylation. Association with REP/GDI is essential for the targeting of Rab proteins to specific membranes (104,105).

Rab proteins must interact with membranes through their prenyl groups to function properly. In order to dissociate from RabGDI, the complex is thought to interact with GDI dissociation factors (GDFs) (2,4,104). Originally described as activities that accelerate dissociation of Rab GTPases from RabGDI, putative GDFs have recently been identified (104). The Yip proteins, and their mammalian homologues PRA-1 and PRA-2, are a family of conserved integral membrane proteins which have been shown to accelerate dissociation of Rab/RabGDI complexes (95,106-109). Based on their localization, GDFs may also contribute to membrane targeting. The Yip1p (PRA-2) protein targets the ER and Golgi whereas Yip3p (PRA-1) localizes to the Golgi and late endosomes. Each Yip/PRA protein may promote release of a different subset of Rab proteins from RabGDI (95,106,110,111). Currently, there is not enough data to distinguish whether each Rab has its own GDF localized to specific target membranes or whether a few GDFs recognize a broad set of Rab GTPases, with additional factors that promote sorting to discrete membrane subcompartments (111).

Guanine nucleotide Exchange Factors (GEFs). In conjunction with proper membrane localization, Rab proteins must be converted to the GTP-bound form to recruit effector proteins necessary for membrane trafficking. Rab GEFs accelerate the intrinsically slow release of GDP by preferentially binding the nucleotide free form (112,113). Subsequent GTP binding results in activation and GEF dissociation. GEFs with limited sequence similarity have been identified for each of the yeast Ypt proteins (Table 1). A number of mammalian GEFs have also been identified including homologues of the yeast GEFs, such as Rab3GEF and the Vps9 domain (2).

Structural information about GEFs is presently limited to other GTPase families, with the exception of Mss4, which exhibits weak exchange activity for several exocytic Rab proteins (114-119). The overall sequence and fold of GEF domains show no obvious similarity and it is unlikely that Rab recognition occurs by a common mechanism. Although acidic residues are typically important for the GTPase/GEF interaction, the function is not the same in all cases. Within the Rab family, GEFs have been identified as components of larger complexes, which also contain effector proteins for the same Rab GTPase (2). For example, Vps39, a Ypt7 GEF, is a component of the HOPS complex whereas the Rab5 GEF Rabex-5 forms an obligate heterodimer *in vivo* with the Rab5 effector Rabaptin-5 (39,120). These complexes may be important for amplifying and stabilizing the population of activated Rab proteins in localized regions of donor, acceptor, and/or vesicle membranes.

GTPase Activating Proteins (GAPs). Once vesicle fusion has occurred, it is important to inactivate Rab proteins so that constitutive recruitment of effector proteins

does not take place. The intrinsic rate of GTP hydrolysis varies among Rab family members, but overall the reaction is relatively slow. A tighter control of Rab inactivation is sometimes necessary to control specific trafficking events (121,122). GAP association stabilizes the transition state for GTP hydrolysis and dramatically accelerates this process. The Gyp family of GAPs has been identified for Ypt proteins in yeast (Table 1) and, unlike GEFs, do not show a high degree of specificity for particular Ypt proteins (123-126). Deletion of Gyp family members in yeast has no effect on viability (124,127,128). Likewise, a constitutively active mutant of Rab5 promotes fusion, suggesting that GAPs are negative regulators of vesicle fusion (16,129). Thus, Rab GAPs may play a less critical role in the regulation of membrane transport than GEFs, by restricting specific processes or preventing inappropriate fusion events that would be detrimental to the cell (2).

Gyp proteins share a common fold and interact with various members of the Ypt family. Information about the catalysis of Rab GTP hydrolysis comes from the structure of the yeast Ypt GAP, Gyp1, and the structures of GAP complexes from other GTPase families (27,130-133). In contrast to Rab GEFs, Rab GAPs such as the Gyp proteins and GapCenA share a high degree of homology (27,130). The active site of Gyp proteins is similar to Ras-GAP and Cdc42-GAP, including a conserved arginine "finger" that defines a universal mechanism to promote hydrolysis (27,133,134).

Rab Regulation of Membrane Trafficking

A primary function of the GTPase cycle is to control the population of activated Rab proteins. The subsequent recruitment of effectors is central to the ability of Rab GTPases to regulate membrane transport. The requirement for Rab proteins in various steps of membrane transport have been examined using a wide range of techniques including protein depletion, cell-free transport assays and dominant-negative or constitutively active mutations (2). The discovery that a single Rab can control more than one transport step suggests that Rab GTPases are capable of recruiting multiple effectors.

Vesicle Formation. Rab proteins and their effectors have been shown to be necessary for generating transport vesicles. The yeast Ypt31 and 32 proteins are required for release of vesicles from the yeast TGN, Rab1 is essential for ER vesicle formation, whereas Rab5 is required for the formation of clathrin-coated pits (43,135,136). Rab/effector complexes utilize a number of mechanisms to promote vesicle budding. The interaction of Rab9 with TIP47 enhances the affinity for MPR cytoplasmic tails resulting in MPR clustering and vesicle formation (58,72). The Rab11 effector Rabphilin-11 interacts with Sec13, a component of the COPII coat complex to promote vesicle formation by coat recruitment (137).

Vesicle Motility. The actin and microtubule cytoskeletons serve as "roadways" for the movement of vesicles throughout the cell. Overexpression of Rab8 results in cytoskeletal rearrangements whereas Rab5 stimulates endosome/microtubule association and movement of endosomes along microtubules (52,138). Microtubules facilitate

vesicle movement over large distances within the cell and Rab GTPases have been shown to interact directly or through adaptor proteins with kinesins and dynein (139). For example, the Rab6 effector Rabkinesin-6 is a kinesin motor protein, Rab4A binds the human dynein LIC-1 and the Rab7 effector RILP has been shown to enhance recruitment of the dynein-dynactin complex to late endosomal structures (140-143).

Rab dependent actin cytoskeleton transport has also been demonstrated by identification of myosin motors as Rab effectors (139). The well characterized Rab27A effector, Myosin Va, regulates actin based melanosome transport to the dendritic extensions of melanocytes (78,144-146). Mutation of either Rab27A or Myosin Va results in a failure to deposit melanosomes at the dendrite and depigmentation (78,139,144). Another unconventional myosin, Myosin Vb, interacts with the Rab11 effector Rab11-FIP2 and colocalizes with Rab11A to the recycling endosome (147-149). Rab/motor protein interactions have also been observed in yeast where the type V myosin homologue Myo2p is required for Sec4p dependent transport of secretory vesicles to the bud tip (150,151).

Vesicle Tethering. Following translocation, physical links must occur over considerable distances to facilitate SNARE complex formation and membrane fusion. Because Rab GTPases and SNARE proteins are distributed throughout membranes, this process of membrane "tethering" may also confer specificity on a fusion reaction (152). The yeast SNAREs Sso1p and Sso2p are distributed over the entire plasma membrane, yet vesicles only fuse with specific regions within the membrane where the exocyst complex is located (153). Two types of proteins required for membrane trafficking have

been identified as potential tethering factors, a group of extended coiled-coil proteins and a number of large multisubunit complexes (152). Early endosome autoantigen 1 (EEA1), Rabenosyn-5 and the Rabaptin family members are long coiled-coil containing proteins that may function as Rab5 dependent endosomal tethers (61,154,155). Ypt1 is implicated in GM130 and Giantin dependent ER vesicle docking with the Golgi via the coiled-coil protein Uso1p, whose mammalian homologue p115 interacts with Rab1 (136,156,157). Membrane trafficking complexes that have been implicated in tethering include the homotypic fusion and vacuole protein sorting (HOPS or Class C) complex, the conserved oligomeric Golgi (COG) complex, the Dsl1p complex, the Golgi-associated retrograde protein (GARP) complex, the transport protein particle (TRAPP I & II) complexes and the exocyst (152). Tethering factors have also been identified for other GTPase families, including the Arf and Arl GTPases that interact with Golgin proteins (158,159).

SNARE complex formation *in vitro* is too slow to account for proper membrane fusion *in vivo*, emphasizing the need for additional factors to accelerate this process (160). The role of each of these complexes or proteins differs, but all appear to function upstream of the stable docking of vesicles to their target membranes. Tethering factors may act kinetically, by maintaining the vesicle within a certain vicinity of the target membrane and increasing the probability of SNARE-mediated fusion (152). The yeast Uso1p contains no enzymatic functions, but directly interacts with integral and membrane bound proteins to shift the steady state of diffusion, favoring membrane clustering (136,157). Tethers may also act by promoting fusion through direct activation of Rab GTPases on the vesicle or target membrane (152). A number of tethering complexes

contain factors capable of promoting Rab activation, including the Ypt1 and Ypt31/32 GEF activity of the TRAPP complexes and the Rab5 GEF activity of Rabex-5, which binds Rabaptin-5 (28,29,53,161). Although the distribution of tethering factors and Rab proteins is an area of active investigation, Rab5 and its putative tethering factor Rabaptin-5 are required on both vesicle and target membranes for homotypic fusion. EEA1 is only necessary on early endosomes in order to promote heterotypic fusion with clathrin coated vesicles (162,163). Finally, long range tethers may undergo conformational changes or give way to shorter range effector interactions to bring membranes into closer proximity.

Vesicle Docking and Fusion. Although Ypt/Rab proteins act during budding, transport and tethering, Rab associated factors and complexes may also play a thermodynamic role in docking and fusion events. As noted previously, SNARE complex formation is considered the final step in promoting membrane fusion. Short range paring of vesicle (v-SNARE), target (t-SNARE) and SNAP-25 proteins to form high affinity four helical bundles promotes contact between the lipid bilayers of two membranes (164-166). The N-terminal domains of some SNARE proteins and the Sec1-like family of t-SNARE inhibitors reduce SNARE complex formation and aid in fusion specificity (167-169). The N-terminal domain of the yeast t-SNARE Sso1p is essential; however, the constitutively open (non-inhibitory) mutant is viable, suggesting that the N-terminal domain also possesses an activating function *in vivo* (170,171). Rab effectors and tethering complexes that interact with the SNARE or Sec1-like proteins may bridge the gap between tethering and activation of the fusion machinery (Figure 8) (152,172,173).

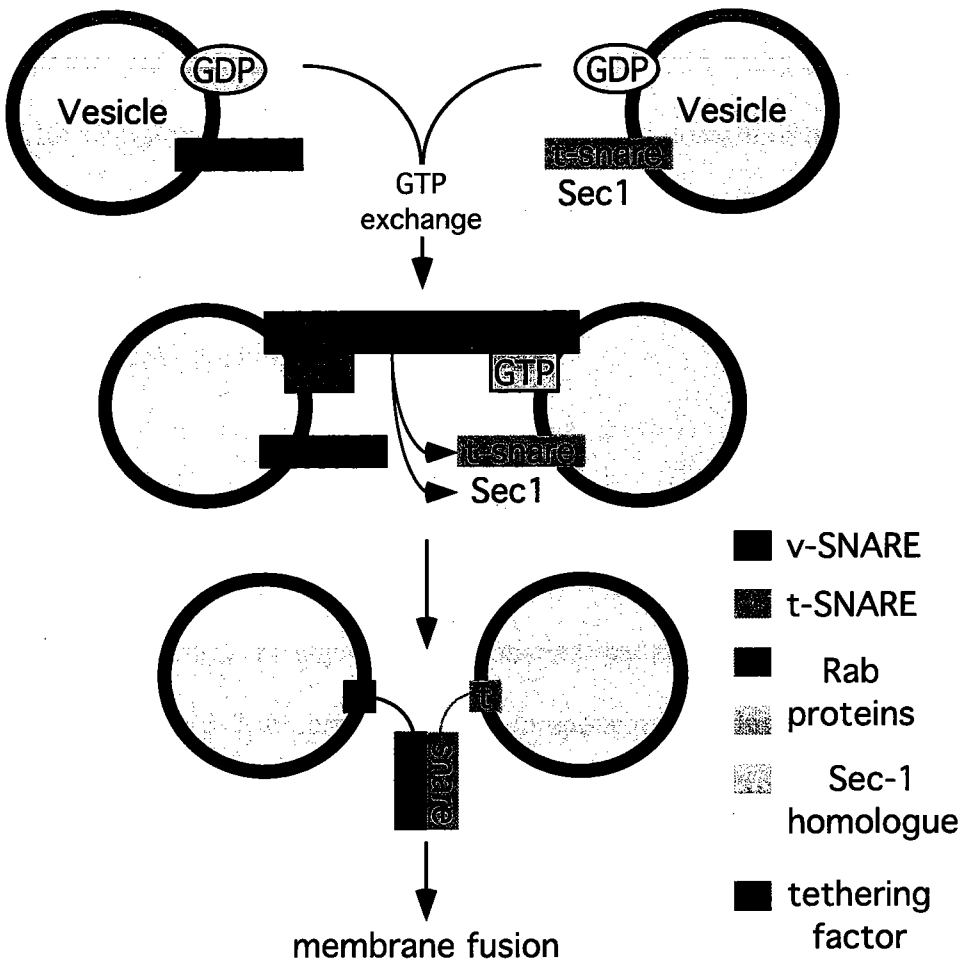


Figure 8. Model for the tethering and fusion of vesicles. Activated Rab GTPases recruit tethering proteins or complexes and can interact with components of SNARE/Sec1 complexes. Assembly of the activated SNARE complex promotes membrane fusion. Figure adapted from Clague, 1999 (356).

In endosome fusion, recruitment of the SNARE proteins syntaxin-6 and 13 by EEA1 and binding of the Sec1-like protein Vps45 by Rabenosyn-5 may promote the docking or "priming" stage of membrane fusion (61,154,174). In addition, the Rab1 effector p115 interacts with syntaxin-5 while the HOPS complex contains a sec1-like protein Vps33p and copurifies with Ypt7 and the t-SNARE Vam3p (175-179). Vps51p of the GARP complex also associates with the N-terminal domain of the t-SNARE Tlg1p (180,181). More information is necessary in order to understand the requirements of vesicle priming as well as potential differences between homotypic and heterotypic fusion with respect to docking and fusion events.

Combinatorial interactions generate specificity

It is evident that Rab GTPases and their effectors can act at many different stages in membrane trafficking. Fusagenic microdomains that amplify individual low affinity interactions allow positive recruitment of like components, shifting the membrane steady state to a particular function. Synthesis of phosphatidylinositol-(3)-phosphate (PtdIns(3)P) by the Rab5 effector hVps34 generates the lipid environment necessary for recruitment of FYVE domain containing proteins to endosomes (182-184). The recruitment of tethering factors like EEA1, which requires PtdIns(3)P for localization, contribute to the biogenesis of endosome membranes. Subsequent recruitment of SNARE proteins by EEA1 promotes the tethering and fusion of endosomal vesicles from the plasma membrane (154,162,163,174,185). The presence of a number of effectors with dual Rab binding capabilities may provide a mechanism for sorting cargo, such as

surface receptors, to distinct compartments. The Rab5 effectors Rabenosyn-5 and Rabaptin-5 also interact with Rab4 present on recycling endosomes, and Rab4, Rab5, and Rab11 positive membranes exist as a distinct but dynamic population (57,96,155). In addition, signaling may be halted by ubiquitin ligation of activated receptors and subsequent transport steps that promote receptor degradation (186-188).

It has become more evident that localization of activated receptors, such as receptor tyrosine kinases (RTKs), serine/threonine kinases (STKs) and G-protein coupled receptors (GPCRs), to the endosome can alter the relative strengths of different signaling pathways beyond receptor down-regulation by degradation (189). For example, a block in internalization of the EGFR to endosomes results in decreased Erk1 phosphorylation and increased phosphorylation of phospholipase Cγ (190). Redistribution of a number of signaling molecules, such as the EGF dependent Shc, Grb2 and mSOS, as well as the Ras-MAPK cascade components Raf, MEK and ERK suggests persistence of signaling from the endosomal compartment (191-194). Other signaling pathways are endosomal specific, including the nerve growth factor (NGF) activation of Rap1 on endosomes, which promotes differentiation of PC12 cells rather than proliferation. Also, endosomal localization of transforming growth factor-β (TGFβ) involving Smad anchor for receptor activation (SARA) is essential for activation of Smad proteins (195-197).

Endosomal targeting of SARA and the ubiquitin binding protein hepatocyte receptor tyrosine kinase substrate (Hrs), also necessary for receptor degradation, occurs through PtdIns(3)P binding to FYVE domains as previously discussed for EEA1 (187,198,199). In addition, β-arrestin2 interacts with clathrin to promote GPCR

internalization and serves as a scaffold for jun-kinase 3 (JNK3) and other upstream activators at the endosome (200-202). In these ways, membrane trafficking can amplify potential responses to signaling proteins by utilizing the spatial and temporal controls provided by compartmentalization. Signaling pathways can also influence the transport machinery, allowing feedback regulation of receptor trafficking (189,203).

Rab5: Master Regulator of Endocytosis

Analysis of Rab5 and its function in EGFR endocytosis has provided the most detailed example of Rab localization and function in signal transduction (11,51,152,189,203). Yeast genetics, *in vitro* reconstitution assays and *in vivo* localization studies, as well as the identification of more than 20 Rab5 specific effectors implicate Rab5 in each step of endocytosis from clathrin coated pit formation to enhancement of SNARE mediated fusion (57,204-206). Rab5 also defines a paradigm for Rab mediated nucleation of membrane subdomains through positive recruitment of cytoplasmic proteins, creating a steady state environment consistent with endosome function. Rab5 mediated interactions are also fundamental to organelle biogenesis, as Rab5 deletion results in reduction of endosomal membranes whereas overexpression of a constitutively active Rab5 mutant leads to an enlarged endosomal compartment (16,46,50).

Ligand dependent EGFR activation results in dimerization and tyrosine kinase activity that promotes cellular proliferation and differentiation; stimulation also triggers Rab5 activation and rapid receptor internalization (Figure 9) (207-209). Phosphorylation

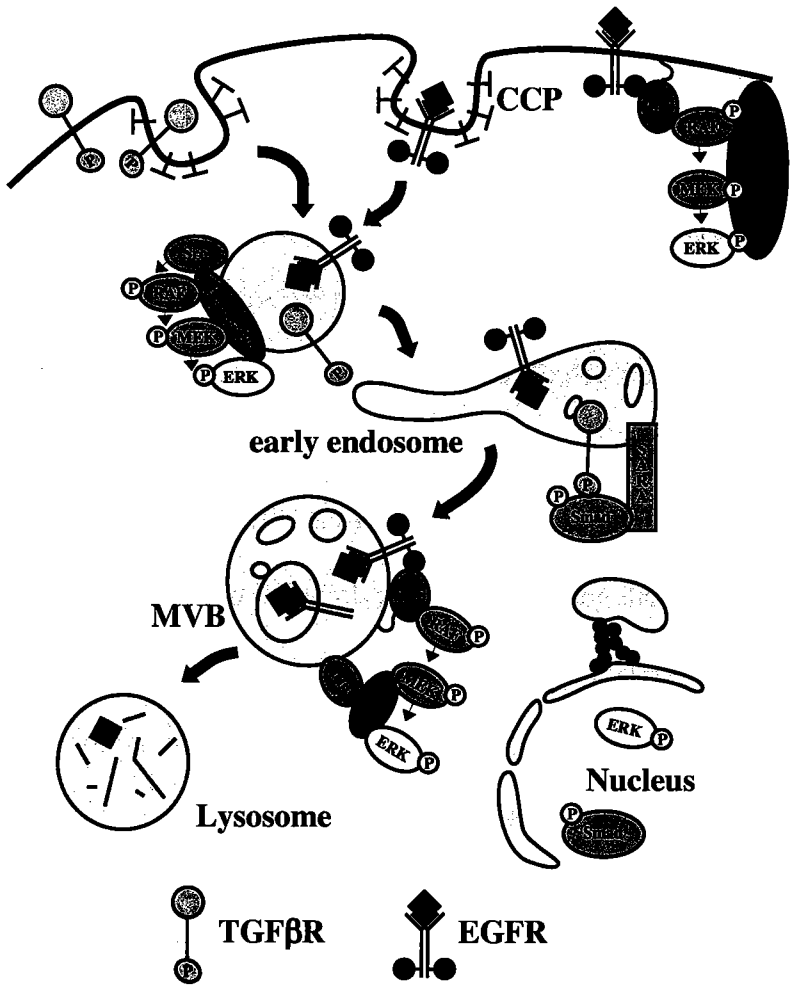


Figure 9. Early endosomal sorting and signaling. Different scaffold complexes reside at different subcellular compartments. This compartmentalization of signaling modules might be used to define a specific biological signal. In addition to the scaffold molecules, other signaling adaptors such as SARA have been found to localize on endosomes. Figure adapted from Teis et al., 2003 (189).

of the clathrin heavy chain, c-Cbl and Src as well as recruitment of Eps8, Eps15 and the AP2 complex allow subsequent engagement of clathrin coat components promoting formation of clathrin coated pits (210-215). Additional recruitment of the PI3K p100β/p85 and phospholipase Cγ results in increased PtdIns(3,4,5)P_3 and PtdIns(4,5)P_2 production and recruitment of additional phospholipid specific signaling and trafficking factors including dynamin, epsin, Grb2, and Akt (189,207,216,217). Internalization of EGF receptors to clathrin coated vesicles (CCVs) is increased by overexpression of Rab5; however, receptor signaling persists throughout endocytosis and endosome fusion (209,218). Downstream targets of Grb2, Src and Eps8, such as Ras and MAPK pathways, also respond to receptor internalization to direct trafficking through feedback loops (189). The Rab5 specific GEF activity of RIN1 is regulated by Ras whereas RN-tre, a Rab5 GAP, can be recruited to the plasma membrane by the EGF receptor substrate Eps8 (218,219). p38 MAPK has also been shown to phosphorylate RabGDI, accelerating extraction of Rab5 from membranes (220).

Rab5 promotes vesicle transport to the endosome likely through a minus end directed kinesin motor, and recruits the PI3K hVps34, increasing the levels of PtdIns(3)P on endosomal membranes (52,221). The specific interaction between PtdIns(3)P and the FYVE domain promotes recruitment of the Rab5 effectors EEA1 and Rabenosyn-5 in addition to other signaling proteins such as SARA (61,182,197,222,223). These factors stabilize endosomal membranes and promote fusion of CCVs with endosomes. EEA1 has a modular architecture with a N-terminal C_2H_2 Zn^{2+} finger, an extended heptad repeat, and a C-terminal region containing a calmodulin binding (IQ) motif, a Rab5

binding site, and a FYVE domain (182,185,224,225). In cell free reconstitution assays, EEA1 is essential for fusion of early endosomes and is thought to function as an endosomal tethering factor (162,174,226,227).

After delivery to endosomes, activated receptors are sorted to the late endosome/ multivesicular body (MVB) and ultimately targeted for degradation at the lysosome (189). Accumulation of PtdIns(3)P in endosomal membranes and ubiquitination of receptors by the E3 ligase c-Cbl promotes recruitment of Hrs through its FYVE domain and ubiquitin interaction motif (UIM), respectively (228-230). Subsequent binding of the ESCRT complex and TSG101 promote vesicle budding and sorting of ubiquinated receptors to the MVB (231-233). Inactive receptors and other factors targeted for the plasma membrane are postulated to return through transport steps involving Rab5 effectors like Rabenosyn-5 and Rapabtin-5 that also contain binding sites for Rab4, a Rab GTPase found on the recycling endosome. Rabaptin-5 forms a complex with the Rab5 GEF Rabex-5 and interacts with Golgi-associated γ-adaptin related ADP-ribosylation factor binding proteins (GGAs) suggesting that the Rabex-5/Rabaptin-5 complex may also function in vesicular transport from the TGN to endosomes. Consistent with this idea, EEA1 has also been shown to interact with the Golgi/endosome targeted Rab22. The wealth of information available with respect to Rab5 and its well-established role as a critical master regulator of endocytosis make the early endocytic machinery an excellent model system for understanding the structural mechanisms underlying Rab GTPase function.

Structural Mechanisms of Rab Mediated Vesicle Transport

As critical regulators of vesicular trafficking, Rab proteins comprise the largest GTPase family, with thirty-eight functionally distinct members and another twenty isoforms in the human genome. The specificity of Rab interactions with effectors and regulatory factors plays a central role with respect to the fidelity of membrane trafficking. The crystallographic studies of the Rab3A/Rabphilin-3A complex discussed above led to the hypothesis that the relatively conserved switch regions convey information regarding the state of the bound nucleotide whereas independent Rab CDRs determine the specificity of Rab/effector interactions (Figure 5B). Chapter II of this thesis describes the crystal structure of Rab5 bound to GppNHp and shows that the switch regions of Rab5C and Rab3A do not encode the same active conformation. In particular, the switch interface adopts a dramatically different conformation in Rab5 such that the resulting hydrophobic surface is non-complementary to the Rab3 effector Rabphilin-3A. Thus, structural plasticity within the relatively conserved switch regions is evolutionarily coupled to the conformational switching mechanism and represents a general determinant of Rab-effector recognition.

Activated Rab GTPases interact with effector proteins involved in vesicle formation, transport, tethering, docking and fusion. Using Rab5 mediated endocytic transport as a model system, I have gained insight into the structural basis for interaction of effectors and regulatory factors with Rab GTPases. The FYVE domain targets the Rab5 effector EEA1 to the endosome where it has been proposed to function as a tethering factor (154,162,163,174,185). In chapter III of this thesis, the highly organized

quaternary structure of the homodimeric C-terminal region of EEA1 provides insight into the structural basis underlying the targeting of EEA1 to early endosomes. In addition, characterization of truncation and site specific mutants of the EEA1 N-terminus demonstrates that the C_2H_2 Zn^{2+} finger of EEA1 is sufficient to mediate association with activated Rab5. These observations suggest that the structural organization of the EEA1 homodimer is ideally configured for multivalent membrane engagement, providing insight into the structural basis of endosome tethering.

In chapter IV of this thesis, structural and biochemical approaches were employed to examine the mechanism of Rab5 activation. A consensus GEF interaction epitope, derived from the crystal structures of four mammalian GEF-GTPase complexes, suggest some determinants of Rab GTPase recognition by GEFs may be general. A central 260 amino acid fragment of Rabex-5, containing a Rab5 GEF domain homologous to Vps9, was isolated and demonstrates robust GEF activity for Rab5C. The crystal structure of this fragment reveals a novel fold for the conserved Vps9 homology region, which corresponds to a sub-domain within the context of an integrated tandem domain architecture. Acidic, hydrophobic, and polar residues on a conserved surface of the Vps9 domain are shown to be critical for exchange activity. A quantitative family-wide analysis of Rab specificity demonstrates that the central catalytic fragment of Rabex-5 has selective exchange activity for Rab GTPases of the Rab5 sub-family. The conserved surface of the Vps9 domain has limited though potentially significant structural similarities with the Sec7 GEF domain of Gea1, which interacts with Arf GTPases through contacts with the switch regions (234-236).

CHAPTER II

STRUCTURAL PLASTICITY OF AN INVARIANT HYDROPHOBIC TRIAD IN THE SWITCH REGIONS OF RAB GTPASES IS A DETERMINANT OF EFFECTOR RECOGNITION

Summary

Rab GTPases function as regulatory components of an evolutionarily conserved machinery that mediates docking, priming and fusion of vesicles with intracellular membranes. Previously published work from the Lambright lab has shown that the active conformation of Rab3A is stabilized by a substantial hydrophobic interface between the putative conformational switch regions (18). A triad of invariant hydrophobic residues at this switch interface (Phe59, Trp76 and Tyr91) represents a major interaction determinant between the switch regions of Rab3A and the Rab3A specific effector Rabphilin-3A (90). Here, I report the crystal structure of the active form of Rab5C, a prototypical endocytic Rab GTPase. Like Rab3A, the active conformation of Rab5C is stabilized by a hydrophobic interface between the switch regions. However, the conformation of the invariant hydrophobic triad (residues Phe58, Trp75 and Tyr90 in Rab5C) is dramatically altered such that the resulting surface is non-complementary to the switch interaction epitope of Rabphilin-3A. This structural rearrangement reflects a set of non-conservative substitutions in the hydrophobic core between the central β sheet and the α2 helix. These

observations demonstrate that structural plasticity involving an invariant hydrophobic triad at the switch interface contributes to the mechanism by which effectors recognize distinct Rab subfamilies. Thus, the active conformation of the switch regions conveys information about the identity of a particular Rab GTPase as well as the state of the bound nucleotide.

This chapter contains work previously published in Merithew et al., 2001 (237). Initial Rab5C constructs were generated by Robin Heller-Harrison (Czech lab) and diffraction quality crystals were generated by Scott Hatherly. Extensive comparisons were made to the structure of Rab3A determined by John Dumas (18).

Introduction

As general regulators of intracellular vesicle transport between donor and acceptor membranes, Rab proteins comprise the largest GTPase family with eleven distinct homologues in yeast and 38 distinct Rab GTPases in mammals (5,12,13,166,238). Like other GTPases of the Ras superfamily, Rab proteins cycle between active (GTP bound) and inactive (GDP bound) conformations (5,17,87). A key question concerns the molecular and structural mechanisms by which Rab GTPases generate specificity for a diverse spectrum of effectors and regulatory factors. Biochemical and genetic studies of chimeric and mutant Rab proteins have identified several hypervariable regions, including the N/C-termini and the α3/β5 loop, that play an important role in determining functional specificity (69,92). However, interactions involving hypervariable regions can not explain the ability of RabGDI to recognize most or all Rab GTPases, yet still discriminate against other GTPase families, or the ability of certain regulatory factors to recognize particular Rab subfamilies (93-96). These observations imply the existence of specificity determinants that are common to all Rab GTPases, but not other GTPase families, as well as determinants that are conserved only within particular Rab subfamilies.

Crystallographic studies have identified structural motifs and modes of effector interaction that are distinct from those of other GTPase families. The active conformation is stabilized by additional hydrogen bonding interactions with the γ phosphate of GTP, mediated by serine residues in the P-loop and switch I region, as well as an extensive hydrophobic interface between the switch I and II regions (18,239). The

structure of the complex between a constitutively active mutant of Rab3A and a putative effector, Rabphilin-3A, revealed an interaction epitope that extends from the relatively conserved switch interface to an adjacent pocket formed by three hypervariable complementary determining regions (CDRs) corresponding to the N- and C-termini of the GTPase domain and the α3/β5 loop (90). A triad of invariant hydrophobic residues at the switch I/switch II interface of Rab3A mediates a central interaction with Rabphilin-3A. In a recent analysis of the primary structure of all known Rab GTPases, five distinctive motifs (termed Rab family motifs or RabFs) were identified that distinguish Rab proteins from other GTPase families (48). The RabF motifs are located either in the putative switch I and II regions or in the adjacent β3 and β4 strands. It was also noted that the Rab family could be further subdivided based on the presence of four subfamily motifs (RabSFs), which include all three Rab CDRs as well as a region corresponding to the α1 helix and most of the α1/β2 loop.

Rab5 is an essential regulator of early endosome fusion (50,240). Several putative Rab5 effectors have been identified including EEA1, a large protein consisting of an N-terminal Zn^{2+} finger, four long heptad repeats having weak homology with myosins, and a compact C-terminal region containing an IQ motif (putative calmodulin binding site), Rab5 interaction site and a FYVE domain that specifically binds PtdIns(3)P (162,182,185,225,241-243). Interactions with both PtdIns(3)P and Rab5 are essential for fusion of endocytic vesicles with early endosomes in an *in vitro* reconstitution assay (226,227). Although lacking the corresponding C-terminal regions, the FYVE domain of EEA1 is structurally similar to the ring finger domain of Rabphilin-3A (244). Moreover,

a predicted helical region N-terminal to the FYVE domain of EEA1 exhibits weak homology with the N-terminal helix of Rabphilin-3A and has been implicated in the interaction with Rab5 (162,245). However, critical hydrophobic residues in Rabphilin-3A are not conserved in EEA1. Thus, the extent to which the Rab3A/Rabphilin-3A complex can be regarded as a paradigm for other Rab-effectors complexes remains uncertain (246).

The crystallographic and biochemical studies cited above are consistent with the hypothesis that the relatively conserved switch regions convey information regarding the state of the bound nucleotide whereas independent Rab CDRs determine the specificity of Rab/effector interactions. Here, I report the crystal structure of Rab5 bound to GppNHp. Despite strong homology, the characteristic RabF motifs in the switch regions of Rab5 and Rab3A do not encode the same active conformation. In particular, the invariant triad of partially exposed hydrophobic residues located at the switch I/switch II interface adopts a dramatically different conformation in Rab5 such that the resulting hydrophobic surface is non-complementary to Rabphilin-3A. These changes reflect an alternative packing arrangement in response to a concerted set of non-conservative substitutions in the hydrophobic core. Moreover, the packing constraints on the conformation of the invariant hydrophobic triad appear to be engaged only in the active form. Thus, structural plasticity within the relatively conserved switch regions is evolutionarily coupled to the conformational switching mechanism and represents a potentially general determinant of Rab-effector recognition.

Experimental Procedures

Expression and Purification. The GTPase domain of Rab5C (residues 16-186) was expressed in *E. coli* using a modified pET15b vector containing an N-terminal 10xHis peptide (MGHHHHHHHHHHGS). BL21(DE3) cells harboring the modified pET15b plasmid containing the Rab5 insert were grown at 37°C in 2xYT media, induced at an OD_{600} of ~0.6 by addition of 1 mM IPTG and harvested after 3 hours at 37°C. The cell pellet was resuspended in 50 mM Tris, pH 8.5, 0.1% 2-mercaptoethanol, lysed by sonication, centrifuged at 35,000xg for 1 hour and the supernatant loaded onto a NiNTA-agarose column (Qiagen). After washing with ten column volumes of 50 mM Tris, pH 8.5, 500 mM NaCl, 10 mM imidazole, 0.1% 2-mercaptoethanol, the fusion protein was eluted with a gradient of 10-500 mM imidazole. The protein was further purified by ion exchange chromatography on Resource Q and Resource S (Pharmacia) followed by gel filtration on Superdex-75 (Pharmacia). Roughly 30 mg of >99% pure protein were obtained from a 6 L culture.

Crystallization and Data Collection. The GppNHp bound form of Rab5C was prepared as described for Rab3A (18). Crystals of the Rab5-GppNHp complex were grown at 4°C by vapor diffusion in hanging drops containing 10% PEG-6000, 50 mM MES, pH 6.0, 0.2 M NaCl, 0.5 mM $MgCl_2$ and 0.1% 2-mercaptoethanol. Single crystals appeared overnight and grew to maximum dimensions of 0.3 x 0.3 x 0.2 after several days. The crystals are in the primitive orthorhombic space group $P2_12_12_1$ with cell constants a = 35.9 Å, b = 64.0 Å and c = 65.9 Å. The volume of the unit cell is consistent with one molecule in the asymmetric unit and a solvent content of 35%. Crystals were

soaked for five minutes at 4°C in a cryoprotectant-stabilizer solution (30% PEG-6000 and 20% glycerol) prior to flash freezing in a nitrogen cryostream. A native data set complete to 1.8 Å was collected on a Rigaku RUH3R/Mar 30 cm image plate detector equipped with focusing mirrors (Charles Supper). The crystal was maintained at 100 °K using a nitrogen cryostream (Oxford Cryosystems). All data were processed with Denzo and scaled with Scalepack (247).

Structure Determination and Refinement. The structure of GppNHp-bound Rab5C was solved by molecular replacement using a polyalanine search model derived from the coordinates of the Rab3A structure (18). A translation search including the top 50 rotation function solutions yielded a unique solution with an R-value of 50.5% after rigid body refinement against data from 8 - 3 Å. Initial difference maps calculated with Sigma A weights to reduce model bias revealed poor density for the putative switch I and II regions and several loop regions. Multiple rounds of simulated annealing and manual fitting were interleaved with gradual extension of the resolution to 1.8 Å. Manual fitting of the omitted regions and placement of the side chains was followed by additional rounds of simulated annealing, positional and B factor refinement. The final refined model includes residues 19 to 182, one molecule of GppNHp, a Mg^{2+} ion, 153 ordered water molecules and has a crystallographic R value of 0.198 and a free R value of 0.246 based on a 5% subset of reflections randomly omitted prior to refinement. Molecular replacement was conducted with AMORE as implemented in CCP4, refinement with X-PLOR and interactive model building with O (248-251). Structural images were

generated with MOLSCRIPT or GRASP, combined with GL_RENDER (Dr. L. Esser) and rendered with RASTER 3D (252-255).

Other crystallographic models. For comparison with Rab5C, models for Rab3A, Ypt51, Rab6 and the Rab3A/Rabphilin-3A complex were rendered from the coordinates of the corresponding crystal structures (PDB ID codes 3RAB, 1EK0, 1D5C, and 1ZBD, respectively).

Results

Overall Structure and Comparison with other Rab GTPases. The GTPase domain of Rab5C bound to GppNHp was crystallized and the structure determined by molecular replacement (Table 3). The final refined model, which includes residues 19-182, one molecule of GppNHp, a Mg^{2+} ion and 153 ordered water molecules, has a working R value of 0.198 and free R value of 0.246. The stereochemistry is excellent and there are no backbone torsion angles outside allowed regions of the Ramachandran plot. Residues 16-18 at the N-terminus, residues 65-69 in the β2/β3 loop and residues 183-186 at the C-terminus are poorly ordered.

Like other members of the GTPase superfamily, Rab5C posses a characteristic nucleotide binding fold consisting of a six-stranded β sheet surrounded by five α helices (Figure 10). In the absence of a structure for the GDP-bound form of Rab5C, a precise definition of the conformational switch regions is not possible. However, a comparison of the GppNHp-bound structure of Rab3A with that of the GDP-bound form of Rab6 from the malaria parasite Plasmodium *falciparum* suggests that the conformational changes accompanying nucleotide exchange will be localized to regions analogous to those in $p21^{ras}$ (256). As a working hypothesis, it is assumed that the conformational changes in Rab5 are localized to the corresponding regions, although the precise nature and extent of the conformational changes will not alter the observations described below. To facilitate discussion, Rab CDRs will be referenced as defined for Rab3A based on the complex with Rabphilin-3A (90). Whether these and/or other variable regions mediate effector interactions with Rab5C remains to be determined.

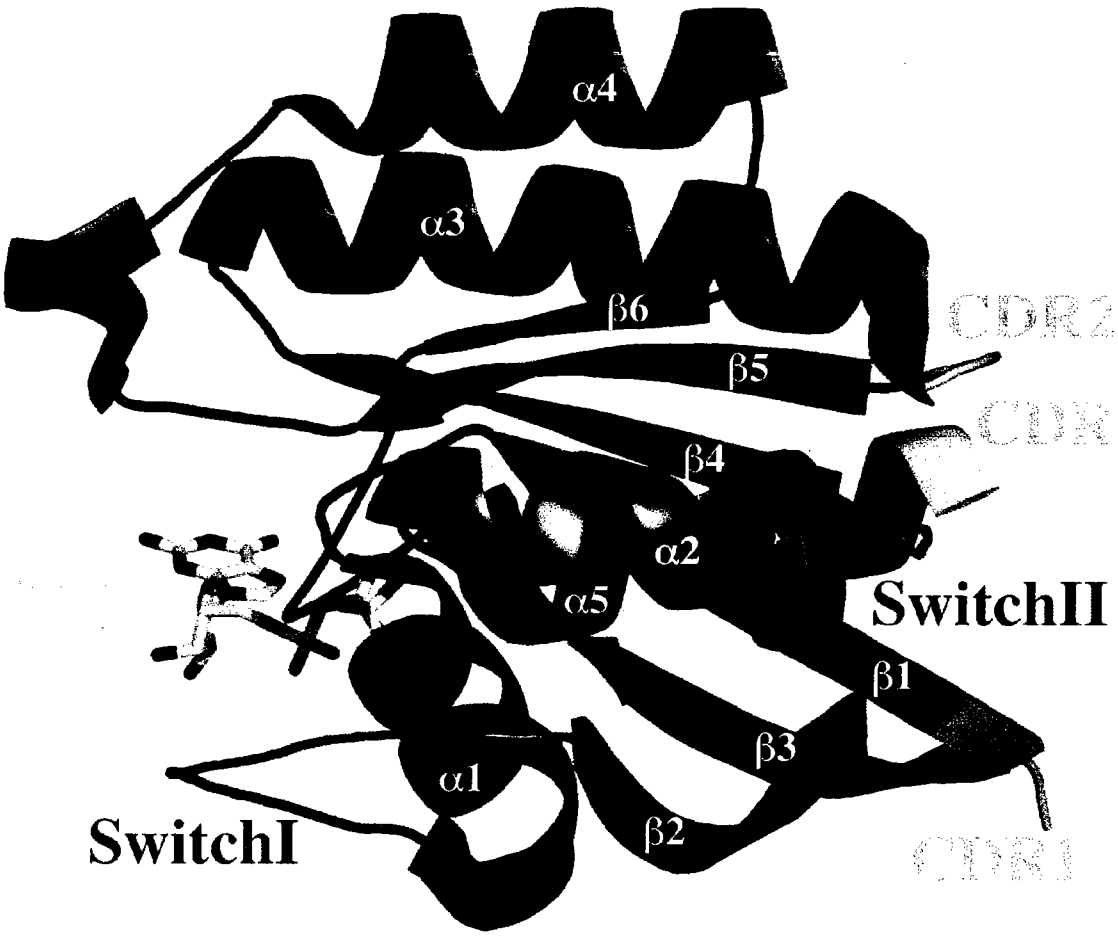


Figure 10. Overall structure of Rab5C. Ribbon representation of the Rab5C structure with the putative switch regions highlighted in green, the regions corresponding to the CDRs of Rab3A in light blue and the nucleotide in yellow.

Table 3. Rab5C: Structure Determination and Refinement

Data Collection[a]				
Resolution (Å)		20 - 1.8		
R_{sym}[b]		0.054 (0.227)		
$\langle I/\sigma \rangle$		20.3 (5.4)		
		4		
Completeness (%)		98.8 (94.9)		
Molecular Replacement				
Euler Angles (Q1, Q2, Q3)		120.6, 14.3, 347.4		
Fractional Coordinates (x, y, z)		0.101, 0.026, 0.476		
		Highest Peak	Highest False Peak	
Correlation Coefficient		32.9	17.9	
Refinement				
RMS Deviations				
Resolution (Å)	R factor	R_{free}[d]	Bond Length	Bond Angle
8 - 1.8	0.195	0.232	0.006 Å	1.2°

[a]Values in parentheses represent the highest resolution shell.

[b]$R_{sym} = \Sigma_h \Sigma_j |I_j(h) - \langle I(h) \rangle| / \Sigma_h \Sigma_j I_j(h)$.

[c]Values after rigid body refinement from 8.0 - 3.0 Å.

[d]R-value for a 5% subset of reflections selected at random and omitted from refinement.

Although the overall fold of Rab5C resembles that of Rab3A, significant structural differences are evident throughout the GTPase domain (Figure 11). The structures are most similar in the regions in contact with the nucleotide whereas the largest differences occur in or adjacent to regions implicated in the interaction with effectors and/or regulatory factors. The GTPase domains of Rab5C and Rab3A share 37% sequence identity overall; however, the homology within the switch regions is considerably higher. For example, the switch II regions of Rab5C and Rab3A are 63% identical and substitutions are either conservative or occur at exposed positions. Moreover, the majority of intramolecular interactions with the switch II region involve residues from the highly conserved RabF motifs. Interestingly, the magnitude of the displacement of Cα atoms does not strictly correlate with sequence variability. Indeed, the displacements in the relatively conserved switch II region are considerably larger than the overall RMS deviation and comparable to the displacements in the poorly conserved α4 helix. Moreover, the largest main chain rearrangement within the switch II region is centered on an invariant tyrosine residue in one of the highly conserved RabF motifs. The origin and functional consequences of these changes with respect to effector recognition are considered below.

Structural Variability in the Rab CDRs. Although the large structural differences in the variable Rab CDRs is due in part to inherent disorder in the absence of interactions with effectors, critical changes can be attributed to specific sequence determinants. For example, proline residues from a PXXXP motif in Rab5 result in an α5 helix that is truncated by two turns relative to Rab3A. The additional turns of the α5 helix in Rab3A

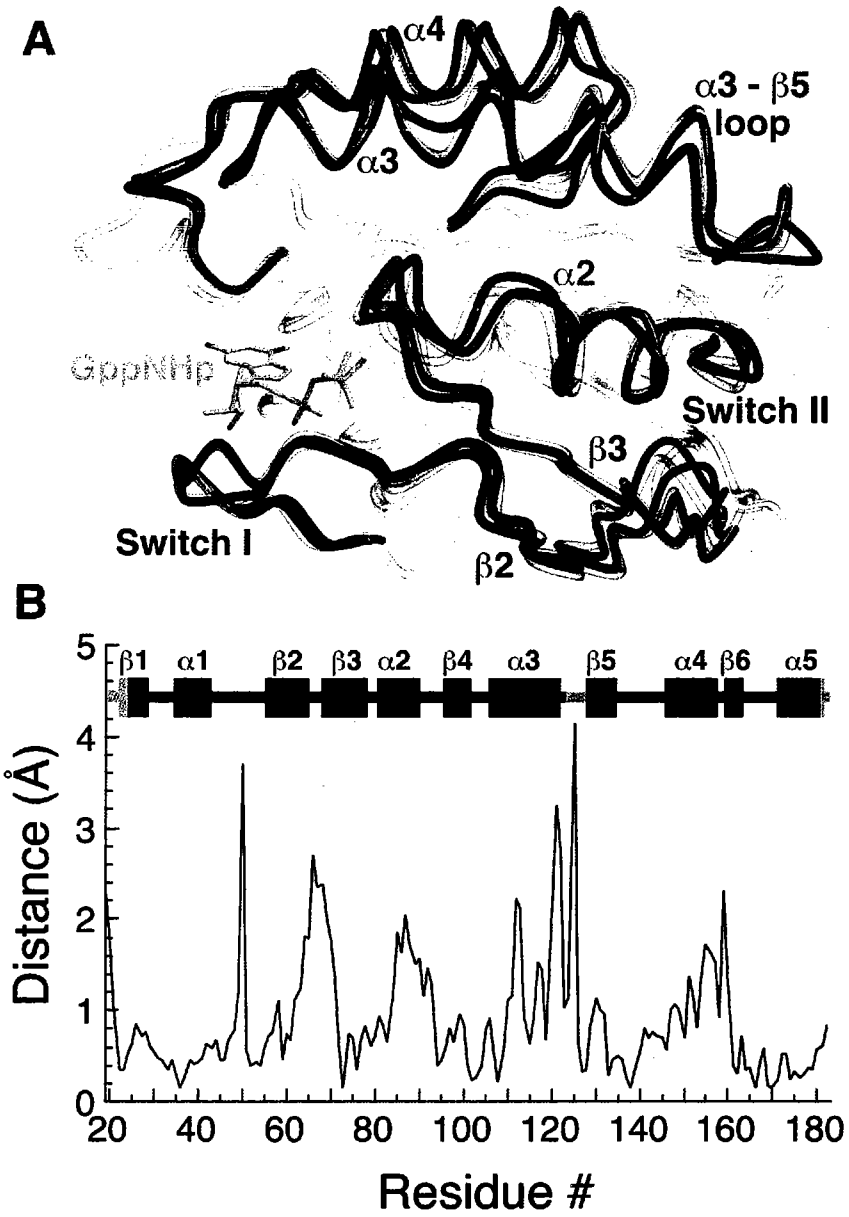


Figure 11. Comparison of the active conformations of endocytic and exocytic Rab GTPases. **(A)** Superposition of the Rab5C structure (blue) with the structures of Rab3A alone (purple), the yeast Rab5 homologue Ypt51 (semi-transparent blue) and Rab3A from the complex with Rabphilin-3A (semi-transparent purple). The superposition is based on Cα atoms. Models for Rab3A alone (18), Ypt51 (239) and Rab3A from the Rab3A/Rabphilin-3A complex (90) were derived from the coordinates of the corresponding crystal structures. **(B)** Distance between pairs of Cα atoms in the Rab5C and Rab3A structure following superposition. Also shown is the secondary structure, with the putative switch regions highlighted in magenta and the CDRs of Rab3A in orange.

comprise an essential element of the binding pocket for a C-terminal hydrophobic motif in Rabphilin-3A (90). The differences in backbone conformation are further augmented by non-conservative substitutions in which hydrophobic residues in Rab3A are substituted for charged or polar residues in Rab5. Thus, the specificity of effector interactions with the Rab CDRs is determined by backbone conformational constraints as well as stereochemical and electrostatic complementarity of the relevant molecular surfaces. These specificity determinants are strongly correlated with a high degree of variability in the sequences comprising the Rab CDRs.

Structural Rearrangement of an Invariant Hydrophobic Triad at the Switch Region Interface. As shown in Figures 12A and B, a triad of invariant hydrophobic residues encoded by three of the characteristic RabF motifs undergoes a dramatic structural rearrangement. In Rab5, the aromatic ring of Tyr91 in the switch II region is buried in the hydrophobic core formed between the switch regions and the central β sheet while the aromatic ring of Phe58 in the β2 strand is partially buried, leaving the indole ring of Trp75 in the β3 strand significantly exposed. In this packing arrangement, the aromatic rings of Trp75 and Tyr91 lie in Van der Waals contact with the methyl group of Ala 56 in the switch I region. Consequently, the conformation adopted by the invariant hydrophobic triad in Rab5 requires a residue with a small side chain at position 56. In Rab3A, a non-conservative alanine to isoleucine substitution at the corresponding position displaces the aromatic side chains of Tyr91 and Trp76. The displacement of Trp76 in turn forces the side chain of Phe59 to adopt an alternative rotomer conformation. Thus, the aromatic side chains of Tyr91 and Phe59 in Rab3A protrude

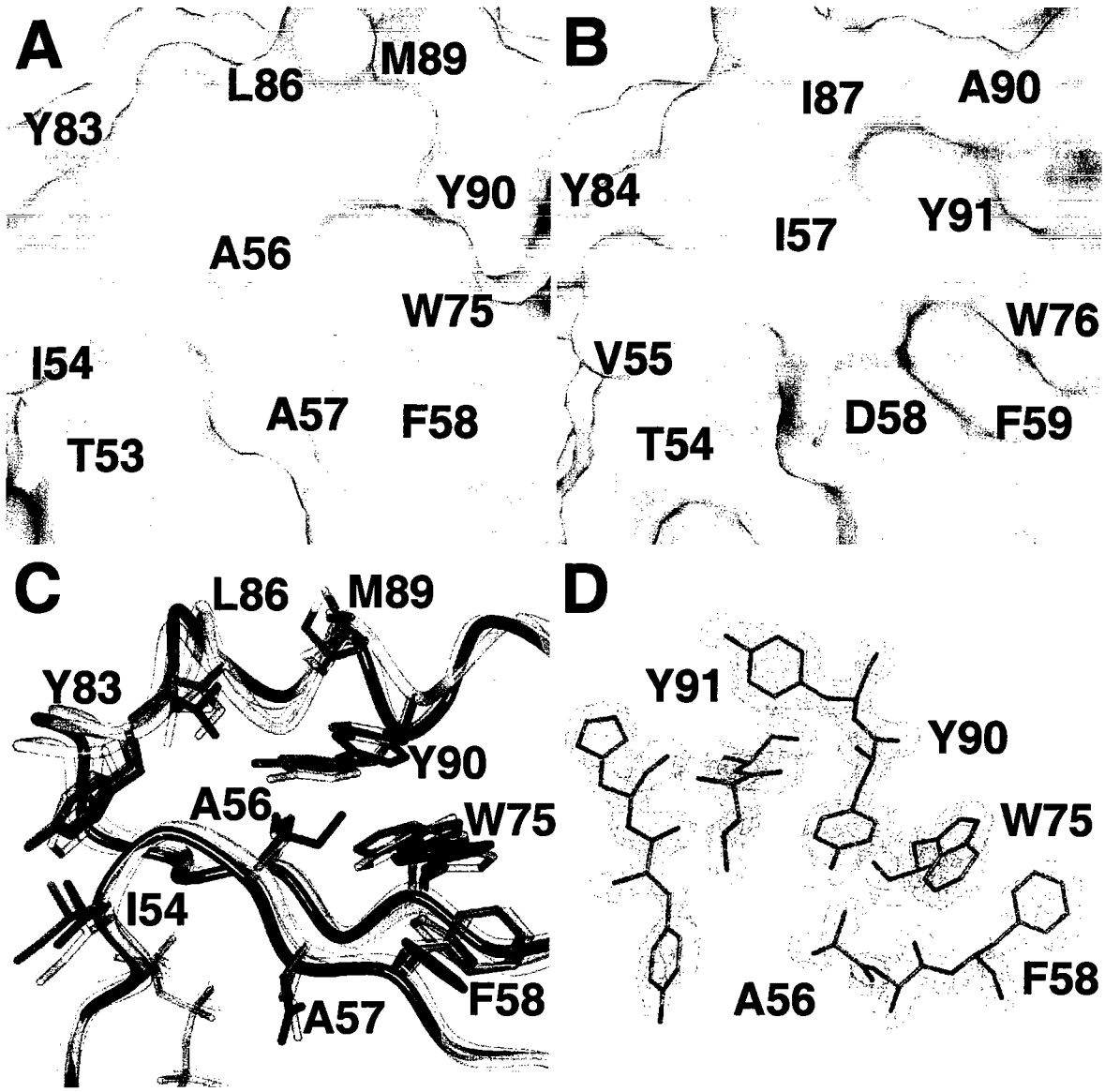


Figure 12. Structural rearrangement of an invariant hydrophobic triad at the switch region interface. **(A)** Semi-transparent molecular surface of Rab5C in the vicinity of the invariant hydrophobic triad (Phe58, Trp75 and Tyr90). **(B)** Semi-transparent molecular surface of Rab3A (18) in the vicinity of the invariant hydrophobic triad (Phe59 Trp76 and Tyr91). **(C)** Overlay following superposition of C α atoms in the independently determined active structures of Rab5C (blue and orange), Ypt51 (semi-transparent blue and orange), Rab3A alone (gray and purple) and Rab3A from the complex with Rabphilin-3A (semi-transparent gray and purple). Models for Rab3A alone (18), Ypt51 (239) and Rab3A from the Rab3A Rabphilin-3A complex (90) were derived from the coordinates of the corresponding crystal structures. **(D)** σA weighted $2F_O$-F_C map countered at 1.2σ showing residues in the vicinity of the invariant hydrophobic triad.

from either side of the indole ring of Trp76, which is considerably more buried than its counterpart in Rab5. Despite the dramatic conformational rearrangement of the invariant hydrophobic triad, the alternative packing arrangement in Rab5 preserves the hydrophobic interface between the switch regions.

Due to crystal contacts involving the hydrophobic surfaces at the switch interface, it is reasonable to question whether the observed active conformations might be influenced by crystal packing or by engagement with effectors. Several lines of evidence indicate that the structural plasticity of the invariant hydrophobic triad is not a crystallographic artifact but rather an inherent property of Rab GTPases. First, the hydrophobic triad in the independently determined structure of the yeast Rab5 homologue Ypt51 (239) adopts a nearly identical conformation despite crystallizing in a different space group (Figure 12C). This observation eliminates crystal packing as a likely explanation. Second, the conformation of the invariant hydrophobic triad in Rab3A is not altered by the interaction with Rabphilin-3A (90), indicating that these residues are pre-oriented (Figure 12C). Consistent with this observation, the side chains of the invariant triad are well ordered due to intramolecular packing constraints (Figure 12D). Finally, as described below, a concerted set of non-conservative substitutions in the hydrophobic core between the switch regions and the central β sheet provides a simple and plausible explanation for the observed structural rearrangement.

Sequence Determinants of the Active Conformation. As illustrated in Figure 13, exocytic Rab GTPases conserve either valine or isoleucine at the equivalent of position 57 in Rab3A whereas a subset of endocytic Rab GTPases (Rab5, Rab7, Rab20, Rab21,

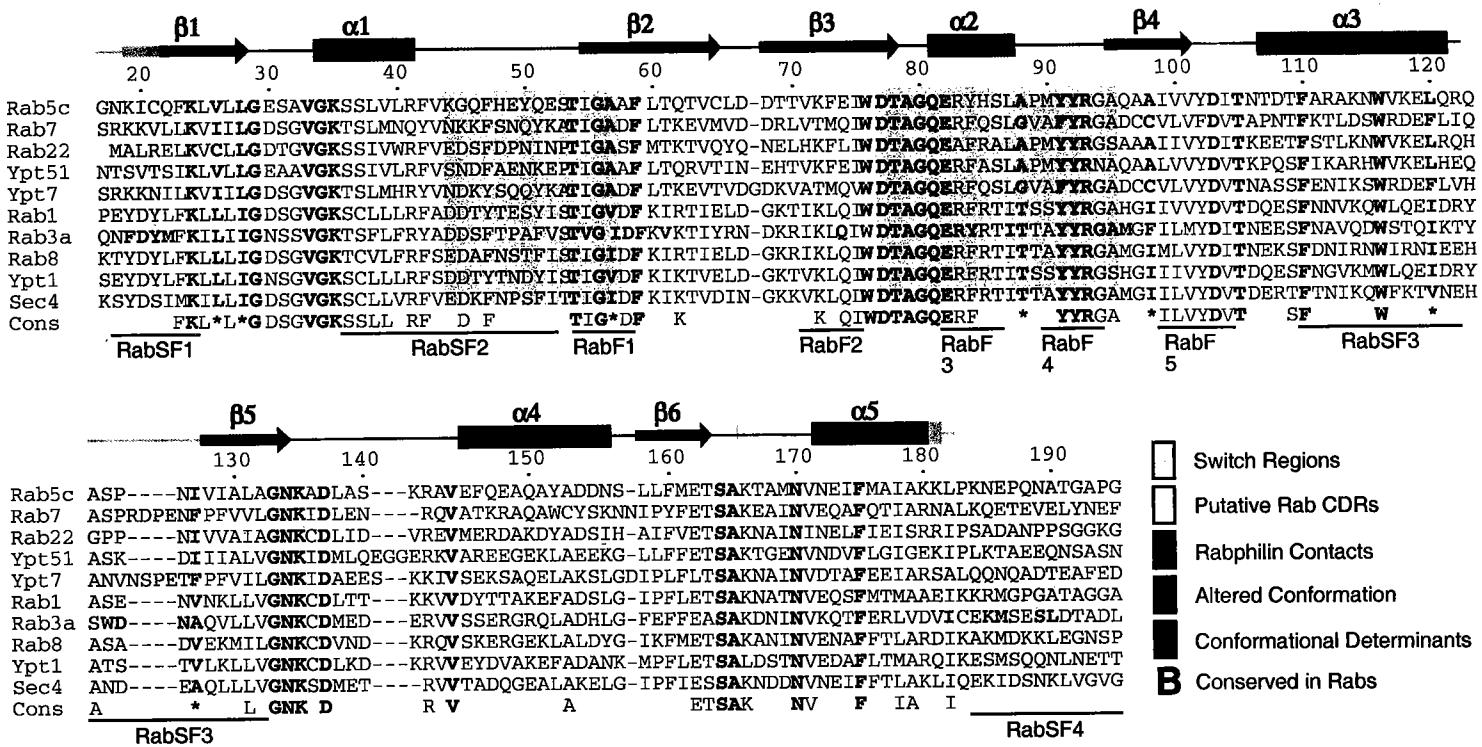


Figure 13. Structure based sequence alignment of representative endocytic and exocytic Rab GTPases. Invariant hydrophobic residues with significantly altered conformations are shown in purple and the corresponding conformation determining residues in green. Rab3A residues in the interface with Rabphilin-3A are shown in blue (90). Residues that are highly conserved in the Rab GTPase family are indicated in bold face type. Residue numbering and secondary structure corresponds to Rab5C. The complementary determining regions of Rab3A are highlighted in orange. The conformational switch regions defined by comparison of GppNHp-bound Rab3A and GDP-bound Rab6 are highlighted in purple. Horizontal lines below the sequences denote Rab family (RabF) and Rab sub-family (RabSF) regions (48).

Rab22 and Rab24) substitute residues with small side chains (alanine or glycine). In yeast, the distinction is absolute: all exocytic Rab homologues conserve an isoleucine or valine and all endocytic Rab homologues conserve an alanine. Although the alanine to isoleucine substitution in the switch I region represents an obvious proximal determinant of the observed differences, other non-conservative substitutions could contribute to stabilization of the alternative conformation.

In order to determine whether the conformation of the invariant hydrophobic triad is likely to vary from one Rab to another or simply reflects two general arrangements, I have systematically analyzed each position in or adjacent to the switch regions. Residues were segregated into three groups based on the sequences of all known Rab GTPases and the available structural data for the active conformations of Rab3A and Rab5. The first group includes invariant residues as well as residues that are conservatively substituted. Although not determinants of the active conformation, it is nevertheless clear that structural rearrangements involving conserved residues (e.g. the invariant hydrophobic triad) can mediate and amplify the effects of adjacent non-conservative substitutions. The second group consists of non-conservative substitutions of residues that occupy solvent exposed positions. Such substitutions could well contribute to the traditional mechanism of specificity determination; however, the effects are not expected to propagate to beyond the immediate vicinity. The third group corresponds to non-conservative substitutions of residues that are substantially buried in either the Rab3A or Rab5 structures. The effects of such substitutions will necessarily propagate to surrounding regions and thus have the potential to be key conformational determinants of

specificity. A surprising number of non-conservative substitutions occur at interior positions. As shown in Figure 14, the majority of such substitutions involve a cluster of hydrophobic residues situated in the hydrophobic core formed between the switch regions, the β1-β4 strands and the α3 helix. Substitutions at these positions are concerted and, to a first approximation, can be correlated with the overall similarity in the GTPase domains of the previously identified Rab subfamilies (48).

Structural Plasticity in the Switch Regions is a Determinant of Effector Recognition. As shown in Figure 15, the invariant hydrophobic triad represents a major site of interaction between Rabphilin-3A and the switch regions of Rab3A (90). The protruding aromatic side chains of Phe59 and Tyr91 engage complementary hydrophobic surfaces in Rabphilin-3A. Due to the structural rearrangement described above, the corresponding surface in Rab5 is non-complementary to the switch interaction epitope of Rabphilin-3A. These observations clearly demonstrate that structural plasticity involving a partially exposed invariant hydrophobic triad contributes to the specificity of Rab-effector interactions. Thus, in addition to signaling the state of the bound nucleotide, the active conformation of the switch regions conveys information about the subfamily of a particular Rab GTPase.

A General Model for Activation of Rab GTPases. The recent crystal structure of a GDP bound Rab6 homologue provides the opportunity to assess whether the structural changes accompanying activation are likely to influence the conformation of the invariant hydrophobic triad (256). Interestingly, the invariant hydrophobic triad in GDP-bound Rab6 adopts a conformation distinct from that of the active forms of either Rab3A or

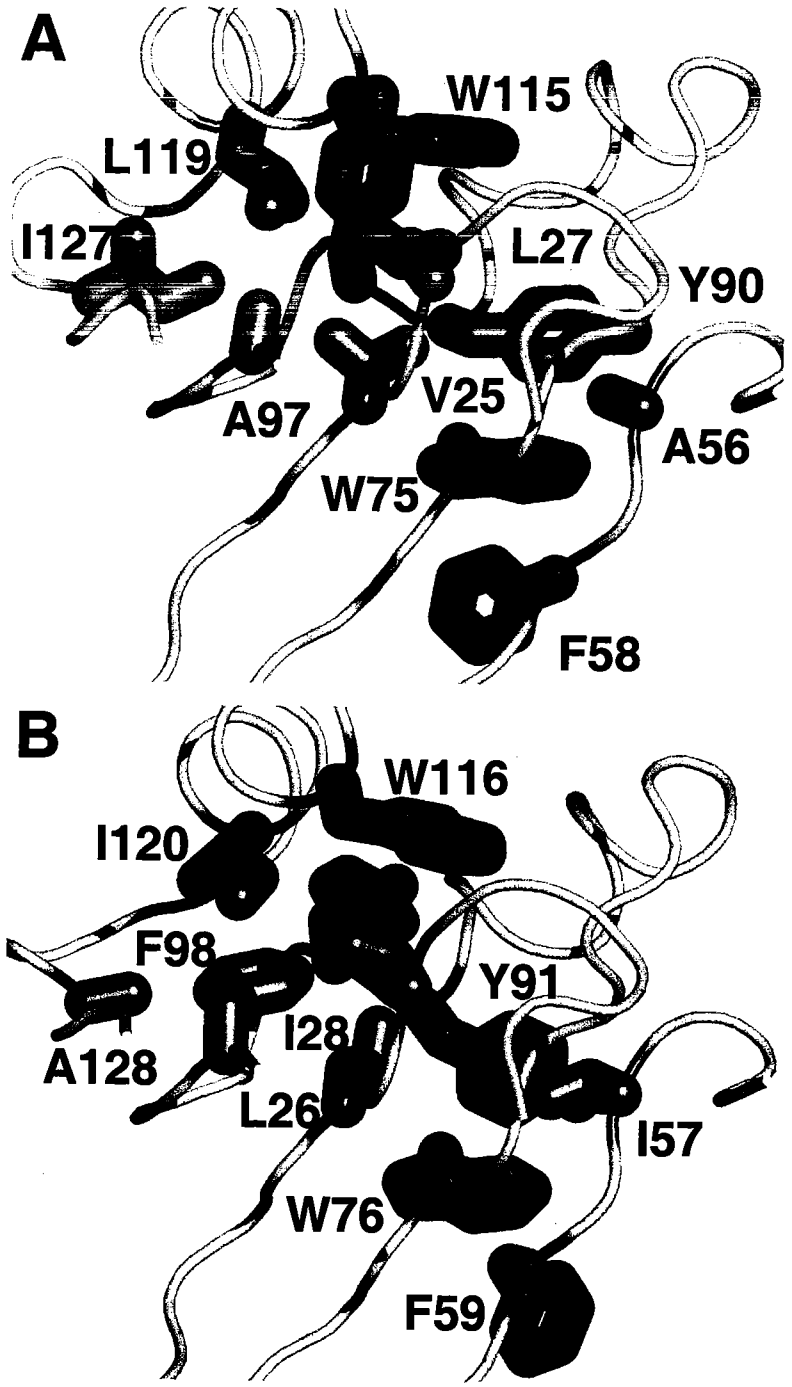


Figure 14. Location of non-conservative substitutions predicted to influence the active conformation of invariant residues in the switch regions. (A) Rab5. **(B)** Rab3A from the GppNHp-bound structure (18). Variable residues are highlighted in green and invariant residues in purple. The view is from the interior of the protein.

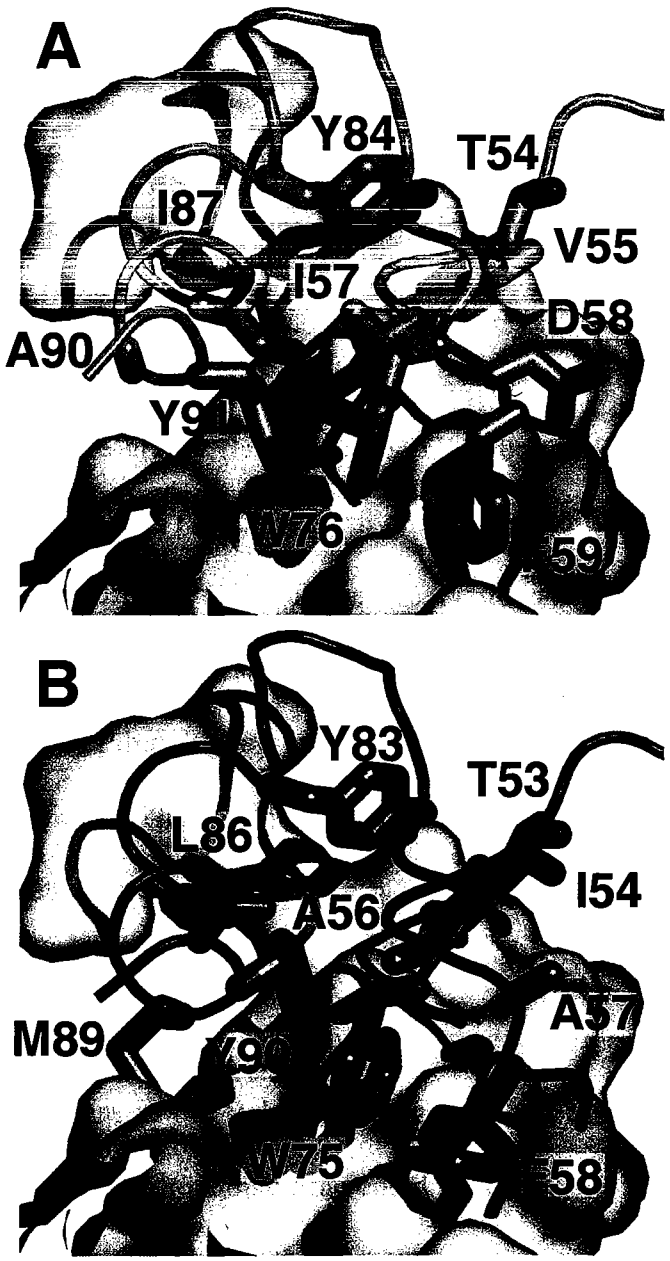


Figure 15. Complementarity of the invariant hydrophobic triad in Rab3A and Rab5C with the switch region interaction epitope of Rabphilin-3A. (A) Molecular surface of Rabphilin-3A in contact with the switch regions of Rab3A (orange and purple). **(B)** Hypothetical Rab5C/Rabphilin-3A interface following superposition of Rab5C (blue and green) with Rab3A. Note that the alternative conformation of the invariant hydrophobic triad in Rab5C lacks complementarity with Rabphilin-3A. The Rabphilin-3A and Rab3A models in this figure were generated from the coordinates of the Rab3A/Raphilin-3A structure (90).

Rab5 (Figure 16). In particular, the conformation determining isoleucine in the switch I region and invariant tyrosine in the switch II region are displaced relative to their counterparts in Rab3A and Rab5. Although the main chain atoms of the invariant tryptophan and phenylalanine are not substantially displaced, the aromatic side chains of these residues adopt different rotomer conformations. Moreover, crystallographic and NMR studies of a large number of GTP binding proteins support the general conclusion that the switch regions are highly flexible in the GDP bound form but adopt specific ordered conformations in the GTP bound form due to the additional hydrogen bonding interactions with the γ phosphate (17). These observations suggest a simple structural model for activation of Rab GTPases, leading to distinct active conformations for each Rab subfamily. Upon GTP binding, an extensive hydrophobic interface forms between the switch I and II regions. The formation of this interface engages the aromatic side chains of the invariant hydrophobic triad in a Rab subfamily specific conformation determined by Van der Waals interactions with the critical conformation determining residues in the hydrophobic core. Finally, the presence of a substantial exposed hydrophobic patch adjacent to the Rab CDRs suggests that the nucleotide dependent engagement of the invariant hydrophobic triad could play a general role in Rab activation and effector recognition.

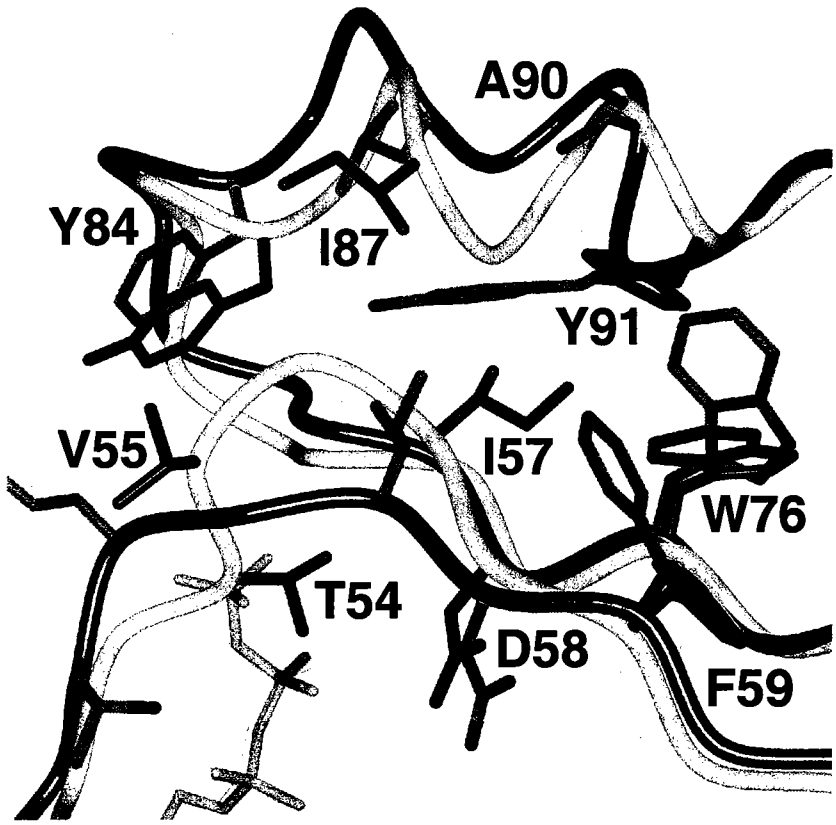


Figure 16. Distinct conformations for the invariant hydrophobic triad in the active and inactive forms of Rab GTPases. Overlay following superposition of Cα atoms in GppNHp-bound Rab3A (blue and purple) and GDP-bound Rab6 (gray and orange). The Rab3A and Rab6 models in this figure were generated from the coordinates of the GppNHp-bound Rab3A (18) and GDP-bound Rab6 structures (256).

Discussion

Given the large number of Rab GTPases, it is plausible that the structural plasticity of the invariant hydrophobic triad evolved primarily as a mechanism to augment other specificity determining interactions with the Rab CDRs. In addition to the obvious implications with respect to effector interactions, differences in the active conformation of the invariant hydrophobic triad could potentially contribute to the specificity of interactions with Rab GAPs (123-125,257-259). On the other hand, comparison of the GppNHp-bound Rab5C and Rab3A structures with the GDP-bound Rab6 structure suggests that the constraints on the conformation of the invariant hydrophobic triad are engaged only in the active form (18,256). Consequently, the conformation of the invariant hydrophobic triad is unlikely to affect the specificity of interactions with GDI, GEFs, GDFs or other factors that interact with the GDP-bound form (33,93,100,104,105,120,260-264). Indeed, one or more residues from the invariant hydrophobic triad could well contribute to the interaction with Rab GDI, which recognizes the GDP-bound conformation of all Rab GTPases (93).

The use of invariant hydrophobic residues in the switch regions as nucleotide dependent recognition determinants appears to be unique to the Rab family. For example, invariant hydrophobic residues in the switch II region of heterotrimeric G protein α subunits engage hydrophobic residues in the α3 helix with little or no difference in conformation in the crystal structures of the active forms of $G_{t\alpha}$, $G_{i\alpha}$ and $G_{s\alpha}$ (265-267). Interestingly; however, the specificity of the $G_{s\alpha}$ interaction with adenylyl cyclase is primarily due to a conformational rearrangement in the α3/β5 loop

(compared with $G_{i\alpha}$ which does not interact with the same site on adenylyl cyclase) rather than differences in the composition of residues in the effector interaction epitope (267). Thus, in both Rab GTPases and G_{α} subunits, conformational plasticity of invariant residues provides a structural mechanism that propagates and amplifies the effects of key non-conservative amino acid substitutions. It is likely that similar structural mechanisms of specificity determination have evolved in other large protein or domain families.

When compared with other Rab family GTPases, the Rab5 structure provides compelling evidence for an evolutionarily conserved mechanism of effector recognition in which the structural plasticity of an invariant hydrophobic triad at the switch region interface is coupled to non-conservative substitutions in the hydrophobic core. The conformation of the invariant hydrophobic triad reflects the state of the bound nucleotide as well as Rab subfamily identity and thus appears to be an integral component of the switching mechanism. Although the extent to which Rab effectors other than Rabphilin-3A will interact with the invariant hydrophobic triad remains to be established experimentally, the presence of a substantial nucleotide dependent hydrophobic surface adjacent to the Rab CDRs suggests the possibility of a general role with respect to activation and effector recognition.

CHAPTER III

MULTIVALENT ENDOSOME TARGETING BY HOMODIMERIC EEA1 AND DETERMINANTS OF RAB5 INTERACTION

Summary

EEA1 localization to early endosomes is mediated by a C-terminal region that includes a calmodulin binding motif, a Rab5 interaction site and a FYVE domain, which selectively binds phosphatidylinositol 3-phosphate. The crystal structure of the C-terminal region bound to inositol 1,3-bisphosphate reveals an organized quaternary assembly consisting of a parallel coiled-coil and a dyad symmetric FYVE domain homodimer. Structural and biochemical observations support a multivalent mechanism for endosomal localization in which dimerization and quaternary structure amplify the weak affinity and modest specificity of head group interactions with conserved residues. A unique mode of membrane engagement deduced from the quaternary structure of the C-terminal region provides insight into the structural basis of endosome tethering.

To gain further insight into the structural determinants for endosome tethering, I have also characterized the interaction of Rab5C with truncation and site specific mutants of the N-terminus of EEA1 using quantitative binding measurements. The results demonstrate that the C_2H_2 Zn^{2+} finger is both essential and sufficient for the N-terminal interaction with Rab5. Although the heptad repeat C-terminal to the C_2H_2 Zn^{2+} finger

provides the driving force for stable homodimerization, it does not influence either the affinity or stoichiometry of Rab5 binding. Hydrophobic residues predicted to cluster on a common face of the C_2H_2 Zn^{2+} finger play a critical role in the interaction with Rab5. Although the homologous C_2H_2 Zn^{2+} finger of the Rab5 effector Rabenosyn-5 binds to Rab5 with comparable affinity, the analogous C_2H_2 Zn^{2+} finger of the yeast homologue Vac1p shows no detectable interaction with Rab5, reflecting nonconservative substitutions of critical residues. Large changes in the intrinsic tryptophan fluorescence of Rab5 accompany binding to the C_2H_2 Zn^{2+} finger of EEA1. These observations can be explained by a mode of interaction in which a partially exposed tryptophan residue located at the interface between the switch I and II regions of Rab5 lies within a hydrophobic interface with a cluster of non-polar residues in the C_2H_2 Zn^{2+} finger of EEA1.

This chapter contains work previously published in Dumas et al. (268), Lawe et al. (269), and Merithew et al. (270). C-terminal EEA1 constructs and diffraction quality crystals were generated by John Dumas. Data collection, refinement and analysis of the EEA1 C-terminal structure was conducted by John Dumas, David Lambright and myself. The FYVE domain lipid binding experiments were done by David Lambright. Craig Stone assisted with collection of surface plasmon resonance and intrinsic tryptophan fluorescence data as well as purification of EEA1 point mutants. GST fusion constructs for the N-terminal C_2H_2 Zn^{2+} fingers of Rabenosyn-5 and Vac1p were provided by Sudharshan Eathiraj (270). EEA1 C-terminal mutations that prevent Rab5 binding were generated and analyzed by Deidre Lawe (Corvera Lab) (269).

Introduction

As master regulators of membrane trafficking, Rab GTPases cycle between active (GTP bound) and inactive (GDP bound) conformations (4,11,59). In the active conformation, Rab GTPases interact with diverse effectors implicated in vesicle budding, cargo sorting, motor-dependent transport, tethering, docking, and fusion. GEFs, GAPs, and other accessory factors, including RabGDI, provide multiple points of regulation throughout the GTPase cycle by modulating nucleotide binding, GTP hydrolysis, and membrane association (2). Protein kinases and phosphatases have also been implicated in the regulation of Rab function, either directly or by phosphorylation of effectors and regulatory factors (271-274). Thus, through regulated interactions with effectors, Rab GTPases couple signal transduction networks to the membrane trafficking machinery.

Both fluid phase and receptor mediated endocytosis depend on activation of Rab5, which plays a critical role in clathrin-coated vesicle formation, endosome motility, and early endosome fusion (275). Activated Rab5 interacts with diverse effectors, including scaffolding proteins and tethering factors, and further influences signaling and trafficking events by recruitment of class I and III PI3Ks to endosomes (61,120,162,241,243). The class III PI3K, hVPS34, selectively generates PtdIns(3)P which binds to FYVE (Fab1, YOTB/ZK632.12, Vac1, EEA1) and PX (phagocyte oxidase homology) domains in modular signaling and trafficking proteins (182,224,225,276-281).

The Rab5 effector EEA1 was identified as a Lupus autoantigen that localizes to early endosomes (282). EEA1 has a modular architecture with an N-terminal C_2H_2 Zn^{2+} finger, four consecutive heptad repeats, and a C-terminal region containing a calmodulin

binding (IQ) motif, a Rab5 binding site, and a FYVE domain that binds specifically to phosphatidylinositol 3-phosphate, PtdIns(3)P (182,185,224,225). In cell free reconstitution assays, EEA1 is essential for fusion of early endosomes (162,174,226,227). Endosomal localization requires an intact FYVE domain and is sensitive to inhibitors of PI3K activity as well as mutants of conserved residues in the FYVE domain that disrupt PtdIns(3)P binding (185,224,283,284). Although essential, the FYVE domain is not sufficient for endosome targeting, which requires an additional region of ~40 amino acids N-terminal to the FYVE domain (185,245). This region is also essential for Rab5 binding, suggesting that localization depends on dual interactions with Rab5 and PtdIns(3)P (162,245). Finally, Rab5 also binds to an independent site at the N-terminus of EEA1, which has recently been shown to bind Rab22 as well (162,285).

The C_2H_2 Zn^{2+} finger of EEA1 shares significant homology with the C_2H_2 Zn^{2+} finger in the Rab5 effector Rabenosyn-5 as well as the corresponding C_2H_2 Zn^{2+} finger in Vac1p, an effector of the yeast Rab5 homologue Ypt51. Like EEA1, Rabenosyn-5 and Vac1p contain a FYVE domain involved in endosome targeting (61). Temperature sensitive mutants implicate Vac1p in intervacuolar trafficking and vacuolar protein sorting (286,287). Immunodepletion of Rabenosyn-5 blocks homotypic early endosome fusion as well as heterotypic fusion of endocytic vesicles with early endosomes, suggesting that Rabenosyn-5 plays a critical role distinct from that of EEA1 (61). The GTP-bound forms of Rab4 and Rab5 bind to sites in the central and C-terminal regions of Rabenosyn-5, respectively (155). It is not known whether Rab5 binds directly to the C_2H_2 Zn^{2+} finger of EEA1, Rabenosyn-5, or Vac1p. Two hybrid data indicate that the

integrity of the C_2H_2 Zn^{2+} finger is essential for Rab5 binding to the N-terminus of EEA1 (162,288). The requirement for an intact C_2H_2 Zn^{2+} finger may reflect a direct interaction with Rab5 or an indirect structural role. For example, the double Zn^{2+} finger of Rabphilin-3A is essential for interaction with Rab3A; however, in the crystal structure of the Rab3A/Rabphilin-3A complex, the double Zn^{2+} finger does not contact Rab3A but instead supports interactions with flanking regions (90).

To determine the structural basis for domain organization, dimerization and quaternary structure with respect to EEA1 localization and endosome tethering, we have characterized the binding of soluble phosphoinositides to monomeric and homodimeric constructs of EEA1 and determined the crystal structure of the homodimeric C-terminal region as a complex with the head group of PtdIns(3)P. Reflecting a highly organized quaternary structure, the EEA1 homodimer is ideally configured for multivalent membrane engagement. The simplest thermodynamic model for bivalent recognition of PtdIns(3)P in a lipid bilayer quantitatively accounts for the large amplification of the weak affinity and moderate specificity of soluble PtdIns(3)P binding to the EEA1 FYVE domain and explains why the region preceding the FYVE domain is required for localization to early endosomes. The mode of multivalent membrane engagement deduced from the organized quaternary structure of the EEA1 C-terminal region provides insight into the structural basis of endosome tethering.

To gain further insight into the structural basis underlying the function of the EEA1 N-terminus in the tethering and fusion of early endosomes and endocytic vesicles, I have characterized the interaction of Rab5C with truncation and site specific mutants of

EEA1 using quantitative binding measurements. The results demonstrate that the C_2H_2 Zn^{2+} finger is sufficient for the N-terminal interaction with Rab5C and support a mode of interaction in which an invariant tryptophan residue, which is partially exposed at the interface between the switch I and II regions of Rab5, lies in or near an interface that involves a cluster of hydrophobic residues in the C_2H_2 Zn^{2+} finger.

Experimental Procedures

Constructs, expression, and purification. EEA1 and Rab5C constructs were amplified with Vent polymerase (New England Biolabs). EEA1 constructs were sub-cloned into a modified pET15b vector containing an N-terminal 6xHis tag (MGHHHHHHGS). Rab5C constructs were sub-cloned into pGEX-4T1 (Amersham Biosciences) for expression as an N-terminal GST fusion. Site specific mutants were generated using the Quick Change Site-Directed Mutagenesis Kit (Stratagene). All constructs and mutants were verified by sequencing the entire coding region from both 5' and 3' directions. BL21(DE3) or BL21(DE3)-RIL cells (Stratagene) were transformed with the pGEX-4T1/Rab5C or modified pET15b/EEA1 plasmids, grown in 2xYT-amp (16 g Bacto tryptone, 10 g Bacto yeast extract, 5 g sodium chloride, and 100 mg ampicillin per liter) at 25°C ($EEA1_{36-91}$) or 37°C (Rab5C and EEA1 constructs) to an OD_{600} of 0.6, and induced with 1 mM isopropyl-1-thio-β-D-glactopyranoside (IPTG) for three hours.

For purification of wild type and mutant proteins, cells were suspended in lysis buffer (50 mM Tris, pH 8.0, 0.1 M NaCl, .1% mercaptoethanol, 0.1 mM PMSF, 1 mg/mL lysozyme) and disrupted by sonication. Triton X100 was added to a final concentration of 0.5% and the cell lysates centrifuged at 35000xg for 40 min. For 6xHis fusion proteins, clarified supernatants were loaded onto an NiNTA-agarose column (Qiagen). After washing with ten column volumes of buffer (50 mM Tris, pH 8.0, 500 mM NaCl, 10 mM imidazole, .1 % mercaptoethanol), 6xHis fusion proteins were eluted with a gradient of 10-150 mM imidazole. For GST fusions, the supernatants were loaded onto a

glutathione-sepharose column (Amersham Biosciences) equilibrated with 50 mM Tris, pH 8.0, 0.1 M NaCl, 0.1% 2-mercaptoethanol. After washing with 10 column volumes of the same buffer, GST fusion proteins were eluted with 10 mM reduced glutathione. Subsequent ion exchange chromatography using Source Q or Source S (Amersham Biosciences) followed by gel filtration chromatography over Superdex-75 (Amersham Biosciences) resulted in preparations that were >99% pure as judged by SDS-PAGE. To generate the untagged form of Rab5C constructs, GST fusion proteins at a concentration of 2-4 mg/mL were incubated with 2 μg/mL of human α-thrombin (Hematalogic Technologies) overnight at 4°C in 50 mM Tris, pH 8.0, 2 mM $CaCl_2$, and .1% mercaptoethanol. Following incubation with glutathione-agarose to remove residual fusion protein, the cleaved Rab5C constructs were further purified by ion exchange and gel filtration chromatography. Typical yields of purified proteins range from 10 to 100 mg/L of bacterial culture. For Rab GTPases, all buffers are supplemented with 2 mM $MgCl_2$.

Ligand Binding Experiments EEA1 constructs at concentrations of 1-5 μM in 20 mM Hepes, pH 7.5, 100 mM NaCl were titrated with aliquots of soluble head group analogs. Samples were excited at 290 nm (2 nm bandpass) and emission spectra recorded from 300-400 nm (1 nm bandpass) using an ISS spectrofluorimeter. The magnitude of fluorescence quenching (ΔI) was calculated from $\Delta I = I - I_o$, where I is the integrated emission spectrum at a given free ligand concentration ([L]) and I_o is the integrated emission spectrum without ligand. Values for the dissociation constant (K_d) were

obtained by a non-linear least square fit to a simple two state binding model $\Delta I = C*[L] / (K_d + [L])$, where C is a constant.

Sedimentation equilibrium. EEA1 constructs were dialyzed against 50 mM Tris, pH 7.5, 100-150 mM NaCl and centrifuged to equilibrium in a Beckman Optima XLI analytical ultracentrifuge. The absorbance at 230 nm (A_{230}) or 280 nm (A_{280}) was measured as a function of the radial distance (r) from the axis of rotation. The x-values of the data were transformed as $\sigma_m*(r_o^2-r^2/2)$, where r_o was taken as the last point in each data set and σ_m was calculated with SEDINTERP (289) using the monomer molecular mass for each construct. Data were fit to the function $A_{280}(r) = C_0 + \Sigma_i C_i*\exp(-n_i*\sigma_m *(r_o^2-r^2)/2)$, where C_0 and C_i are constants and n_i represents the order of the ith oligomeric species.

Crystallization and Structure Determination Crystals were grown at 4°C in microseeded hanging drops containing 12 mg/ml $EEA1_{1287-1411}$, 11% PEG 4000, 50 mM Hepes, pH 7.0, 60 mM ammonium acetate, and 1.5 mM $Ins(1,3)P_2$. The crystals are in the space group $P2_12_12_1$ with unit cell dimensions a = 36.5 Å, b = 85.1 Å, and c = 88.0 Å. Prior to data collection, crystals were briefly transferred to 10% PEG 8000, 10% MPD, and 10% glycerol and flash frozen in a nitrogen cryostream. The structure was solved by multiwavelength anomalous diffraction at the Zn^{2+} edge (Table 4). Data were collected at the X12C beam line at the Brookhaven National Synchrotron Light Source using an inverse beam strategy. Throughout data collection, the crystal was maintained at 100°K. Four Zn^{2+} sites were identified by Patterson and direct methods (SHELXS)

based on the Bijvoet differences in data collected at the f'' maximum. The heavy atom model was refined against a maximum likelihood target function using SHARP (290). Initial phases were improved by solvent flipping using Solomon as implemented in CCP4 (291). A σA weighted Fourier summation yielded a readily interpretable map with continuous density throughout both polypeptide chains and unambiguous density for two Ins(1,3)P_2 molecules. An initial model was constructed using O (292) and refined without NCS restraints against data from 50 - 2.2Å using X-plor (250). Structural figures were generated with Molscript (252), GL_Render (L. Esser) and Raster3D (253).

Co-precipitation. GST-$Rab5_{18-185}$ was exchanged at 37°C for 30 minutes with a 25-fold molar excess of GppNHp in 50mM Tris, pH 8.5 containing 5mM EDTA, 100mM NaCl and 2U of agarose-immobilized alkaline phosphatase per mg of protein. The exchange reaction was quenched by addition of 10 mM $MgCl_2$ and excess nucleotide removed by gel filtration on Superdex-75. 6xHis-EEA1 constructs were incubated in a 1:1 molar ratio with GDP-bound or GppNHp-bound GST-$Rab5_{18-185}$ at a concentration of 20μM for 30 minutes at 4°C in buffer A (50mM Tris, pH 8.5, 100mM NaCl, 2mM $MgCl_2$, 0.1 mg/mL BSA, and 0.1% Tween 20). 50 μl of equilibrated glutathione-sepharose beads (Amersham Biosciences) were added to 100 μl of the protein mixture and incubated for 1h. Following centrifugation, the supernatant was collected and the pellet washed three times with 100 μl of buffer A. After washing, the beads were incubated with buffer A containing 10 mM glutathione for 15 minutes and the fractions analyzed by SDS-PAGE with Coomassie Blue staining.

Surface plasmon resonance. SPR sensograms were collected with a BiacoreX instrument (Pharmacia Biosensor AB) using a carboxy-methylated (CM5) sensor chip to which a GST antibody was covalently coupled using reagents and protocols supplied by the manufacturer. All proteins were dialyzed into flow buffer (10 mM Tris pH 7.5, 150 mM NaCl, 2 mM $MgCl_2$, 0.005% Tween-20) prior to injection. Tandem flow cells were utilized, one loaded with the 500 nM GST-Rab5C (sample channel) and the other with an equivalent molar quantity of GST (reference channel) expressed and purified as described above for the GST-Rab5 constructs. GST and GST-Rab5C were injected at a flow rate of 5 μl/min whereas subsequent injections were conducted at a flow rate of 20 μl/min. Conversion to the active conformation was achieved by injecting 50μl of 3 μM Rabex-5 followed immediately by a 10 μl injection of 200 nM GppNHp. Binding and dissociation were monitored following 20 μl injections of increasing concentrations of 6xHis-EEA1. Following curve alignment, the reference sensogram, which reflects bulk refractive index changes and/or reversible non-specific binding, was subtracted from the sample sensogram. The SPR signal at equilibrium (R_{eq}) was extracted from the fit with a simple 1:1 Langmuir binding model and plotted as a function of 6xHis-EEA1 concentration. Dissociation constants (K_d) were obtained from a fit to the hyperbolic binding function $R_{eq} = R_{max}$ [6xHis EEA1] / (K_d + [6xHis EEA1]), where R_{max} corresponds to the SPR signal at saturation and is treated as an adjustable parameter. Mean values and standard deviations (σ_{n-1}) were calculated from 2-4 independent measurements. Control experiments verify that the fitted K_d values are independent of flow rate (5-50 μl/min)

and surface coverage (10 fold range), indicating that the equilibrium data are not limited by mass transfer or rebinding.

Intrinsic tryptophan fluorescence. Rab5C at concentrations of 1 or 20 μM in 10 mM Tris pH 7.5, 150 mM NaCl, 2 mM $MgCl_2$ was titrated with 6xHis-$EEA1_{36-91}$, 6xHis-$EEA1_{36-218}$, or the W104A mutant of 6xHis-$EEA1_{36-218}$. Samples were excited at 290 nm (1 μM Rab5C) or 300 nm (20 μM Rab5C) with a 2 nm bandpass and emission spectra recorded from 300-400 nm (1 nm bandpass) using an ISS spectrofluorimeter. The magnitude of fluorescence quenching (ΔI) at 340 nm was calculated as $\Delta I = I - I_o$, where I and I_o are the emission intensities in the presence and absence of EEA1, respectively. Values for the dissociation constant (K_d) and the number of binding sites (n) were obtained by a non-linear least squares fit to a simple two state binding model $\Delta I = \Delta I_{max} (b - \{b^2 - 4 [EEA1]_t / (n [Rab5C]_t)\}^{1/2})/2$, where $b = 1 + [EEA1]_t/(n [Rab5C]_t) + K_d/(n [Rab5C]_t)$, $[EEA1]_t$ and $[Rab5C]_t$ are the total concentrations of EEA1 and Rab5C, respectively, and ΔI_{max} corresponds to the emission intensity at saturation. Concentrations were determined from the absorbance at 280 nm using calculated extinction coefficients ε_{280} (M^{-1} cm^{-1}) = #Trp x 5200 + #Tyr x 1200 + #Cys x 120.

Homology modeling. The sequence of the EEA1 C_2H_2 Zn^{2+} finger was threaded against a protein structure database using the 3D-PSSM fold recognition server to identify a suitable structure for homology modeling. The NMR structure of a C_2H_2 Zn^{2+} finger from the yeast transcription factor Adr1 (PDB ID 1paa) was selected for further homology modeling on the basis of it's low scoring E-value and the absence of gaps in

the alignment. The EEA1 C_2H_2 Zn^{2+} finger shares 29% identity with that of Adr1, which represents the closest homologue of known structure. Non-conserved residues were substituted with the corresponding residues in EEA1, which were modeled in the most frequently observed rotomer conformation compatible with the structure. The resulting homology model represents a rough, working approximation to the actual structure, with an overall fold consistent with the common topology of C_2H_2 Zn^{2+} fingers.

Results

The C-terminus of EEA1

A screen of homodimeric EEA1 constructs reconstituted with a 2 fold excess of $Ins(1,3)P_2$ yielded crystals for $EEA1_{1287\text{-}1411}$, which begins at the IQ motif and extends through the C-terminus. The structure was solved by multiwavelength anomalous diffraction at the Zn^{2+} edge (Table 4). Two polypeptide chains and two molecules of $Ins(1,3)P_2$ are contained within the asymmetric unit. The high quality of the experimental electron density map, obtained without symmetry averaging, allowed each chain to be separately fit and refined, giving rise to two independent observations. The refined model, which includes two polypeptide chains (residues 1289-1411), two molecules of $Ins(1,3)P_2$, four Zn^{2+} ions and 45 water molecules, has a R-value of 0.221 and a free R-value of 0.281.

The C-terminal region of EEA1 forms a highly organized quaternary assembly consisting of a parallel coiled-coil that terminates abruptly in a dyad symmetric FYVE domain homodimer (Figure 17). Two molecules of $Ins(1,3)P_2$ are bound to identical sites located on a common surface directly opposite the coiled-coil. Flanking each phosphoinositide binding site, a short 'turret loop' (residues ^{1366}SVTV^{1369}) and a pair of tandem lysine residues (^{1396}KK^{1397}) protrude from the extended surface of the FYVE domain homodimer. The EEA1 FYVE domain is comprised of four short β strands and a C-terminal α helix assembled around two Zn^{2+} ions, which coordinate the thiol groups of cysteine residues from the four conserved CxxC motifs. The first ten residues of each

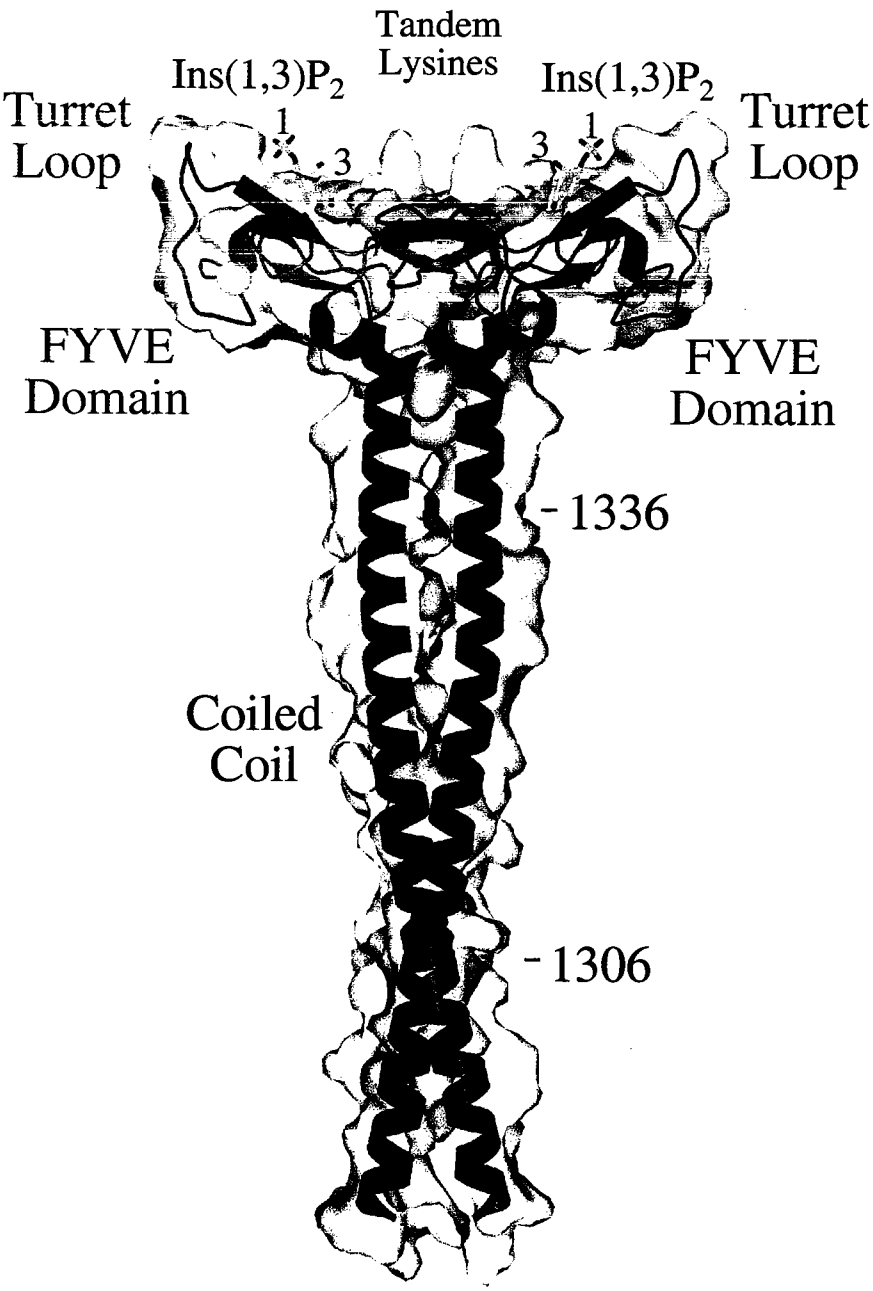


Figure 17. Overall structure of the homodimeric EEA1 C-terminal region bound to Ins(1,3)P2. The polypeptide chains are colored green (chain A) or blue (chain B), the head group is shown in yellow (carbon and phosphorus atoms) and red (oxygen atoms), and the Zn^{2+} ions are depicted as gray spheres. Side view with semi-transparent molecular surfaces covering the individual polypeptide chains

Table 4. EEA1 C-terminus: structure determination and refinement

Data Collection[a]

Wavelength (Å)	λ1 (1.28322)	λ2 (1.28356)	λ3 (1.23985)
Source	NSLS X12C	NSLS X12C	NSLS X12C
Resolution (Å)	20-2.2	20-2.2	20-2.2
R_{sym} (%)[b]	5.0	4.9	4.5
$\langle I/\sigma \rangle$	24.7 (2.9)	24.9 (2.8)	27.7 (3.4)
Completeness	86.4 (33.5)	86.8 (40.4)	91.7 (58.0)
Redundancy	4	4	4

Phasing Power and Figure of Merit[a]

	Bijvoet λ1	Bijvoet λ2	Bijvoet λ3	λ1 vs. λ3	λ2 vs. λ3
Aentric PP	3.07 (0.79)	3.67 (0.93)	3.57 (0.98)	5.05 (1.49)	4.08 (1.45)
Centric PP				3.14 (0.85)	2.66 (1.02)

FOM	Acentric	0.66 (0.36)	Centric	0.60 (0.27)

Refinement

			RMS Deviations	
Resolution	R Factor	Free R Factor	Bond Length	Bond Angle
20-2.35 Å	22.1	28.1	0.006	1.1

Table 4 footnotes

[a]Values in parentheses represent the highest resolution shell.

[b]$R_{sym} = \Sigma_h \Sigma_j |I_j(h) - \langle I(h) \rangle| / \Sigma_h \Sigma_j I_j(h)$.

[c]Values are $\langle (\Delta|F|^2) \rangle^{1/2} / \langle |F|^2 \rangle^{1/2}$ for data from 20 to 2.2 Å.

[d]R-value for a 5% subset of reflections selected at random and omitted from refinement.

FYVE domain extend around the outer edge of the phosphoinositide binding site as they link the C-termini of the coiled-coil helices to the FYVE domain core. In both polypeptide chains, the N-terminal linker region is observed in a similar, well ordered conformation constrained both by intramolecular interactions with the FYVE domain core and by extensive intermolecular interactions between polypeptide chains. Residues ^{1351}EDNEV^{1355} in the linker segment adopt a stable α helical secondary structure such that the side chain of the second residue in the helix (Asp1352) faces the phosphoinositide binding site.

Competition assays and NMR experiments suggest that the affinity of the EEA1 FYVE domain for the head group of PtdIns(3)P lies in the micromolar range (284,293). However, the dissociation constant for Ins(1,3)P_2 or other phosphoinositide head groups has not been measured directly. Consequently, the intrinsic affinity and specificity of the FYVE domain for the head group of PtdIns(3)P is not known. In view of the organized quaternary structure of the EEA1 C-terminal region (Figure 18), it is likewise unclear whether the free energy for simultaneous head group binding to an EEA1 homodimer reflects an additive, cooperative or anti-cooperative mechanism. To address these questions, we have determined the head group binding properties and oligomeric state of three C-terminal EEA1 constructs that contain different lengths of the coiled-coil region.

The C-terminal region of EEA1 contains a single conserved tryptophan at the N-terminus of the FYVE domain. When titrated with phosphoinositide head groups or soluble dibutyl PtdIns(3)P, the intrinsic tryptophan fluorescence of $EEA1_{1278\text{-}1411}$ is quenched and accompanied by a ~5 nm blue shift in the emission maximum (Figure 19A).

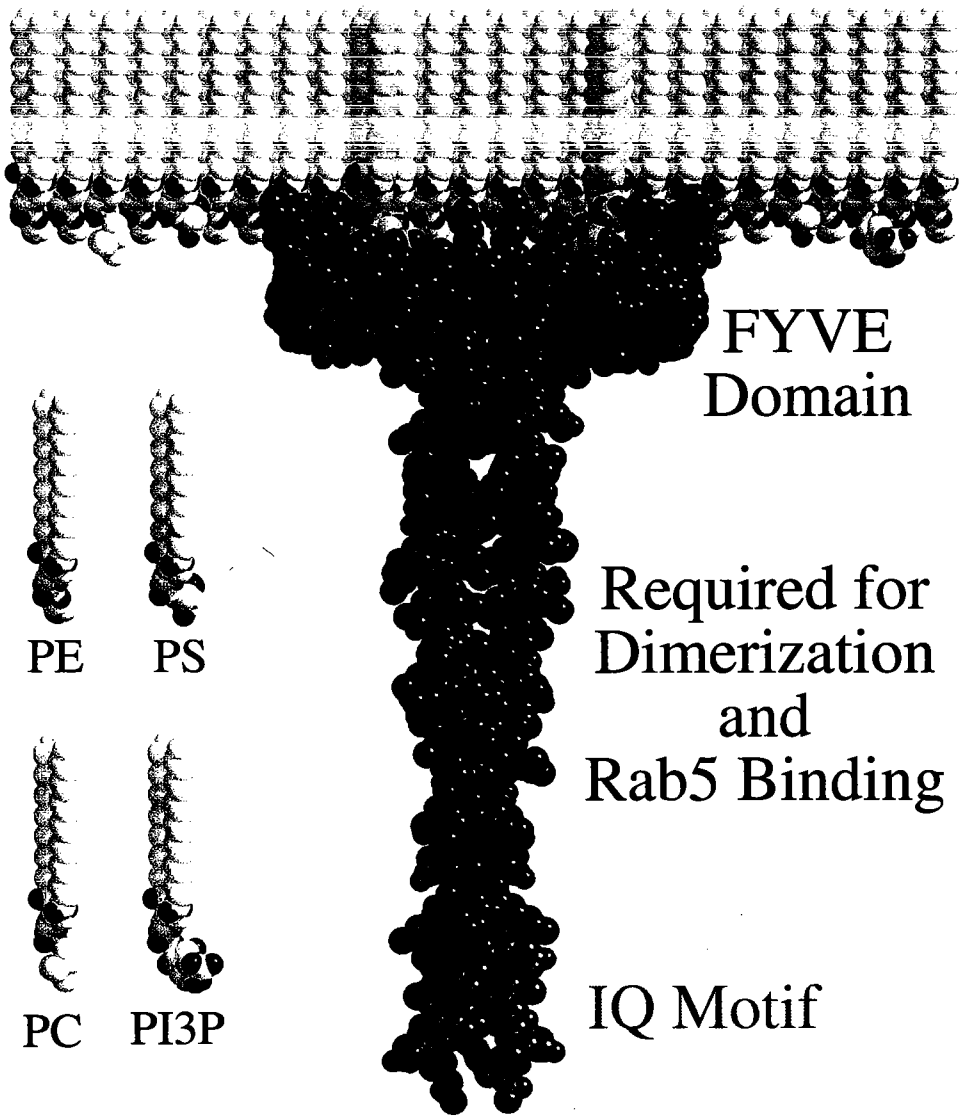


Figure 18. Model for multivalent membrane binding by homodimeric EEA1. Orientation of the EEA1 homodimer with respect to an idealized membrane leaflet deduced from the location and dyad symmetry of the PtdIns(3)P head groups (see text). The region required for both dimerization and interaction with Rab5 is highlighted in magenta. The putative IQ motif is highlighted in green.

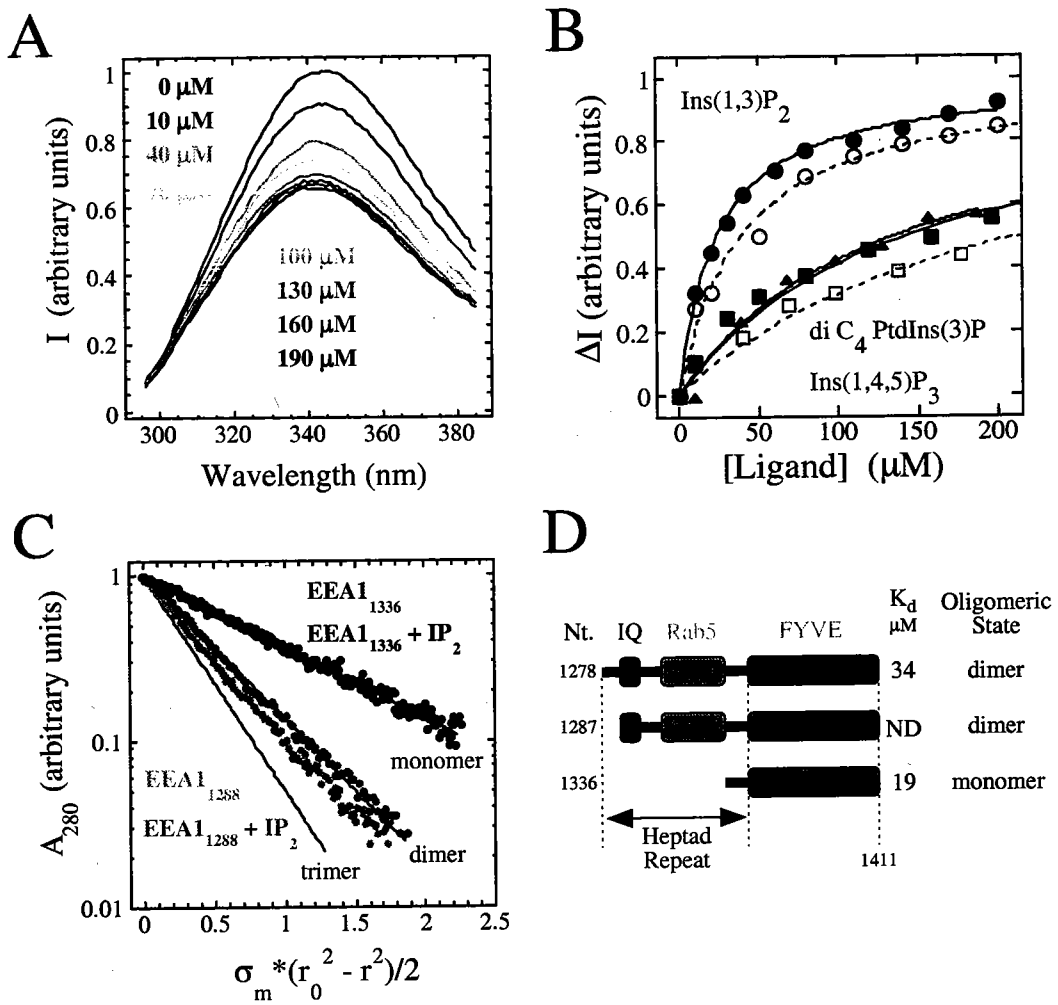


Figure 19. Head group binding properties and oligomeric state of EEA1 C-terminal constructs. **(A)** Emission spectra of EEA1$_{1278\text{-}1411}$ in the absence and presence of increasing concentration of Ins(1,3)P_2. Samples were excited at 290nm and the emission monitored from 300-400nm. **(B)** Quenching of the intrinsic tryptophan fluorescence as a function of ligand concentration for Ins(1,3)P_2 (black circles), dibutyl PtdIns(3)P (blue squares) and Ins(1,4,5)P_3 (red triangles). Solid lines correspond to the non-linear least squares fit to a simple two state binding model. **(C)** Sedimentation equilibrium experiments for EEA1$_{1336\text{-}1411}$ and EEA1$_{1287\text{-}1411}$ in the absence and presence of 1mM Ins(1,3)P_2. Solid lines represent predicted model functions calculated with the expected molecular mass for each species. **(D)** Summary of the oligomeric state and Ins(1,3)P_2 association constant (K_d) for three EEA1 C-terminal constructs showing that the affinity for head group and oligomeric state represent independent properties.

Both the decrease in emission intensity and the blue shift exhibit saturation at micromolar concentrations of ligand (Figure 19B). Control experiments with inorganic phosphate, inositol or mixtures of phosphate and inositol have no detectable effect on the emission spectrum at concentrations as high as 1 mM (the highest concentration tested). The dependence of the intrinsic tryptophan fluorescence on ligand concentration is well described by a simple two state binding model. The measured K_d of 34 μM for Ins(1,3)P_2 is roughly 4-8 fold lower than that estimated for either Ins(1,4,5)P_3 or the soluble dibutyl PtdIns(3)P analogue. The moderately decreased affinity for soluble PtdIns(3)P likely reflects the reduction in charge on the 1-phosphate incurred by covalent linkage to the diacyl glycerol moiety. Similar results are obtained for $EEA1_{1336-1411}$, which lacks most of the heptad repeat region. The characteristics of head group binding contrast with the considerably higher affinity (roughly 50 nM K_d) and specificity observed for the association of EEA1 with liposomes containing PtdIns(3)P (224,225,284,294).

The oligomeric state of C-terminal EEA1 constructs was established by sedimentation equilibrium experiments in the presence and absence of 1 mM Ins(1,3)P_2. At protein concentrations similar to those used for the fluorescence experiments, $EEA1_{1336-1411}$ sediments as a monomer whereas $EEA1_{1287-1411}$ and $EEA1_{1278-1411}$ sediment as dimers (Figure 19C). Addition of 1 mM Ins(1,3)P_2 eliminates a small contribution from higher order aggregation but otherwise has no apparent effect on the oligomeric state. Moreover, the binding data show no indication of cooperativity and are

independent of oligomeric state (Figure 19D). These results demonstrate that the PtdIns(3)P head group binds to independent sites on the EEA1 homodimer, indicative of an additive thermodynamic mechanism.

Given the modest affinity and specificity for soluble phosphoinositides, it is clear that the interaction with a single head group can not account for the PtdIns(3)P dependent targeting of EEA1 to early endosomes (226,283). Furthermore, non-specific membrane interactions involving the turret loop and tandem lysine residues cannot explain the requirement for a significant portion of the coiled-coil region (185,245). EEA1 localization has previously been attributed to dual interactions with Rab5 and PtdIns(3)P (162,245). However, a double mutant defective in Rab5 binding to the C-terminal region exhibits wild type localization to early endosomes (269). Moreover, the requirement for the coiled-coil region can be bypassed by fusing two EEA1 FYVE domains in tandem (294). Likewise, the Hrs FYVE domain exhibits a cytoplasmic distribution in cultured cells (199,294) but localizes to endosomes when fused to the coiled region of EEA1 or as a tandem fusion with a second Hrs FYVE domain (182,294). These observations strongly suggest that FYVE domains are unable to target endosomes as isolated monomers.

Given additive free energies, the K_d for bivalent binding to a homodimer will go as the square of the K_d for monovalent binding to an isolated monomer. Using the measured K_d of 130 μM for dibutyl PtdIns(3)P, the additive mechanism predicts a K_d of 17 nM for bivalent PtdIns(3)P binding, which agrees remarkably well with the observed value of 50 nM (182). The additive mechanism further predicts that the specificity for

PtdIns(3)P vs. PtdIns(4,5)P_2 will be amplified from the 4-8 fold level observed for soluble head group analogs to 16-64 fold for bivalent binding in the context of a lipid bilayer. Lipid binding assays suggest that the K_d for EEA1 binding to PtdIns(4,5)P_2 in membranes is at least 20 fold weaker than the K_d for PtdIns(3)P (182,224,225). Thus, the simplest thermodynamic model for bivalent PtdIns(3)P binding quantitatively accounts for the high affinity of the EEA1 FYVE domain for membranes containing PtdIns(3)P and also accounts for much of the increase in specificity.

The N-terminus of EEA1

The active forms of Rab5A, Rab5B, and Rab22 have been shown to interact directly with $EEA1_{1\text{-}209}$ (162,285,288). This region of EEA1 encompasses a hydrophilic sequence of ~35 residues, the C_2H_2 Zn^{2+} finger, and two consecutive heptad repeats. As shown in Figure 20A, 6xHis $EEA1_{1\text{-}218}$ co-precipitates with GST-Rab5C loaded with the non-hydrolyzeable GTP analog, GppNHp, but not does not co-precipitate with the GDP-bound form or in the presence of the Zn^{2+} chelating agent TPEN. These results are consistent with two hybrid experiments in which mutation of a cysteine residue involved in Zn^{2+} coordination disrupts the interaction with constitutively active Rab5A and Rab5B mutants (162,288). Although these observations demonstrate a requirement for an intact C_2H_2 Zn^{2+} finger, it is not clear whether this reflects a direct interaction between the C_2H_2 Zn^{2+} finger and Rab5 or whether the C_2H_2 Zn^{2+} finger plays an indirect structural

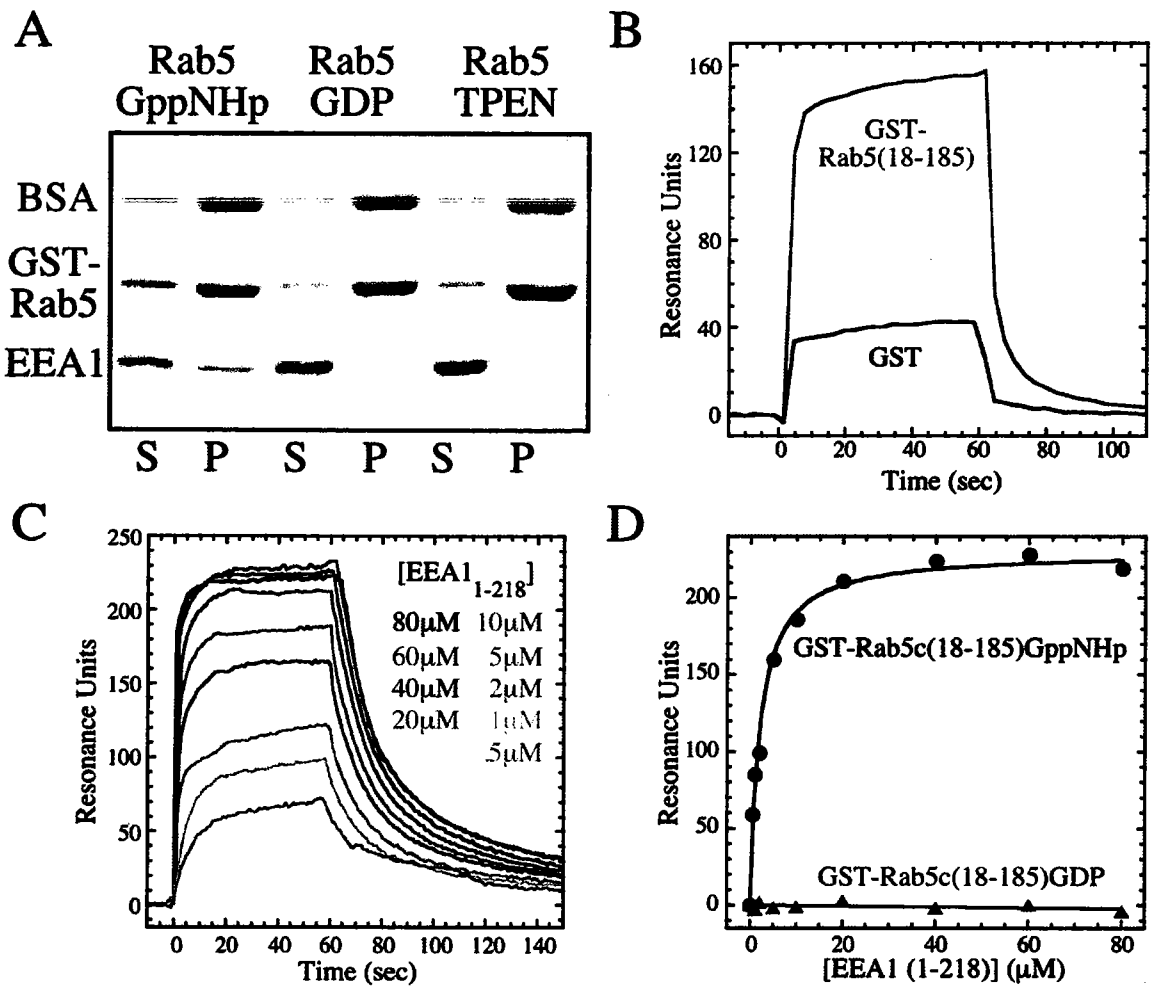


Figure 20. Nucleotide and Zn^{2+} dependent binding of Rab5C to the N-terminus of EEA1. **(A)** Co-precipitation of 6xHis EEA1$_{1-218}$ with GDP-bound GST-Rab5C or GppNHp-bound GST-Rab5C in the absence and presence of TPEN. 20 μM 6xHis EEA1$_{1-218}$ was incubated with 20 μM GST-Rab5C loaded with GDP or GppNHp in the absence and presence of 5 mM TPEN. S, supernatant following co-precipitation. P, glutathione elution of the pellet following co-precipitation. **(B)** SPR sensograms following injection of 20 μl of 10 μM 6xHis EEA1$_{1-218}$ at a flow rate of 20 μl/min into a dual channel flow cell in which anti-GST covalently coupled to a CM5 sensor chip was loaded with GST-Rab5C$_{18-185}$ (sample channel) or GST (reference channel). **(C)** SPR sensograms following injection of increasing concentrations of 6xHis EEA1$_{1-218}$ under the same conditions as in Figure 20B. Sample sensograms were corrected for bulk changes by subtraction of the reference sensogram. **(D)** Concentration dependence of the equilibrium SPR signal (R_{eq}) for 6xHis EEA1$_{1-218}$ binding to GST-Rab5C$_{18-185}$ loaded with GDP or GppNHp.

role by supporting interactions with flanking regions as is the case for the double Zn^{2+} finger of the Rab3A effector Rabphilin-3A (90).

As shown in Figure 20B-D, the interaction of GST-Rab5C-GppNHp with the N-terminus of EEA1 can also be detected and quantitatively analyzed by surface plasmon resonance (SPR) in a BIAcore instrument using a monoclonal GST antibody coupled to a CM5 sensor chip. When injected at concentrations in the low micromolar range, 6xHis $EEA1_{1-218}$ exhibits reversible binding to the GppNHp bound form of GST-$Rab5_{18-185}$ as judged by the amplitude of the SPR signal compared with the GST reference channel (Figure 20B). The signal in the reference channel rises and decays within the response time of the instrument, scales linearly with the concentration of 6xHis $EEA1_{1-218}$, and therefore represents either a bulk refractive index change or weak, reversible non-specific binding indistinguishable from a bulk refractive index change. Under the conditions of these experiments, the association of N-terminal EEA1 constructs with GST-Rab5C approaches equilibrium on the time scale of the injection (Figure 20C). The quantity bound at equilibrium (R_{eq}) saturates at low micromolar concentrations of 6xHis $EEA1_{1-218}$ (Figure 20D). The data are well fit by a simple Langmuir binding isotherm, yielding a dissociation constant (K_d) of 3.3 μM. In contrast, GST-$Rab5_{18-185}$ loaded with GDP shows no detectable binding to 6xHis $EEA1_{1-218}$, as expected for a bona fide GTPase-effector interaction. An equivalent affinity (K_d = 2.3 μM) is observed for the binding of 6xHis $EEA1_{36-218}$ to full length GST-Rab5-GppNHp, which includes the hypervariable

N- and C-terminal extensions, indicating that the interaction determinants reside within the GTPase domain.

To map the minimal interaction site at the N-terminus of EEA1 and determine whether the C_2H_2 Zn^{2+} finger is sufficient for Rab5 binding, SPR experiments were used to quantitatively analyze the binding of GST-$Rab5_{18-185}$ loaded with GppNHp to a panel of 6xHis EEA1 truncation constructs (Figure 21A). Elimination of the first 35 residues, corresponding to the hydrophilic N-terminus, has no significant effect on the interaction. Likewise, C-terminal truncations eliminating part or all of the heptad repeats show relatively small differences in affinity, which likely reflect systematic variations in the physical properties of the constructs. Indeed, 6xHis $EEA1_{36-91}$, which lacks the N-terminal hydrophilic region and both heptad repeats, binds in a nucleotide dependent manner to GST-$Rab5_{18-185}$ with an affinity comparable to that of 6xHis $EEA1_{1-218}$ (compare Figure 21B and C with Figure 20C and D). Consistent with this observation, $EEA1_{91-218}$, which lacks the N-terminal hydrophilic region and C_2H_2 Zn^{2+} finger, shows no detectable binding to GppNHp-bound GST-$Rab5_{18-185}$ at concentrations up to 150 μM (the highest concentration tested). A shorter construct corresponding to the minimal C_2H_2 Zn^{2+} finger defined by homology ($EEA1_{36-74}$) expressed poorly in bacteria and could not be fully purified. Although not suitable for quantitative analysis by SPR, this construct co-precipitates with GST-$Rab5_{18-185}$ loaded with GppNHp but not GDP and, by this measure, does not differ significantly from $EEA1_{36-91}$ (data not shown). I

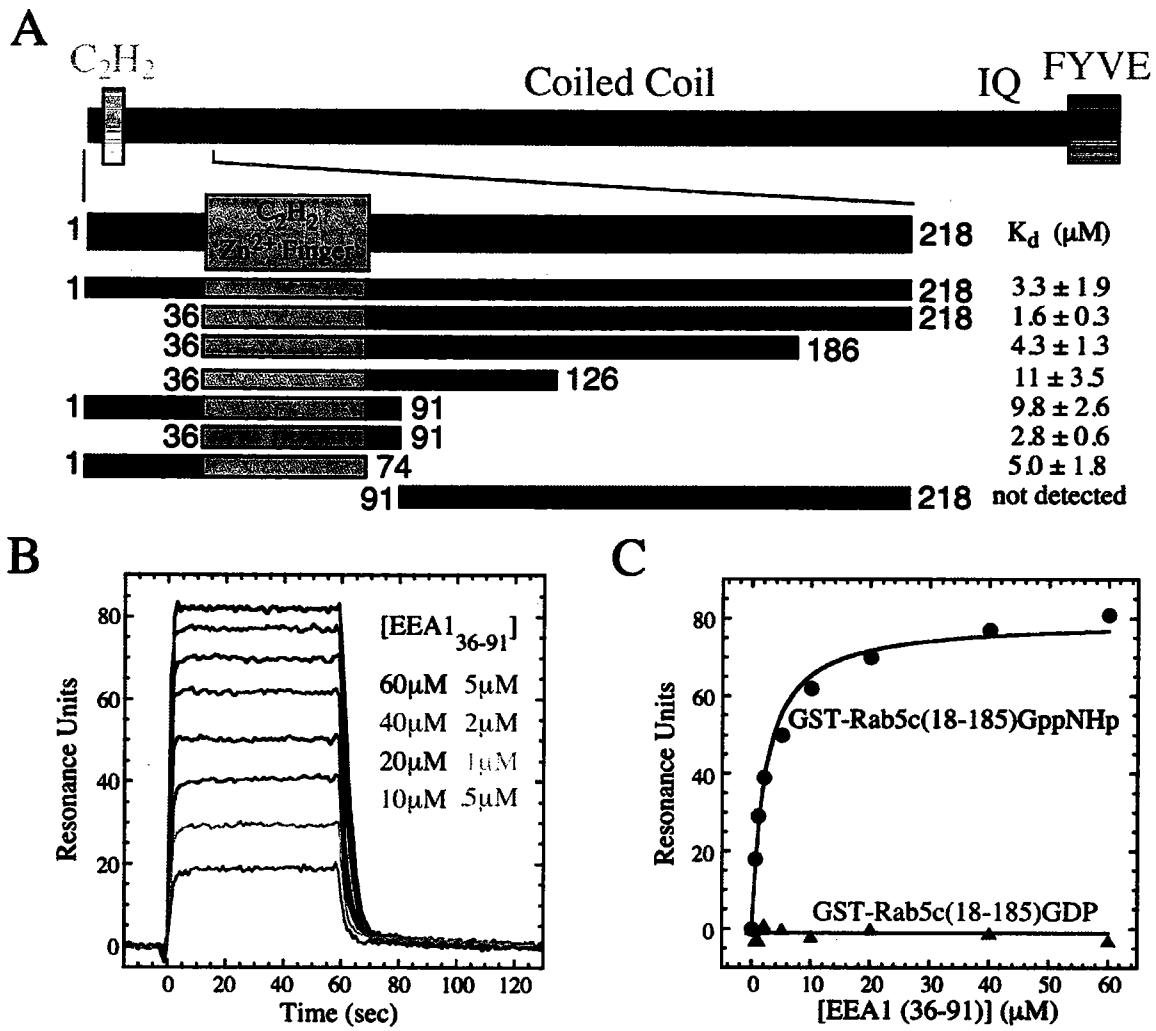


Figure 21. Identification of the minimal Rab5C binding site. **(A)** Panel of EEA1 constructs used to determine the minimal Rab5C binding site at the N-terminus of EEA1. Mean dissociation constants and standard deviations were calculated from 2-4 independent SPR experiments. **(B)** SPR sensograms following injection of increasing concentrations of 6xHis EEA136-91 under the same conditions as in Figure 20B. Sample sensograms were corrected for bulk changes by subtraction of the reference sensogram. **(C)** Concentration dependence of the equilibrium SPR signal (Req) for 6xHis EEA136-91 binding to GST-Rab5C18-185 loaded with GDP or GppNHp.

therefore conclude that the C_2H_2 Zn^{2+} finger is both necessary and sufficient for the interaction of Rab5C with the N-terminus of EEA1.

Full length EEA1 contains over 1200 residues of heptad repeat and forms a parallel coiled-coil homodimer in cells (295). To establish the oligomeric state of N-terminal EEA1 constructs, $EEA1_{1-91}$, $EEA1_{36-126}$, and $EEA1_{36-218}$ were centrifuged to equilibrium in an analytical ultracentrifuge. Whereas $EEA1_{1-91}$ sediments as a uniform monomer at a relatively high concentration of 20 μM, $EEA1_{36-126}$ and $EEA1_{36-218}$ sediment as uniform dimers at a lower concentration of 1 μM (Figure 22A). Thus, the heptad repeat proximal to the C_2H_2 Zn^{2+} finger provides sufficient driving force for stable homodimerization but does not contribute either directly or indirectly to the affinity for Rab5C.

Rab5C contains two tryptophan residues (Trp74 and Trp114) whereas the N-terminus of EEA1 contains a single tryptophan residue (Trp104). When titrated with 6xHis $EEA1_{36-91}$, which lacks tryptophan residues, the intrinsic tryptophan fluorescence of untagged GppNHp-bound $Rab5_{18-185}$ undergoes significant quenching accompanied by a small shift in the emission maximum (Figure 22B). Both effects saturate at low micromolar concentrations of 6xHis $EEA1_{36-91}$, indicative of a binding interaction. As shown in Figure 22C, the change in intrinsic tryptophan fluorescence is well described by a simple hyperbolic binding model, which yields a K_d of 1.1 μM, in good agreement with the affinity of 6x His $EEA1_{36-91}$ for GppNHp-bound GST-$Rab5_{18-185}$ measured by SPR.

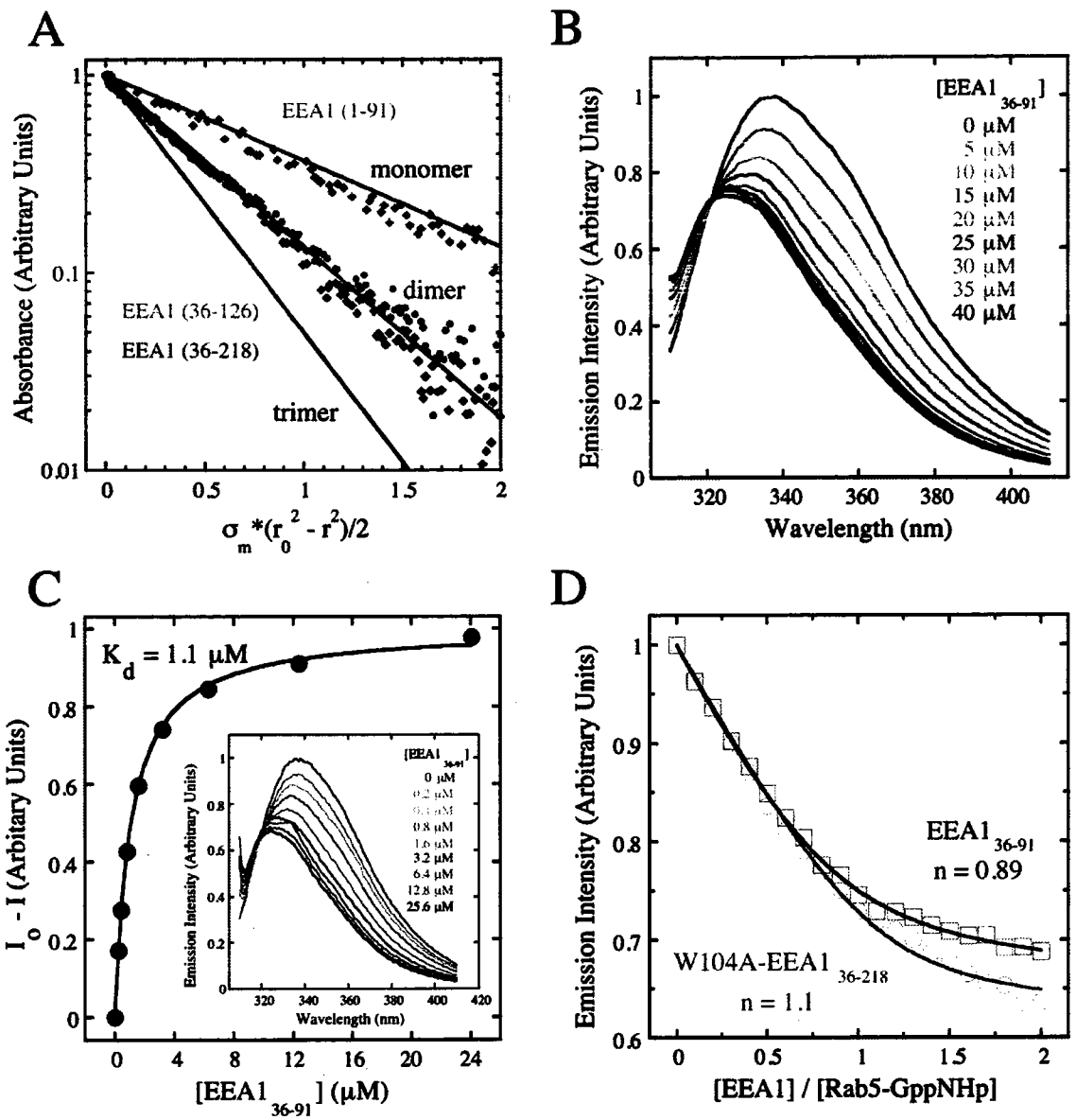


Figure 22. Oligomeric state and stoichiometry of Rab5C binding. **(A)** Sedimentation equilibrium experiments for EEA136-218, EEA136-126, and EEA11-91. Solid lines represent predicted model functions corresponding to uniform oligomeric states. **(B)** Intrinsic tryptophan emission spectra for 20 μM Rab5C18-185 in the presence of increasing concentrations of EEA136-91. Samples were excited at 300 nm and the emission detected at 340 nm. **(C)** Quenching of the intrinsic tryptophan fluorescence of 1μM Rab5C18-185 as a function of EEA136-91 concentration. Samples were excited at 300 nm and the emission detected at 340 nm. **(D)** Intrinsic tryptophan fluorescence for titration of 20μM Rab5C18-185 with EEA136-91 or the W104A mutant of EEA136-218. Samples were excited at 300 nm and the emission detected at 340 nm.

Consistent with these observations, the intrinsic tryptophan fluorescence of untagged Rab5-GDP is not perturbed by addition of 6xHis $EEA1_{36-91}$.

To determine whether homodimerization influences the stoichiometry of Rab5C binding to the N-terminus of EEA1, titration experiments employing intrinsic tryptophan fluorescence to monitor binding were conducted under conditions where the concentration of the fixed component (GppNHp-bound $Rab5_{18-185}$) was roughly 7 fold greater than the measured K_d. The resulting data were analyzed with a titration binding model (see Experimental Procedures) that relates the change in intrinsic tryptophan fluorescence to the binding stoichiometry (n), K_d, and the maximum change in intrinsic tryptophan fluorescence at binding saturation (ΔF_{max}). Because K_d cannot be accurately determined under titration conditions, it was fixed at the value obtained from the experiment in Figure 22C. Titration with 6xHis $EEA1_{36-91}$ yields a stoichiometry of 0.89, consistent with 1:1 binding. Titration with 6xHis $EEA1_{36-218}$ is complicated by the presence of a single tryptophan residue in the longer EEA1 construct. Because the magnitude of the change in GppNHp-bound $Rab5_{18-185}$ intrinsic tryptophan fluorescence is considerably larger than the intrinsic tryptophan fluorescence contributed by $EEA1_{36-218}$, the observed signal decreases monotonically until the majority of GppNHp-bound $Rab5_{18-185}$ is bound by $EEA1_{36-218}$, at which point the signal increases monotonically reflecting the contribution from excess $EEA1_{36-218}$ (data not shown). However, titration with the W104A mutant of $EEA1_{36-218}$, which binds with an affinity comparable to the wild type protein in the SPR experiment, yields a stoichiometry of 1.1, consistent with

two molecules of GppNHp-bound Rab5C binding to identical independent sites at the N-terminus of homodimeric EEA1.

To facilitate further characterization of the structural requirements underlying the interaction of Rab5C with the N-terminus of EEA1, a working homology model for the EEA1 C_2H_2 Zn^{2+} finger was constructed from the NMR structure of a C_2H_2 Zn^{2+} finger from the Adr1 transcription factor (296). The Adr1 structure was identified by threading against a structural database using the 3D-PSSM fold recognition server (297) and was selected from the lowest E-value structures on the basis of the sequence identity (29%) and the absence of gaps in the alignment. Non-identical residues in the Adr1 C_2H_2 Zn^{2+} finger were replaced with the corresponding residues in EEA1, which were modeled in the most common rotomer conformation. Although the resulting homology model does not represent an accurate representation of the actual structure, it is likely that the overall topology and approximate location of residues are preserved. The latter assertion is supported by extensive structural studies of weakly homologous C_2H_2 Zn^{2+} fingers, which share a common $\beta\beta\alpha$ fold.

The C_2H_2 Zn^{2+} fingers of EEA1 and Rabenosyn-5 conserve a number of residues in addition to those required for Zn^{2+} coordination or stability (Figure 23A). The majority of these residues are partially exposed in the homology model and cluster on a common surface, suggestive of a putative Rab5 interaction epitope (Figure 23B). To test this hypothesis, seven residues (Glu39, Phe41, Ile42, Pro44, Met47, Tyr60, and Glu61) were substituted with alanine and the mutant proteins characterized with respect to

A

```
                                    β-sheet   β-sheet   α-helix
hEEA1    M-{26}--NTVDVNNESSSEGFICPQCMKSLGSADELFKHYEAVHDAGNDSGHGG
cEEA1    M-{33}--QSQKSENDQEIEGFLCPMCMVELGGPDELTVHFEKEHSDFQTSANRN
hRbnsyn          MASLDDPGEVREGFLCPLCLKDLQSFYQLHSHYEEEHSGEDRDVKGQ
cRbnsyn  MIGATGSGGTGGHNDVVRQGFICPFCMEDFGEYERLICHVENDHPEEDSSDLAG
scVac1p                    MDLENVSCPICLRKFDNLQALNAHLDVEHGFNDNEDSLG
scAdr1   M-{118}-LKRHYRSHTNEKPYPCGLCNRCFTRRDLLIRHAQKIHSGNLGETISH
Consensus                          C  C   ϕ     L  H     H
```

B

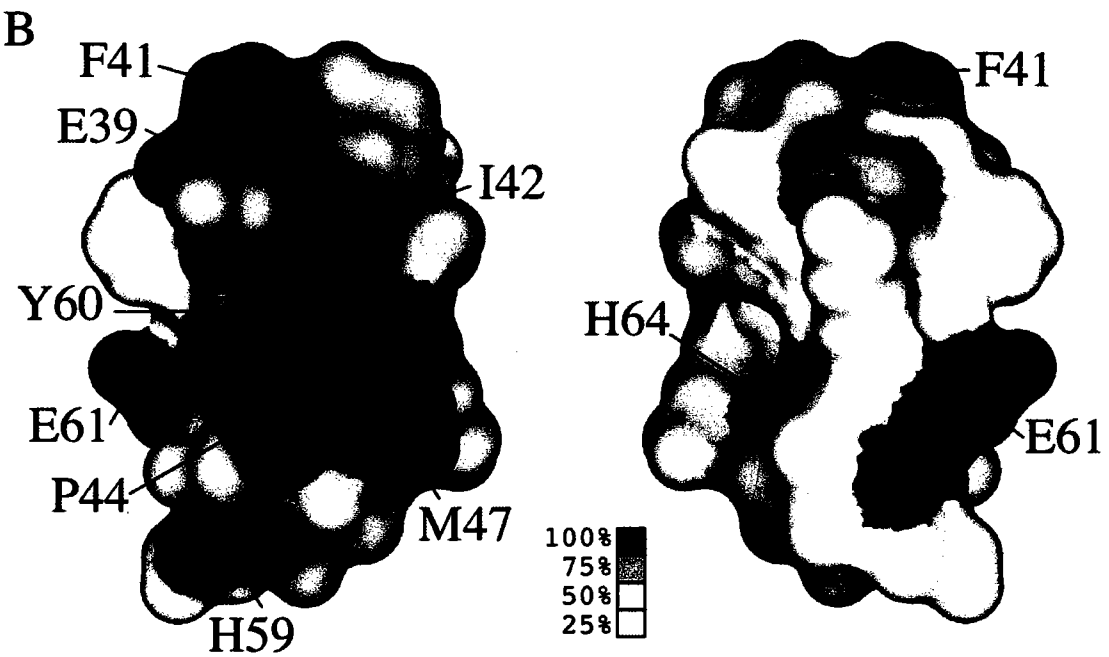


Figure 23. Homology model for the C_2H_2 Zn^{2+} finger of EEA1. **(A)** Alignment of the C_2H_2 Zn^{2+} fingers of EEA1, Rabenosyn-5, Vac1p, and Adr1. The secondary structure for Adr1 and conserved residues corresponding to the 'consensus' motif for classic C_2H_2 Zn^{2+} fingers are shown above and below the alignment, respectively. Residues are color coded according to the conservation of physiochemical similarity as defined by the following amino acid classes: small [GA], acidic [DE], basic [KR], polar [STNQ], Cβ-branched non-polar [VI], non-polar [LM], aromatic [FYW], and unique (P, C, or H). **(B)** Structural homology model for the C_2H_2 Zn^{2+} finger of EEA1 with conservation of physiochemical similarity mapped to the surface. The homology model was derived from the NMR structure of the Adr1 C_2H_2 Zn^{2+} finger (PDB ID code 1PAA) by substituting EEA1 side chains in the most common rotomer conformation. Physio-chemical similarity is defined as in A.

Rab5C binding and structural integrity. All seven mutants expressed in soluble form at levels comparable to the wild type protein. Substitution of Glu61 had little effect on the affinity for GppNHp-bound $Rab5C_{18-185}$. In contrast, alanine mutants involving Glu39 or any of the hydrophobic residues exhibited severe defects, with 10 fold or greater reduction in the affinity for GppNHp-bound $Rab5C_{18-185}$ (Figure 24 and Table 5). Consistent with these observations, the binding of GppNHp-bound $Rab5C_{18-185}$ to N-terminal EEA1 constructs was undetectable by isothermal titration microcalorimetry (data not shown), suggesting that the interaction is entropically driven, presumably by burying exposed hydrophobic surfaces. The results of the mutational analysis suggest that Rab5C should be capable of binding to the C_2H_2 Zn^{2+} finger of Rabenosyn-5 but not to the C_2H_2 Zn^{2+} finger of Vac1p, due to non-conservative substitutions involving critical residues (see Figure 23A). To test this hypothesis, the C_2H_2 Zn^{2+} fingers of Rabenosyn-5 and Vac1p were expressed as GST fusions and the interaction with 6xHis $Rab5C_{18-185}$ determined by SPR (Figure 25). Whereas the C_2H_2 Zn^{2+} finger of Rabenosyn-5 binds to $Rab5C_{18-185}$ with a K_d of 0.6 μM, the C_2H_2 Zn^{2+} finger of Vac1p exhibits no detectable binding at concentrations of $Rab5C_{18-185}$ as high as 150 μM (the highest tested).

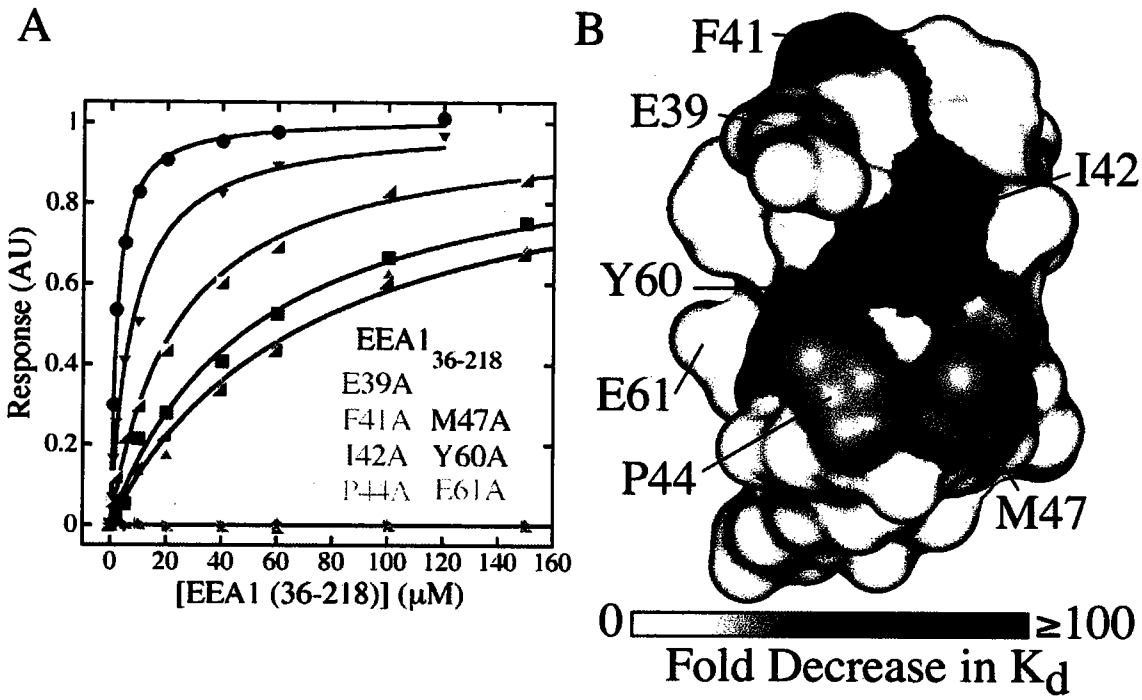


Figure 24. Determinants of Rab5C binding to the C_2H_2 Zn^{2+} finger of EEA1. **(A)** Effects of alanine substitutions on the affinity of GST-Rab5C18-185 for EEA136-218 as assessed by SPR under the same conditions as in Figure 20B. **(B)** Relative affinity, Kd (mutant) / Kd (wild type), mapped to the surface of the structural homology model for the C_2H_2 Zn^{2+} finger of EEA1.

Table 5. Dissociation constants for GppNHp-bound GST Rab5C$_{18-185}$ binding to mutants in the C_2H_2 Zn^{2+} finger of 6xHis EEA1$_{36-218}$

Protein	K_d (μM)	K_d (mutant) / K_d (wt)
Wild Type (wt)	1.6 ±0.3	NA
E39A	22.4 ±7.2	14
F41A	>200*	>100
I42A	>200*	>100
P44A	66.5 ±12.5	41
M47A	67.9 ±16.7	42
Y60A	50.3 ±12.1	31
E61A	5.4 ±2.3	3.3

* Conservative estimate reflecting the absence of a detectable SPR signal above that of the reference channel at a concentration of 200 μM 6xHis EEA1$_{36-218}$.

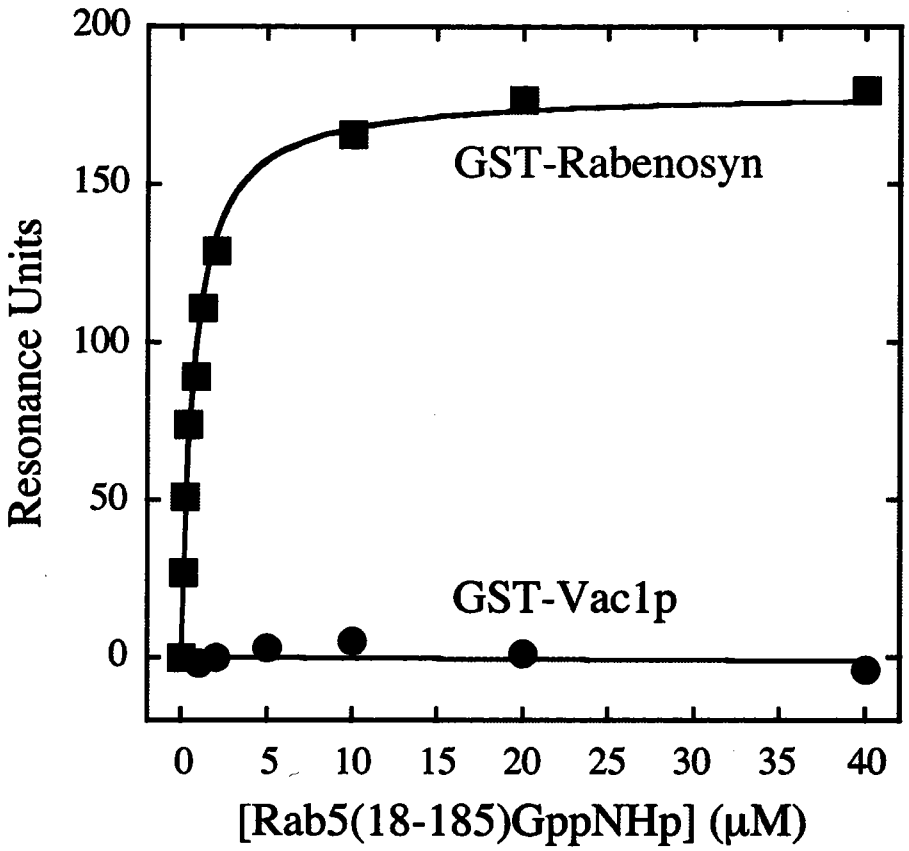


Figure 25. Rab5C binds to the C_2H_2 Zn^{2+} finger of Rabenosyn-5 but not Vac1p. The interaction of GppNHp-bound Rab5C18-185 with GST fusions of the Rabenosyn-5 and Vac1p C_2H_2 Zn^{2+} fingers was assessed by SPR under the same conditions as in Figure 20B.

Discussion

The C-terminus of EEA1

A striking feature of the EEA1 homodimer is the location of dyad symmetric phosphoinositide binding sites on a common surface formed as the direct consequence of the unique domain organization and quaternary structure of the C-terminal region. The resulting surface is ideally configured for simultaneous binding of two PtdIns(3)P head groups, provided that the dyad axis of the FYVE domain homodimer is oriented normal to the membrane surface as depicted for an idealized membrane leaflet in Figure 18. Any other orientation would favor interactions with one FYVE domain at the expense of the other, necessitating an energetically unfavorable deformation of the membrane to sustain simultaneous head group binding. The depth of penetration into the bilayer can be established by translation parallel to the membrane normal until the 1-phosphate groups of the dyad related head groups in the EEA1 homodimer are positioned at the level of the corresponding 1-phosphate groups in the idealized membrane leaflet. The resulting model predicts that residues ^{1367}VT^{1368} at the tip of the turret loop would penetrate into the interfacial region of the lipid bilayer while the remaining residues in the turret loop (Ser1366 and Val1369) would reside within the head group region. This prediction, deduced from the quaternary structure of the liganded EEA1 homodimer and the assumption of simultaneous PtdIns(3)P binding, agrees well with the observation that resonances for residues ^{1365}FSVTV^{1369} shift and broaden in the presence of dodecyl phosphatidyl choline micelles (293,298) and supports the hypothesis that the exposed non-polar residues at the tip of the turret loop insert into interfacial region (244).

Although the energetic contribution from non-specific membrane interactions with the turret loop and/or tandem lysine residues is not known, it is expected to be weak given the small number of residues in direct contact with the membrane (244). This view is supported by the non-specific association of the unliganded EEA1 FYVE domain with dodecyl phosphatidyl choline micelles, which is not saturated at a high detergent concentration of 600 mM (293). Furthermore, the FYVE domain construct used for these studies includes twenty residues from the coiled-coil region and was reported to form dimers at the concentrations of the NMR experiments. These observations are consistent with a monomer K_d for non-specific binding to dodecyl phosphatidyl choline micelles in the mid to high mM range, which is several orders of magnitude weaker than the dissociation constant for the specific interaction with soluble dibutyl PtdIns(3)P. If the energetic contribution from non-specific binding was comparable to the specific interaction with the head group, then EEA1 dimers should bind non-specifically to membranes with nanomolar affinity and have femtomolar affinity for membranes containing PtdIns(3)P. Thus, both experimental observations and thermodynamic considerations support a bivalent mechanism of PtdIns(3)P recognition as the primary driving force for selective amplification of the weak affinity of the EEA1 FYVE domain for the head group of PtdIns(3)P. Moreover, the critical role of the coiled-coil region for stable dimerization explains the requirement of the region N-terminal to the FYVE domain for endosomal localization.

An unresolved question concerns whether the apparently rigid organization of the EEA1 homodimer is maintained in the absence of head group. NOE data for an isolated

EEA1 FYVE domain provide evidence that residues in the linker segment undergo a localized conformational change upon head group binding (298). Because the conformation and mobility of the linker segment in the crystal structure differs substantially from both the liganded and unliganded NMR structures, it is possible that a structural change involving the linker segment does not occur in the context of the intact EEA1 homodimer. We note, however, that crystals of $EEA1_{1287\text{-}1411}$ are not obtained in the absence of $Ins(1,3)P_2$. Both head group binding sites are situated adjacent to a solvent channel in the crystals and neither $Ins(1,3)P_2$ molecule participates in intermolecular crystal contacts, consistent with a difference in tertiary and/or quaternary structure. Moreover, the conserved aspartic acid residue of the WxxD motif at the N-terminus of the linker region directly contacts the head group, suggesting a plausible linkage between head group binding and the conformation of the linker region. As noted above, the FYVE domain dimer interface contributes weakly, if at all, to the stability of the EEA1 homodimer. In the absence of head group, the engagement of the linker region with the FYVE domain core could be weakened, perhaps relaxing the rigid quaternary structure of the liganded FYVE domain dimer. Definitive resolution of this question will require structural data on the unliganded EEA1 homodimer.

Based on the quaternary organization of the liganded EEA1 homodimer and taking into account the available physical, biochemical and cell biological data, we propose the following kinetic model for the sequence of events resulting in localization of EEA1 to early endosomes. The unliganded form of the EEA1 homodimer engages intracellular membranes through transient, non-specific interactions. This mode of non-

selective membrane sampling reflects a high off-rate, which limits the extent of lateral diffusion and strongly favors a cytoplasmic distribution. The presence of PtdIns(3)P is detected by interactions with the invariant residues of the R(R/K)HHCR and RVC motifs, thereby increasing the mean residence time at endosome membranes. The interactions between the head group and the conserved aspartic acid residue of the WxxD motif may trigger a conformational change that fully engages the FYVE domain core with respect to the linker region. Lateral diffusion rapidly facilitates a similar sequence of events for the second FYVE domain. Although the affinity for a single PtdIns(3)P head group is relatively weak, the multivalent quaternary structure of EEA1 facilitates stable association with endosome membranes. Finally, the weak binding of Rab5 to the C-terminal region may further restrict localization to Rab5 positive endosome membranes.

The N-terminus of EEA1

The crystal structure of a constitutively active mutant of Rab3A bound to the minimal Rab3A binding domain (RBD) of Rabphilin-3A provides the only available structural data on the interactions of Rab GTPases with effectors (90). The RBD of Rabphilin-3A consists of an N-terminal helix, a double Zn^{2+} finger with a fold similar to that of a FYVE domain, and a C-terminal loop/helix. Although the double Zn^{2+} finger does not contact Rab3A directly, it serves an indirect structural role by supporting interactions with the N-terminal helix and C-terminal loop/helix. The interaction of Rab5 with the C-terminus of EEA1 requires an intact FYVE domain and is disrupted by point mutants in the proximal coiled-coil region (185,269). Although the interaction with the

C-terminus has not been quantitatively analyzed, Rab5 does not co-precipitate with either the proximal coiled-coil or the FYVE domain alone, suggesting that the binding site encompasses both regions (269). Using purified reagents and quantitative measures of binding, I have shown that C_2H_2 Zn^{2+} finger (residues 36-74) is both necessary and sufficient for the interaction of Rab5C with the N-terminus of EEA1. Thus, the function of the C_2H_2 Zn^{2+} finger at the N-terminus of EEA1 differs fundamentally from the indirect structural role of the double C_2H_2 Zn^{2+} finger of Rabphilin-3A. The mode of interaction also differs from that at the C-terminus of EEA1, which requires both the FYVE domain and the proximal coiled-coil.

Two variations are common in classic C_2H_2 Zn^{2+} fingers, the 'consensus' motif (**C**-X_2-**C**-X_3-F-X_5-L-X_2-**H**-X_{3-5}-**H**) and the 'swapped' motif (**C**-X_2-**C**-X-F-X_7-L-X_2-**H**-X_{3-5}-**H**) (299,300). These motifs differ in the location of the central phenylalanine residue, which in either case packs in the small hydrophobic core with the conserved leucine and one of the conserved histidine residues. Both motifs adopt a similar $\beta\beta\alpha$ fold stabilized by a tetrahedral Zn^{2+} ion, which forms a tetrahedral coordination complex with the thiol groups of the conserved cysteine residues contributed by adjacent strands of the β hairpin and the imidazole side chains of the conserved histidine residues at the C-terminus of the α helix. The C_2H_2 Zn^{2+} finger of EEA1 closely resembles the consensus motif with the exception that the phenylalanine residue following the conserved cysteine residues is replaced by a leucine residue. The corresponding substitution has been studied in the context of the C_2H_2 Zn^{2+} finger of ZFY, where it has little effect on the

structure or affinity for Zn^{2+}, although it does increase the overall dynamic mobility (300). Whether the latter effect would be compensated by other substitutions in EEA1 or otherwise contribute to the interaction with Rab5 is not clear. Although it is possible that the observed defects in some mutants reflect an indirect effect on the folding and/or structure of the C_2H_2 Zn^{2+} finger, alanine residues occur naturally in C_2H_2 Zn^{2+} fingers at each of the positions examined in this study. Consistent with this observation, C_2H_2 Zn^{2+} fingers have been shown to be tolerant of alanine substitutions at non-consensus positions (301). Furthermore, each mutated residue occupies an exposed or partially exposed position in the structural homology model. Finally, the mutants with severe defects involve residues that cluster on a common surface. Therefore, I favor the interpretation that the defects arise from altered interactions with Rab5.

Rab5C contains two tryptophan residues, one located in the β3 strand at the interface between the switch I and II regions (Trp74) and the other located in the α3 helix adjacent to the switch II region (Trp114). In the crystal structure of GppNHp-bound Rab5C, Trp74 is partially exposed whereas Trp114 is buried (237). Although it is possible that the observed changes in intrinsic tryptophan fluorescence arise indirectly from structural rearrangements in the vicinity of one or both tryptophan residues, a more straightforward explanation is that the partially exposed Trp74 is located in or adjacent to the epitope for interaction with the C_2H_2 Zn^{2+} finger of EEA1. Trp74 is flanked by two other partially exposed hydrophobic residues, Phe58 in the switch I/β2 region and Tyr90 in the switch II region. This triad of invariant hydrophobic residues constitutes a prominent, exposed hydrophobic patch in the active form of Rab GTPases with known

structure and lies at the core of the interface between the switch regions of Rab3A and Raphillin3A (90). I have previously shown, in chapter II, that the conformation of the invariant hydrophobic triad varies between Rab GTPases and thereby contributes to the specificity of Rab-effector interactions (237). Moreover, the hydrophobic triad is located adjacent to variable loop regions also implicated in Rab-effector specificity (90). I therefore propose that the invariant hydrophobic triad in Rab5 lies at the core of a largely non-polar interface with the cluster of conserved hydrophobic residues in the C_2H_2 Zn^{2+} finger of EEA1. This hypothesis is generally consistent with known modes of GTPase-effector interaction and would explain the apparent lack of a significant enthalpic component to the free energy of binding inferred from the inability to detect the interaction by isothermal titration microcalorimetry. A more detailed understanding of the structural basis underlying the interaction of Rab5 with the N-terminus of EEA1 will require the structure of the C_2H_2 Zn^{2+} finger in complex with Rab5.

EEA1 dependent endosome tethering

In vitro fusion assays as well as studies in cultured cells implicate both EEA1 and Rab5 in the tethering of early endosomes (154,162,163,226,227,269,302). Tethering is thought to represent a general intermediate preceding the docking, priming and fusion of vesicles with target membranes. EEA1 has been proposed to tether early endosomes/endocytic vesicles by binding PtdIns(3)P at its C-terminus and Rab5 at its N-terminus (163). The organized quaternary structure of EEA1 strongly supports this hypothesis and serves as the basis for a plausible structural model for the tethered

intermediate (Figure 26). Bivalent binding is expected to significantly enhance the affinity for membranes containing PtdIns(3)P; however, the isolated FYVE domain of EEA1 does not form stable dimers. These observations explain the strict requirement of the proximal heptad repeat for endosome targeting and why it can be bypassed by fusing two FYVE domains in tandem (294).

Although the heptad repeat at the N-terminus also provides the driving force for stable dimerization, it neither enhances nor interferes with Rab5 binding. Moreover, titration experiments indicate that the N-terminus of the EEA1 homodimer supports simultaneous, independent binding of two Rab5 molecules. Whether one or two molecules of Rab5 bind in a cellular context would depend on the effective concentration of active Rab5 on the membranes of endosomes or endocytic vesicles. It seems unlikely that the low micromolar K_d for Rab5 binding to the N-terminus of EEA1 would be sufficient for stable endosome tethering. However, the combined action of multiple EEA1 molecules, each with the potential for bivalent Rab5 binding, could generate a stably tethered intermediate with considerable dynamic flexibility, given the glycine rich hydrophilic sequence following the C_2H_2 Zn^{2+} finger, several predicted hinge regions in the heptad repeat near the N-terminus of EEA1, and the glycine/proline rich hydrophilic sequence connecting the GTPase domain of Rab5 to the dual prenylation motif at the C-terminus.

With the exception of the N-terminal C_2H_2 Zn^{2+} finger and C-terminal FYVE domain, a strong heptad repeat is present throughout the sequence of EEA1. Several proline-glycine motifs (residues ^{126}PDG^{128}, ^{218}PG^{219}, and ^{279}GP^{280}) break the heptad

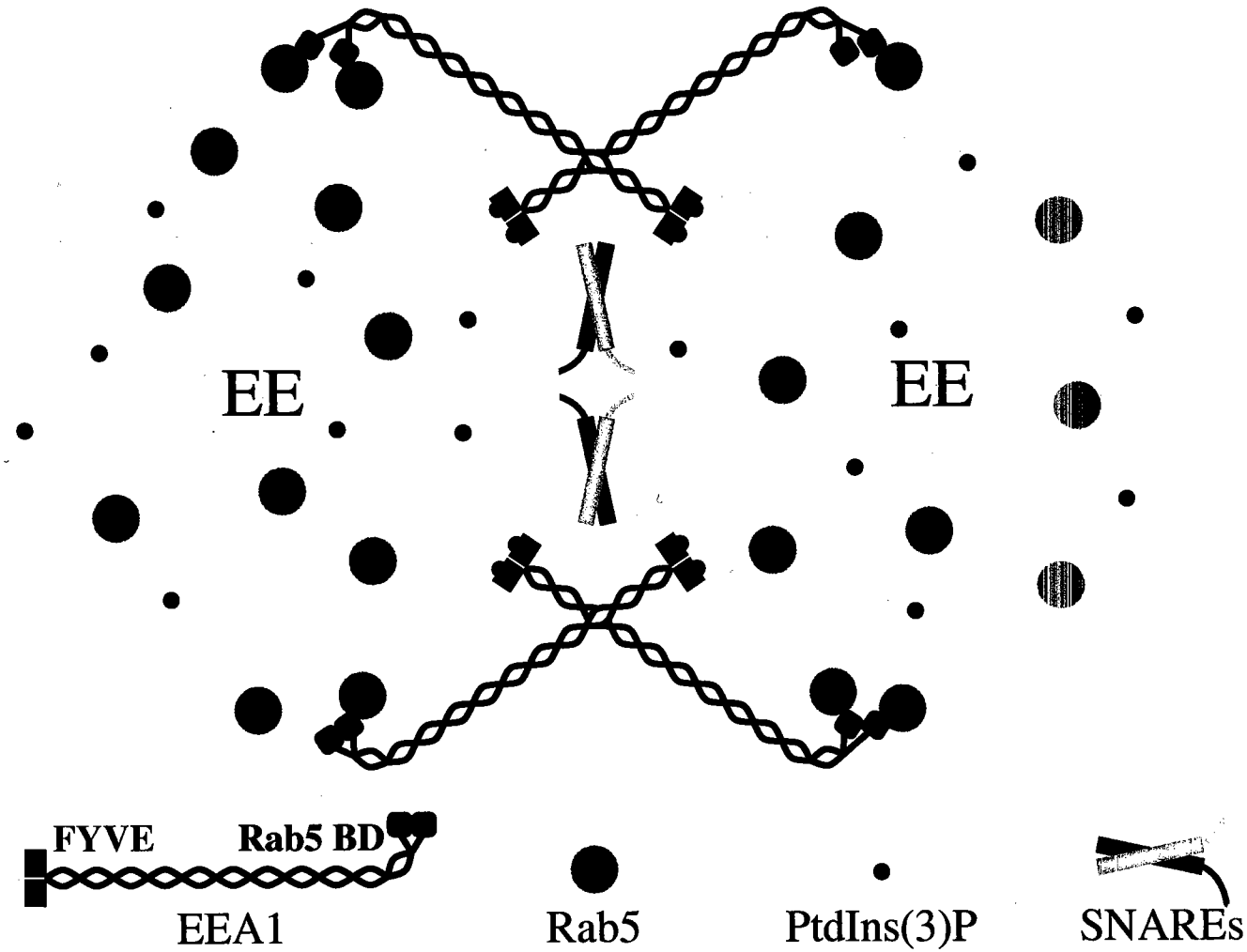


Figure 26. Hypothetical model for endosome tethering by homodimeric EEA1. Early endosomes (EE) with a diameter of 200nm are depicted as beige spheres scaled to the predicted length of an EEA1 homodimer. See text for description of the model.

repeat in the vicinity of the N-terminal Rab5 binding site. However, similar flexible kink regions are absent from the remainder of the protein. Thus, residues 281-1346 of EEA1 are predicted to form a continuous parallel coiled-coil. Given the mode of multivalent membrane binding depicted in Figure 18, the long axis of the coiled-coil would be oriented normal to the endosome surface such that the N-terminal Rab5 binding site would be located ~160-180 nm from the endosome surface. Early endosomes coated with such EEA1 homodimers would be ideally configured to capture and tether Rab5 positive endocytic vesicles or other early endosomes. The rigid quaternary structure of the EEA1 C-terminal region ensures that the coiled-coil extends directly into the cytoplasm, thereby preventing the N-terminus from folding back and binding Rab5 proteins on same endosome.

EEA1 possesses two distinct Rab5 binding sites, one corresponding to the C_2H_2 Zn^{2+} finger and the other overlapping the C-terminal FYVE domain and proximal coiled-coil region. We estimate that the solution affinity of unprenylated Rab5 for C-terminal constructs of EEA1 is at least an order of magnitude weaker than that observed for the C_2H_2 Zn^{2+} finger. Though weaker, the C-terminal interaction may nevertheless contribute significantly *in vivo*, as a consequence the restricted dimensionality resulting from co-localization on endosome membranes. Indeed, point mutants that disrupt Rab5 binding to the C-terminal site exhibit a dominant negative phenotype when expressed at high levels in cultured cells (269). Recently, the interactions of calmodulin and Rab5 with the C-terminal region of EEA1 have been reported to antagonize PtdIns(3)P binding (303). One role of these interactions may be to disrupt the quaternary structure of the C-

terminal region, thereby effecting release from the tethered state in conjunction with membrane fusion.

Like EEA1, Rabenosyn-5 also possesses two apparently independent Rab5 binding sites corresponding to the C_2H_2 Zn^{2+} finger (present study) and a second site within the C-terminal third of the protein (155). Although the C_2H_2 Zn^{2+} fingers of EEA1 and Rabenosyn-5 bind Rab5 with an affinity and nucleotide specificity characteristic of *bona fide* GTPase effectors, the functional significance of these interactions has not been tested *in vivo*. If, as hypothesized, the N-terminal interaction with Rab5 plays a critical role in the tethering and/or fusion of early endosomes, then cells expressing full length EEA1 constructs carrying the F41A or I42A mutations in the C_2H_2 Zn^{2+} finger should exhibit a strong dominant negative phenotype. I anticipate that these mutants and the analogous point mutations in Rabenosyn-5 will provide useful reagents for exploring the *in vivo* functional role of the homologous C_2H_2 Zn^{2+} fingers in the context of the full length proteins.

CHAPTER IV

DETERMINANTS OF RAB-GEF RECOGNITION AND THE STRUCTURE AND SPECIFICITY OF THE RABEX-5 HELICAL BUNDLE-VPS9 CATALYTIC TANDEM

Summary

Rab GTPases function as essential regulators of vesicle transport between sub-cellular compartments of eukaryotic cells. Mss4, an evolutionarily conserved Rab accessory factor, facilitates nucleotide release and binds tightly to the nucleotide free form of exocytic but not endocytic Rab GTPases. A structure based mutational analysis of residues that are conserved only in exocytic Rab GTPases reveals three residues that are critical determinants of the broad specificity recognition of exocytic Rab GTPases by Mss4. One of these residues is located at the N-terminus of the switch I region near the nucleotide binding site whereas the other two flank an exposed hydrophobic triad previously implicated in effector recognition. The spatial disposition of these residues with respect to the structure of Rab3A correlates with the dimensions of the elongated Rab interaction epitope in Mss4 and supports a mode of interaction similar to that of other exchange factor/GTPase complexes.

Activation of the Rab5 GTPase, an essential regulator of endocytic vesicular trafficking, is mediated by modular exchange factors that contain a Vps9 homology

domain. Here we report the crystal structure, exchange determinants, and family-wide Rab specificity of the catalytically active core of the Rab5 exchange factor Rabex-5. The structure reveals a tandem architecture in which the Vps9 domain is stabilized by an indispensable helical bundle. An invariant aspartic acid residue in an ordered loop of the Vps9 domain adjoins a shallow groove lined with highly conserved residues from a pair of adjacent helices, the arrangement of which resembles a critical helical substructure in the Sec7 domain of Arf exchange factors. Mutation of conserved residues within the putative Rab5 binding surface, but not adjacent regions, severely impairs exchange activity. Finally, a quantitative Rab family-wide interaction analysis indicates that the Rabex-5 catalytic tandem has highly selective exchange activity for Rab GTPases of the Rab5 subfamily.

This chapter contains work previously published in Zhu et al. (304) and Delprato et al. (submitted). Consensus determinants for Rab GEF interactions were defined using data generated by Zhongyuan Zhu (304) and previously published reports (114-117). Some Rabex-5 constructs and the diffraction quality crystals were generated by Anna Delprato. Data collection, refinement and analysis of the Rabex-5 HB-Vps9 structure were conducted by Anna Delprato, David Lambright and myself. Rabex-5 point mutations were generated and analyzed by Anna Delprato and myself. Rab family wide analysis of Rabex-5 exchange specificity was conducted by Anna Delprato using a collection of Rab proteins generated by Anna Delprato, Sudharshan Eathiraj, Xiaojing Pan, Chris Ritacco and myself.

Introduction

As critical regulators of intracellular membrane trafficking, Rab proteins comprise the largest GTPase family with at least 38 functionally distinct proteins and an additional twenty isoforms encoded in the human genome (4,11,59,60). Like other GTPases, Rab proteins cycle between active (GTP bound) and inactive (GDP bound) conformations. Accessory factors, in some cases coupled to signaling networks, regulate the GTPase cycle by modulating membrane association, nucleotide binding, and GTP hydrolysis. Upon targeting to donor membranes, Rab GTPases are activated by GEFs. In the active conformation, Rab GTPases interact with diverse effector proteins to facilitate vesicle budding from donor membranes, cargo sorting, and motor-dependant transport as well as the tethering, docking, and fusion of vesicles with acceptor membranes. GTP hydrolysis accelerated by GAPs completes the GTPase cycle, allowing recovery of prenylated Rab GTPases as a soluble complex with RabGDI.

The interaction of Rab GTPases with regulatory factors and effectors reflects multiple levels of specificity. The majority of known Rab exchange factors, GAPs and effectors exhibit a high degree of specificity towards Rab GTPases, consistent with the diverse regulatory mechanisms required for the function of Rab GTPases in distinct trafficking pathways. Rab GDI, on the other hand, broadly recognizes members of the Rab family but not other GTPase families (93). Several Rab accessory factors exhibit an intermediate level of specificity for distinct Rab sub-families (95,104,109,305). The high specificity of Rab effector interactions is reflected in the sequence diversity characteristic of the Rab family and determined in part by hypervariable 'Rab complementarity

determining regions' (Rab CDRs), which include the N- and C-termini of the GTPase domain as well as the α3/β5 loop (69,90,92). An indirect mechanism of effector specificity determination involves sequence variability in the hydrophobic core, which is reflected in the active conformation of a triad of invariant hydrophobic residues located at the interface between the switch I and II regions (237). Although a phylogenetic analysis indicates strong evolutionary conservation in the sequences of Rab GTPases that function in common trafficking pathways (48), little is known about the determinants of Rab subfamily recognition.

Mammalian suppressor of Sec4 (Mss4) binds tightly to the nucleotide free forms of exocytic Rab GTPases but does not interact with endocytic Rab proteins (34,94). Homologues of Mss4 are present in evolutionarily diverse species including both fission and budding yeast, worms, flies, zebra fish and mammals (119). Overexpression of Mss4 or the budding yeast homologue Dominant suppressor of Sec4 (Dss4) suppresses the lethal phenotype of temperature sensitive dominant negative Sec4 mutants (34,306). Although viable, a Dss4 null exhibits a synthetic negative phenotype when combined with a temperature sensitive mutant of Sec2, an essential Sec4 exchange factor (33). Mss4 is broadly expressed in different tissues but is most abundant in the brain and is one of the few factors that stimulates neurotransmitter release when injected into squid giant nerve terminals (94). These observations implicate Mss4 proteins as Rab accessory factors that regulate exocytosis in cooperation with more potent Rab GEFs (33,307,308). Aberrant overexpression of Mss4 has been reported in a wide variety of malignant tissues, including human pancreatic and colon cancers, suggesting a potential role in

cancer progression through enhanced secretion of trophic factors required for tumor proliferation and maintenance (309).

Studies employing Rab3A/Rab5 chimeras demonstrate that the determinants of the broad recognition of exocytic Rab proteins by Mss4 reside within the N-terminal third of the GTPase domain, which includes the P-loop as well as both conformational switch regions (305). Within this region, exocytic Rab GTPases conserve a number of residues that are variable in endocytic Rab proteins. Several of these residues have been implicated in intramolecular interactions and thus may be conserved for functions other than Mss4 recognition (18). For example, serine residues in the P-loop and switch I region of exocytic Rab GTPases regulate the intrinsic and, potentially, the GAP accelerated rates of GTP hydrolysis through hydrogen bonding interactions with the γ phosphate of GTP (18,239,310,311). Consequently, it is not clear which of the residues conserved in exocytic Rab GTPases comprise the determinants for broad specificity recognition by Mss4.

Rab5 has been implicated as a master regulator of endocytic vesicular trafficking, from clathrin-coated vesicle budding to early endosome fusion. The function of Rab5 in endosome fusion is mediated by several essential effectors, including the tethering factor EEA1 as well as the divalent effectors Rabaptin-5 and Rabenosyn-5, which possess separate binding sites for the GTP-bound forms of Rab5 and Rab4 (53,61,155,162,226,227,241,242,270). Rabex-5 (Rabaptin-5 associated exchange factor for Rab5) was originally identified as a 60kDa protein that co-purified as a stable endogenous complex with Rabaptin-5 (120). The Rabex-5-Rabaptin-5 complex

cooperates with other factors to promote homotypic endosome fusion as well as the heterotypic fusion of endocytic vesicles with early endosomes and has been shown to catalyze nucleotide exchange for Rab5 (53,120,154,242). Rabex-5 contains a central region with homology to the yeast protein Vps9p, which was originally identified in a genetic screen for defects in vacuolar protein sorting (44,120). Null and temperature sensitive mutations in the Vps9 gene result in enlarged vacuoles and improper sorting of vacuolar proteins to the extracellular space (44).

The Vps9 Hidden Markov Model (HMM), as defined in the SMART and PFAM databases, is a region of approximately 120 residues that is conserved in at least 18 modular mammalian proteins, including Rabex-5, the RIN (Ras Interaction/Interference) family of Ras effectors, and the ALS2 (Amyotrophic Lateral Sclerosis Type 2) protein (312,313). Vps9p, the RIN proteins, and the ALS2 protein have been shown to posses GEF activity for Rab5 or the yeast homolog Vps21/Ypt51 (45,314-317). The RIN proteins co-localize with Rab5 and contain high affinity Ras association domains that facilitate allosteric regulation of Rab5 nucleotide exchange activity by GTP-bound Ras (317-319). RIN1 knockout mice exhibit markedly enhanced long term potentiation in the amygdala and increased formation of aversive memories (320). RIN1 also associates with the BCR-ABL (Breakpoint Cluster Region-Ableson Kinase) fusion protein, potentiates the oncogenic activity of BCR-ABL in hematopoietic cells, and accelerates BCR-ABL induced leukemia in mice (318,321). The human ALS2 protein contains an N-terminal RCC1-like domain, a putative Dbl homology-pleckstrin homology (DH-PH) tandem, and a series of MORN repeats in addition to a C-terminal Vps9 domain. A

number of autosomal recessive single nucleotide polymorphisms (SNPs) in the ALS2 protein have been identified in individuals with a rare juvenile onset form of ALS and in individuals with a clinically related syndrome referred to as Infantile Ascending Hereditary Spastic Paralysis (IHASP). The known SNPs, including one in the Vps9 domain, are either single base deletions or non-sense mutations, resulting in truncation of the ALS2 protein (322-326). When over-expressed in cultured cells, the truncated ALS2 proteins are targeted for rapid degradation by the proteosome (327).

Although the exchange activity of full length Rabex-5 appears to be specific for Rab5, it is weak compared with that observed for other well characterized GEFs (314). The exchange activity of Rabex-5 is enhanced in the complex with Rabaptin-5, which does not possess intrinsic exchange activity (53). Whether the weak exchange activity of full length Rabex-5 reflects an auto-inhibited state or is an inherent property of the catalytic domain remains to be investigated. Furthermore, it is known that some GEFs, notably those for Arf and Rho GTPases, have overlapping specificities for one or more members of their respective GTPase families (328-330). However, the specificity of mammalian Rab GEFs, including those with Vps9 domains, has not been systematically addressed, owing to the large size and phylogenetic complexity of the Rab family.

To identity Mss4 recognition determinants, eight of the residues selectively conserved in exocytic Rab GTPases were examined in a structure based mutational analysis of the interaction between Mss4 and Rab3A. The mutants do not show significant effects on the rate of intrinsic nucleotide release. Five mutants exhibited Mss4 catalyzed release activities comparable to or greater than the wild type protein whereas

three residues were found to be critical determinants of the interaction with Mss4. Two of the critical recognition determinants lie within a consensus GEF interaction epitope derived from the crystal structures of four mammalian GEF-GTPase complexes whereas the residues that exhibit the least significant defects when mutated lie outside the consensus epitope.

In addition, as a first step towards understanding the structural characteristics, catalytic properties, and global Rab specificity of Vps9 domains, we have identified a central 260 amino acid fragment of Rabex-5 with robust GEF activity for Rab5. The crystal structure of this fragment reveals a novel fold for the SMART/PFAM Vps9 homology region, which corresponds to a sub-domain within the context of an integrated tandem domain architecture. Acidic, hydrophobic, and polar residues on a conserved surface of the Vps9 domain are shown to be critical for exchange activity. A quantitative family-wide analysis of Rab specificity demonstrates that the central catalytic fragment of Rabex-5 has selective exchange activity for Rab GTPases of the Rab5 sub-family.

Experimental Procedures

Constructs All constructs were amplified with Vent polymerase (New England Biolabs) and sequenced from both ends to ensure the absence of secondary mutations. Human Rabex-5 constructs were sub-cloned into a modified pET15b vector containing an N-terminal His_6 tag (MGHHHHHHGS). Mammalian Rab GTPases were subcloned into the PGEX-4T1 vector (Pharmacia), which provides an N-terminal GST tag followed by a thrombin cleavage site. Rab8, Rab10, and Rab17, which could not be expressed in a soluble form as fusions with GST, were sub-cloned into the pET44a vector (Novagen), which incorporates an N-terminal fusion containing a 6xHis tag, the E. *coli* NusA protein, and a thrombin cleavage site. Site-specific mutants were generated using the Quick Change Site-directed Mutagenesis kit (Stratagene).

Expression and Purification BL21(DE3)-RIL cells (Stratagene) transformed with the modified pET15b vectors containing Rab or Rabex-5 constructs were grown at 20°C in 2xYT-amp (16 g of Bactotryptone, 10 g of Bactoyeast extract, 5 g of sodium chloride, and 100 mg of ampicillin per liter) to an OD_{600} of 0.2 and induced with 0.05 mM IPTG overnight. For purification of wild type and mutant proteins, cells were resuspended in lysis buffer (50 mM Tris-HCl, 50 mM NaCl, 0.1% mercaptoethanol), disrupted by sonication, and centrifuged for 40 min at 35000g. For 6xHis fusion proteins, clarified supernatants were loaded onto Ni-NTA agarose columns (Qiagen) and washed with 10 column volumes of Wash Buffer (50 mM Tris-HCl, pH 8.5, 500 mM Nacl, 10 mM imidazole, 0.1% mercaptoethanol). The 6xHis fusion proteins were eluted with a

gradient of 10-150 mM imidazole in Elution Buffer (50 mM Tris-HCl, pH 8.5, 150 mM NaCl, 0.1% mercaptoethanol). For GST fusion proteins, the supernatants were loaded onto glutathione-sepharose columns (Amersham) and washed with 10 column volumes of Wash Buffer. GST fusion proteins were eluted with 10 mM glutathione in Elution Buffer. For Rab protein purifications, all buffers were supplemented with 0.5 mM $MgCl_2$. Wild type Rabex-5 constructs were further purified by anion exchange chromatography over Source Q (Amersham) and gel filtration chromatography over Superdex-75 (Amersham).

Crystallization and Structure Determination Crystals of a selenomethionine-substituted Rabex-5 construct (residues 132-394) were grown at 4°C in hanging drops containing 12 mg/ml protein in 30% PEG 8000, 450 mM $MgCl_2$, and 50 mM Tris pH 8.0. Crystals appeared in 2-3 days and grew to maximum dimensions of 0.05 x 0.1 x 0.2 mm over 2 weeks. The crystals are in the primitive orthorhombic space group $P2_12_12$ with unit cell dimensions a = 47.4 Å, b = 68.7 Å, c= 89.6 Å. The volume of the unit cell is consistent with one molecule in the asymmetric unit and a solvent content of 43%. Prior to data collection, crystals were transferred briefly to a cryo-stabilizer solution (30% PEG 8000, 50 mM Tris, pH 8.0, 450 mM $MgCl_2$, 10% glycerol) and flash frozen in liquid propane. The structure was solved by multiwavelength anomalous diffraction (MAD) at the selenium edge. Data were collected at the X12B beam line at the Brookhaven National Synchrotron Light Source using an inverse beam strategy to maximize the completeness and redundancy of Friedel pairs. Throughout data collection, the crystals

were maintained at 100°K. Data were collected on two isomorphous crystals at four wavelengths near the selenium edge: the f" maximum, the f' minimum, and high as well as low energy remote wavelengths (Table 6). Data were processed with Denzo and scaled with Scalepack (247). Nine Se sites were identified by Patterson and direct methods (SHELXS) using the Bijvoet differences in data collected at the f" maximum. The heavy atom model was refined against a maximum likelihood target function using SHARP (290). Following phase-improvement by solvent-flipping with Solomon, a σA-weighted Fourier summation yielded an interpretable map with continuous main chain density and clearly defined density for most side chains (291). An initial model was constructed using ArpWarp and completed by manual model building in O (291,292). The model was refined against data from 20 to 2.35 Å using simulated annealing and positional refinement in CNS and Refmac5 (291,331). The refined model, which includes residues 134-387, 81 water molecules, and an ordered Mg^{2+} ion, has an R factor of 22.6% and a free R factor of 27.4% with excellent stereochemistry. The Mg^{2+} ion is coordinated by non-conserved residues and does not have an obvious structural or functional significance. Structural figures were generated with PyMol (http://www.pymol.org).

Nucleotide Exchange Assays The kinetics of nucleotide exchange were measured by monitoring either the quenching of fluorescence following the release of the nucleotide analog 2'-(3')-bis-O-(N-methylanthraniloyl)-GDP (mant-GDP, Molecular Probes) or the decrease in intrinsic tryptophan fluorescence accompanying conversion to the active state. For mant-GDP assays, Rab proteins were loaded with the mant-GDP as

described (304) and diluted to 1.0 μM in 50 mM Tris, pH 8.0, 150 mM NaCl, and 0.5 mM $MgCl_2$. Samples were excited at 360 nm and the emission monitored at 440 nm. For intrinsic tryptophan fluorescence measurements, Rab5 was diluted to 1.0 μM in 20 mM Tris-HCl, pH 8.0, 150 mM NaCl, and 0.5 mM $MgCl_2$ and the emission monitored at 340 nm with excitation at 300 nm. For both intrinsic tryptophan and mant-GDP assays, nucleotide exchange reactions were initiated by addition of 200 μM GppNHp and varying concentrations of Rabex-$5_{132-391}$. Data were collected using a Saphire multimode microplate spectrophotometer (Tecan) or a PC1 spectrofluorimeter (ISS). Observed pseudo-first order rate constants (k_{obs}) at each concentration of Rabex-5 were extracted from a nonlinear least-squares fit to the exponential function

$$I(t) = (I_0 - I_\infty) \exp(-k_{obs}\, t) + I_\infty$$

where I(t) represents the emission intensity as a function of time and I_0 and I_∞ represent the emission intensities at $t = 0$ and $t = \infty$, respectively. The catalytic efficiency, k_{cat}/K_m, was obtained from the slope of a linear least-squares fit to

$$k_{obs} = (k_{cat}/K_m)\,[\text{Rabex-5}] + k_{intr}$$

where k_{intr} is the intrinsic rate constant for GDP-release in the absence of Rabex-5.

Results

Determinants of Rab GTPase Recognition

A structure based alignment of representative Rab sequences (Figure 27) reveals a subset of residues that are highly conserved in exocytic Rab GTPases but variable in the endocytic subfamily. Most of these residues fall within the N-terminal third of Rab3A (residues 1-102) previously implicated in the interaction with Mss4 and cluster in or near the P-loop and switch regions (305). Mutational and crystallographic studies of the interactions between other monomeric GTPases and their cognate GEFs have highlighted the importance of residues in the P-loop and switch regions (114-117,332). The majority of the residues selectively conserved in exocytic Rab GTPases occupy exposed or partially exposed positions in the structure of GppNHp-bound Rab3A (18). It is likely that at least some of the residues conserved in exocytic Rab GTPases contribute directly to the interaction with Mss4. In order to identify determinants of the broad recognition of exocytic Rab GTPases by Mss4, eight of the selectively conserved residues (Ser31, Asp45, Phe51, Val52, Ser53, Lys60, Val61, and Thr89) were targeted for mutational analysis. To avoid complex phenotypes, the selected residues were converted to alanine, which occurs naturally at each position in at least one Rab GTPase and is therefore unlikely to disrupt the structure of the protein. Consistent with this expectation, all mutants expressed in soluble form at wild type levels and exhibited no observable defects during purification, with the exception of the D45A mutant, which formed higher order oligomers as detected by gel filtration.

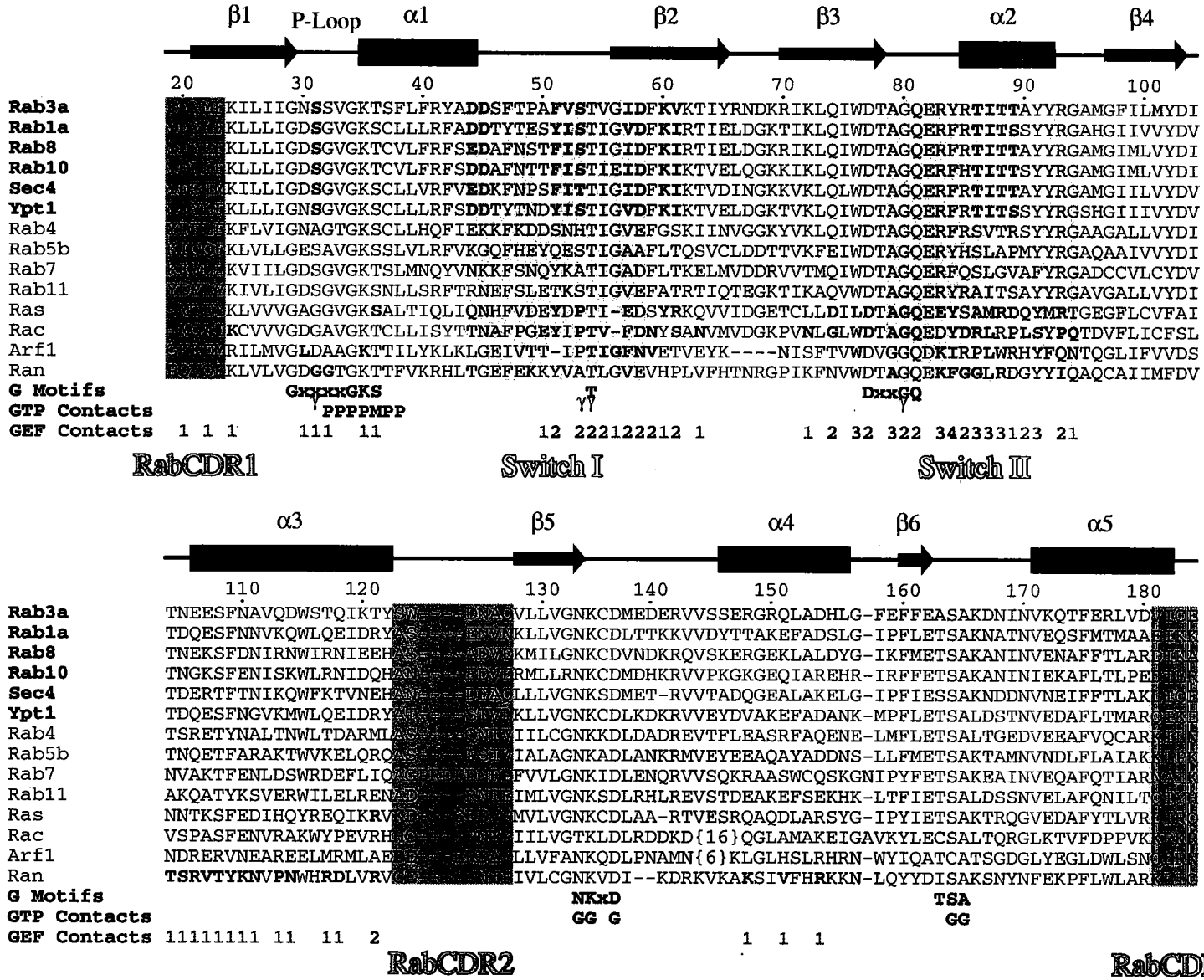

β1
P-Loop
α1
β2
β3
α2
β4
20 30 40 50 60 70 80 90 100
Rab3a KILIIGNSSVGKTSFLFRYADDSFTPAFVSTVGIDFKVKTIYRNDKRIKLQIWDTAGQERYRTITTAYYRGAMGFILMYDI
Rab1a KLLLIGDSGVGKSCLLLRFADDTYTESYISTIGVDFKIRTIELDGKTIKLQIWDTAGQERFRTITSSYYRGAHGIIVVYDV
Rab8 KLLLIGDSGVGKTCVLFRFSEDAFNSTFISTIGIDFKIRTIELDGKRIKLQIWDTAGQERFRTITTAYYRGAMGIMLVYDI
Rab10 KLLLIGDSGVGKTCVLFRFSDDAFNTTFISTIEIDFKIKTVELQGKKIKLQIWDTAGQERFHTITTSYYRGAMGIMLVYDI
Sec4 KILLIGDSGVGKSCLLVRFVEDKFNPSFITTIGIDFKIKTVDINGKKVKLQLWDTAGQERFRTITTAYYRGAMGIILVYDV
Ypt1 KLLLIGNSGVGKSCLLLRFSDDTYTNDYISTIGVDFKIKTVELDGKTVKLQIWDTAGQERFRTITSSYYRGSHGIIIVYDV
Rab4 KFLVIGNAGTGKSCLLHQFIEKKFKDDSNHTIGVEFGSKIINVGGKYVKLQIWDTAGQERFRSVTRSYYRGAAGALLVYDI
Rab5b KLVLLGESAVGKSSLVLRFVKGQFHEYQESTIGAAFLTQSVCLDDTTVKFEIWDTAGQERYHSLAPMYYRGAQAAIVVYDI
Rab7 KVIILGDSGVGKTSLMNQYVNKKFSNQYKATIGADFLTKELMVDDRVVTMQIWDTAGQERFQSLGVAFYRGADCCVLCYDV
Rab11 KIVLIGDSGVGKSNLLSRFTRNEFSLETKSTIGVEFATRTIQTEGKTIKAQVWDTAGQERYRAITSAYYRGAVGALLVYDI
Ras KLVVVGAGGVGKSALTIQLIQNHFVDEYDPTI-EDSYRKQVVIDGETCLLDILDTAGQEEYSAMRDQYMRTGEGFLCVFAI
Rac KCVVVGDGAVGKTCLLISYTTNAFPGEYIPTV-FDNYSANVMVDGKPVNLGLWDTAGQEDYDRLRPLSYPQTDVFLICFSL
Arf1 RILMVGLDAAGKTTILYKLKLGEIVTT-IPTIGFNVETVEYK----NISFTVWDVGGQDKIRPLWRHYFQNTQGLIFVVDS
Ran KLVLVGDGGTGKTTFVKRHLTGEFEKKYVATLGVEVHPLVFHTNRGPIKFNVWDTAGQEKFGGLRDGYYIQAQCAIIMFDV
G Motifs GxxpxxGKS T DxxGQ
GTP Contacts PPPPMPP
GEF Contacts 1 1 1 111 11 12 222122212 1 1 2 32 322 342333123 21
RabCDR1
Switch I
Switch II
α3 β5 α4 β6 α5
110 120 130 140 150 160 170 180
Rab3a TNEESFNAVQDWSTQIKTYSWDNAQVLLVGNKCDMEDERVVSSERGRQLADHLG-FEFFEASAKDNINVKQTFERLVDIICE
Rab1a TDQESFNNVKQWLQEIDRYASENVNKLLVGNKCDLTTKKVVDYTTAKEFADSLG-IPFLETSAKNATNVEQSFMTMAAEIKK
Rab8 TNEKSFDNIRNWIRNIEEHASADVEKMILGNKCDVNDKRQVSKERGEKLALDYG-IKFMETSAKANINVENAFFTLARDIKA
Rab10 TNGKSFENISKWLRNIDQHANETVCRMLLRNKCDMDHKRVVPKGKGEQIAREHR-IRFFETSAKANINIEKAFLTLPELIKR
Sec4 TDERTFTNIKQWFKTVNEHANDEACLLLVGNKSDMET-RVVTADQGEALAKELG-IPFIESSAKNDDNVNEIFFTLAKLIQE
Ypt1 TDQESFNGVKMWLQEIDRYATENVNKLLVGNKCDLKDKRVVEYDVAKEFADANK-MPFLETSALDSTNVEDAFLTMARQIKE
Rab4 TSRETYNALTNWLTDARMLASQNIVIILCGNKKDLDADREVTFLEASRFAQENE-LMFLETSALTGEDVEEAFVQCARKILN
Rab5b TNQETFARAKTWVKELQRQASPSIIALAGNKADLANKRMVEYEEAQAYADDNS-LLFMETSAKTAMNVNDLFLAIAKKLPK
Rab7 NVAKTFENLDSWRDEFLIQASPSDPENFVVLGNKIDLENQRVVSQKRAASWCQSKGNIPYFETSAKEAINVEQAFQTIARNAL
Rab11 AKQATYKSVERWILELRENADPSNPETIMLVGNKSDLRHLREVSTDEAKEFSEKHK-LTFIETSALDSSNVELAFQNILTEIF
Ras NNTKSFEDIHQYREQIKRVKDSDDVPMVLVGNKCDLAA-RTVESRQAQDLARSYG-IPYIETSAKTRQGVEDAFYTLVREIRQ
Rac VSPASFENVRAKWYPEVRHHCPSTPIILVGTKLDLRDDKD{16}QGLAMAKEIGAVKYLECSALTQRGLKTVFDPPVKVVE
Arf1 NDRERVNEAREELMRMLAEDELRDAVLLVFANKQDLPNAMN{6}KLGLHSLRHRN-WYIQATCATSGDGLYEGLDWLSNQLS
Ran TSRVTYKNVPNWHRDLVRVCENIPIVLCGNKVDI--KDRKVKAKSIVFHRKKN-LQYYDISAKSNYNFEKPFLWLARKLI
G Motifs NKxD TSA
GTP Contacts GG G GG
GEF Contacts 11111111 11 11 2 1 1 1
RabCDR2
RabCDR3

Figure 27. Structure based sequence alignment of representative Rab family and related GTPases. Exocytic Rab proteins are indicated by bold face type. Residues that are highly conserved in the exocytic Rab subfamily but variable in endocytic Rab GTPases are highlighted in red. Residues highlighted in blue participate in direct GEF contacts in the crystal structures of Ras-Sos, Arf1-Sec7, Rac-Tiam1, and Ran-RCC1 (114-117). The number of GEF contacts at each position is indicted below the alignment.

The distal spatial relationship of the three critical recognition determinants provides insight into the structural elements of exocytic Rab GTPases that mediate the interaction with Mss4. As shown in Figure 28, Phe51 lies at the N-terminus of the Switch I region, near the phosphate and Mg^{2+} binding site, whereas Val61 is located in the β2 strand, just beyond the C-terminus of the switch I region while Thr89 resides at the C-terminus of the α2 helix in the switch II region. Although Phe51 and Val61 are completely exposed, Thr89 is only partially exposed in the GTP-bound form. Structural data are not available for the GDP-bound form of Rab3A; however, the corresponding residues in the GDP-bound form of Sec4 (the Rab3A homologue in budding yeast) are either exposed or disordered (88). Consequently, the simplest interpretation is that Phe51, Val61 and possibly Thr89 contribute directly to the interface with Mss4, which presumably explains (at least in part) their conservation in exocytic but not endocytic Rab GTPases. Interestingly, Val61 and Thr89 flank an invariant triad of partially exposed aromatic residues discussed previously (Phe59, Trp76 and Tyr91), located at the hydrophobic interface between the switch I and II regions.

Crystal structures of nucleotide free EF-Tu, Ras, Arf1, Rac and Ran have been solved in complex with their cognate GEFs (114-117,332). Despite the lack of structural homology between the various GEFs and significant differences in the interaction interfaces, some general principles have emerged. In all cases, the GEFs contact the switch II region as well as the P-loop and/or switch I regions, resulting in structural changes that disrupt interactions with the phosphates and Mg^{2+} ion. The extent of the overlap in the various interaction interfaces can be represented in the form of a 'consensus

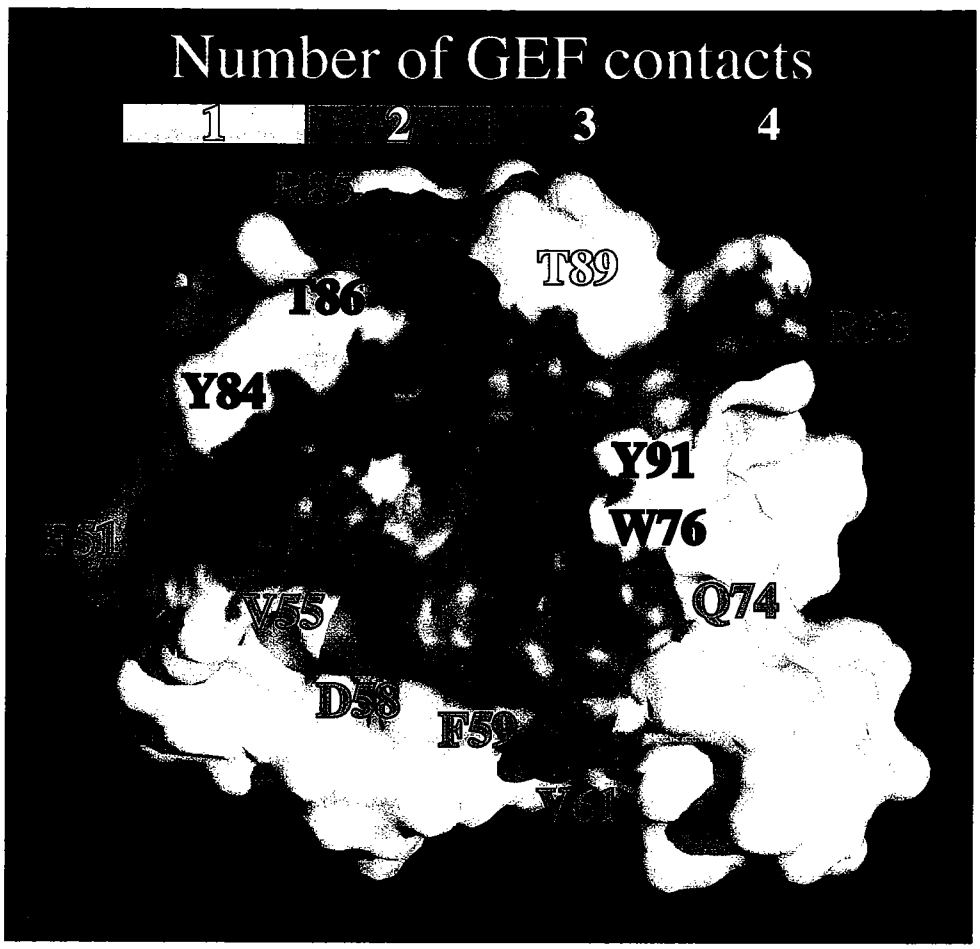


Figure 28. Distribution of GTPase/GEF recognition determinants with respect to the crystal structure of GppNHp bound Rab3A (90). Extent of overlap in GEF interaction epitopes deduced from the crystal structures of Ras-Sos, Arf1-Sec7, Rac-Tiam1, and Ran-RCC1 (114-117) and mapped onto the surface of Rab3A.

epitope' obtained by assigning each position in an alignment of GTPase sequences a score equivalent to the number of known structures in which the corresponding residue lies within the GEF interface. In Figure 28, the consensus epitope derived from the four mammalian GTPase-GEF complexes of known structure (see Figure 27) is mapped onto the surface of the Rab3A structure. For the numerous nucleotide free Rab-GEF complexes whose structures have not been determined, I propose that the consensus epitope can be interpreted as a map of the likelihood that any given residue in a Rab GTPase will lie within the GEF interaction interface. Interestingly, two of the three critical Mss4 recognition determinants identified by mutational analysis (Phe51 and Val61) reside within the GEF consensus epitope whereas the three residues that exhibited the weakest effects when mutated (Ser31, Val52, and Lys60) lie outside the GEF consensus epitope.

Vps9 stimulated Rab exchange

A fragment of Rabex-5 containing the Vps9 domain has high exchange activity. The amino acid sequence of Rabex-5 encodes a modular architecture with five distinctive regions: an N-terminal Cys_4 type Zn^{2+} finger, a glycine and serine rich region, a central region of predicted α helical secondary structure that includes the Vps9 homology domain, a short stretch of heptad repeat, and a proline rich C-terminus. This modular architecture is evolutionarily conserved in Rabex-5 homologs from diverse multicellular organisms. Regions with weak but potentially significant similarity flanking the SMART/PFAM Vps9 homology domain were detected in a manual alignment of twenty

amino acid sequences from a variety of eukaryotic proteins with Vps9 domains (Figure 29 and Appendix 1). The additional homologous regions coincide with five predicted helices at the N-terminus and one at the C-terminus.

A construct consisting of residues 132-391, which includes the five predicted helices N-terminal to the Vps9 domain as well as the additional helix on the C-terminal side, expresses at high levels, behaves as a uniform monomer over a broad concentration range, and exhibits potent GEF activity for Rab5 (Figure 30A). The observed rate constant for the exchange reaction (k_{obs}) depends linearly on the concentration of Rabex-$5_{132-391}$ to at least 2 μM (the highest concentration tested). This observation places lower limits on the values of the apparent Michaelis-Menten kinetic constants such that $k_{cat} > 0.1\ s^{-1}$ and $K_m > 2\ \mu M$. Constructs that truncate one or more of the predicted helices at the N- or C-terminus express poorly and/or are prone to aggregation when purified. Consistent with this observation, Rabex-$5_{132-391}$ is resistant to proteolysis by LysC, GluC (V8 protease), ArgC, and chymotrypsin (data not shown). We therefore conclude that a central 260 amino acid region of Rabex-5, which extends significantly beyond the 120 aa SMART/PFAM Vps9 homology domain, represents the core catalytic unit. This conclusion is further supported by the crystal structure and the results of a mutational analysis of conserved residues (see below).

Tandem domain architecture of the catalytically active region of Rabex-5. The crystal structure of selenomethionine-substituted Rabex-$5_{132-394}$ was solved by multiwavelength anomalous diffraction and refined to 2.35Å (Table 6 and Figure 30B). As illustrated in Figure 30C, the Rabex-5 catalytic core consists of two distinct domains:

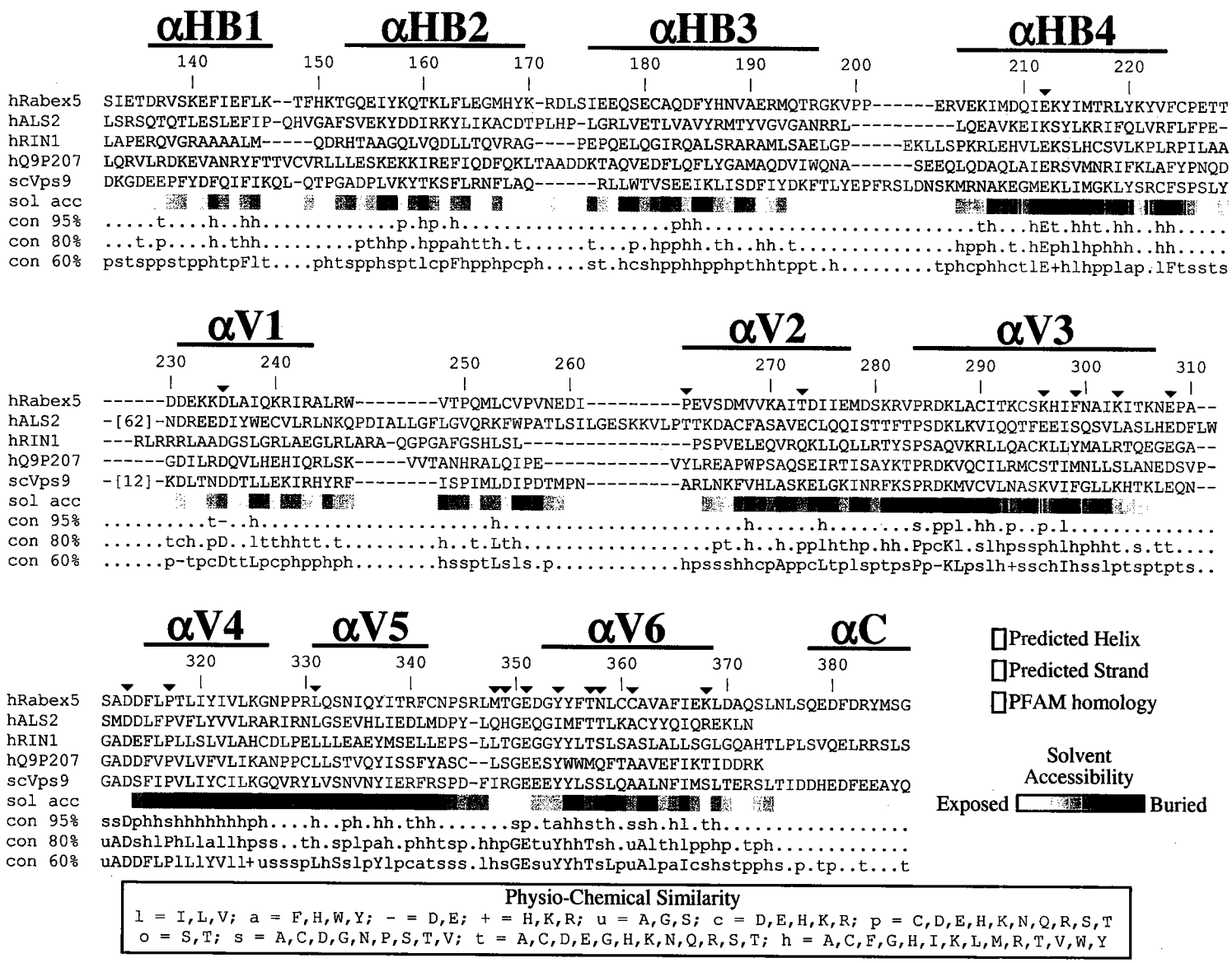

αHB1
αHB2
αHB3
αHB4
αV1
αV2
αV3
αV4
αV5
αV6
αC
hRabex5
hALS2
hRIN1
hQ9P207
scVps9
sol acc
con 95%
con 80%
con 60%
Predicted Helix
Predicted Strand
PFAM homology
Solvent Accessibility
Exposed
Buried
Physio-Chemical Similarity
l = I,L,V; a = F,H,W,Y; - = D,E; + = H,K,R; u = A,G,S; c = D,E,H,K,R; p = C,D,E,H,K,N,Q,R,S,T
o = S,T; s = A,C,D,G,N,P,S,T,V; t = A,C,D,E,G,H,K,N,Q,R,S,T; h = A,C,F,G,H,I,K,L,M,R,T,V,W,Y

Figure 29. Structure based sequence alignment of representative Vps9 domain proteins. Residue numbers and observed secondary structure for Rabex-5 are shown above the alignment. Fractional solvent accessibility (sol acc) for each residue in the HB-Vps9 tandem of hRabex-5 is indicated as a blue box with a continuous gradient scale from exposed (light blue) to buried (dark blue). The consensus of physio-chemical similarity at conservation levels of 60% (con 60%), 80% (con 80%), and 95% (con 95%) is also shown. The consensus of secondary structure prediction for each sequence is indicated by highlighting predicted helices in blue and predicted strands in orange. The region of the SMART/PFAM Vps9 HMM is highlighted in yellow. An annotated alignment of the twenty sequences used to determine the level physio-chemical conservation is included as Appendix 1.

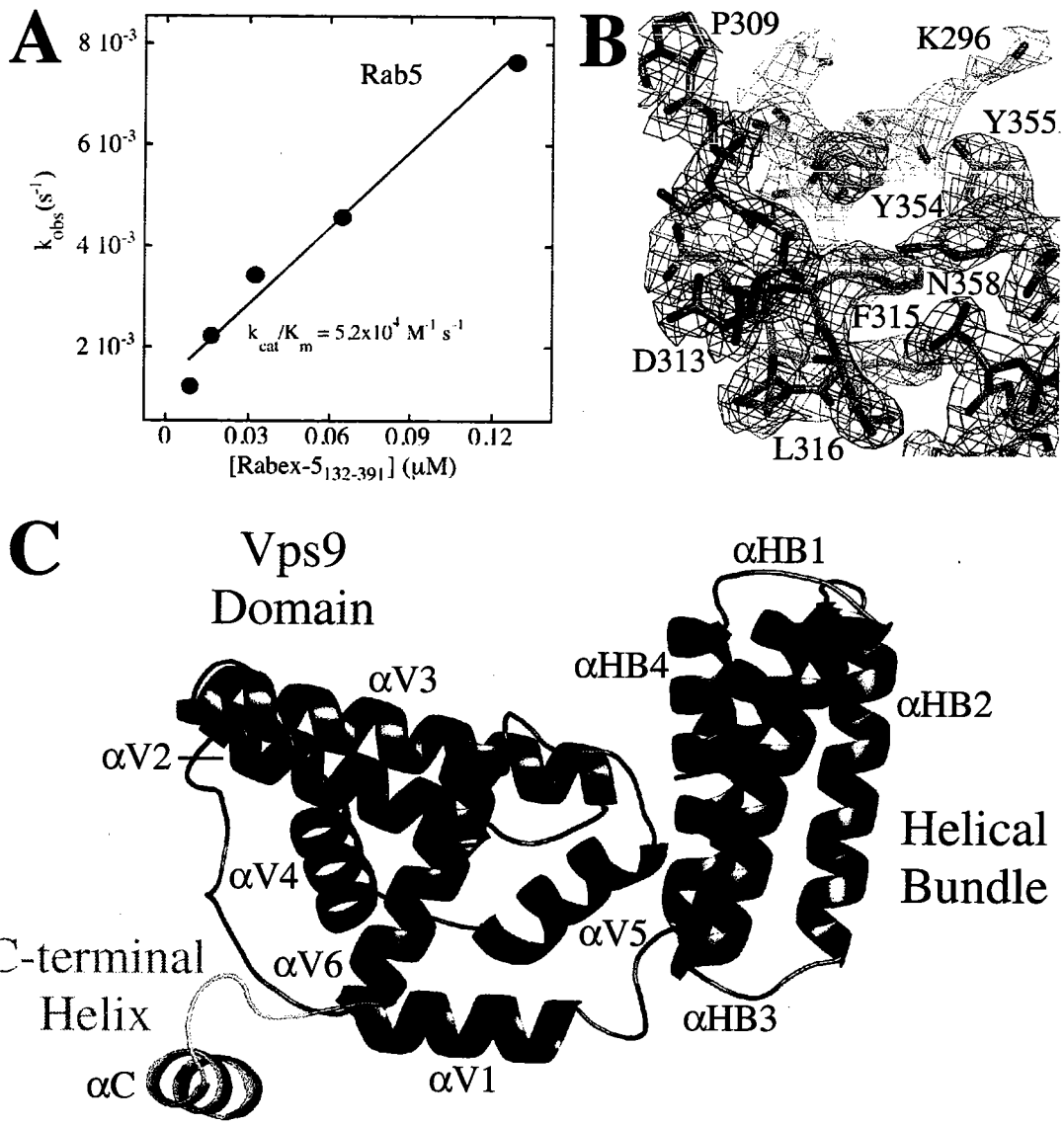


Figure 30. Exchange kinetics and structure of the Rabex-5 catalytic core. **(A)** Observed pseudo-first order rate constant for the exchange of GDP for GppNHp on Rab5 as a function of the concentration of Rabex-5132-391. The large change in the intrinsic tryptophan emission of Rab5 accompanying nucleotide exchange was continuously monitored at 340 nm with excitation at 300 nm. Data were collected and analyzed as described in the Experimental Procedures. **(B)** Representative region of the experimental map following phase improvement by solvent flipping. The density contoured at 1.0 σ is shown with the final refined model. **(C)** Overall fold and domain organization of Rabex-5132-394. The designation of domain boundaries is discussed in the text.

Table 6. Rabex-5 HB-Vps9: structure determination and refinement

Data Collection[a]						
Crystal	Se 1	Se 1	Se 1	Se 2	Se 2	Se 2
Wavelength (Å)	λ1 (0.97956)	λ2 (0.97944)	λ3 (0.96805)	λ1 (0.97954)	λ2 (0.97948)	λ4 (0.99530)
Resolution (Å)	20-2.6	20-2.6	20-2.5	20-2.8	20-2.5	20-2. 35
R_{sym} (%)[b]	7.5 (35.6)	6.5 (42.7)	7.0 (31.0)	6.9 (33.5)	6.9 (35.1)	5.9 (37.2)
$\langle I/\sigma \rangle$	22 (2.7)	25 (2.8)	20.3 (2.9)	24.5 (3.9)	28.4 (3.4)	23.3 (2.0)
Completeness (%)	98.1 (94.9)	97.8 (91.8)	97.0 (94.0)	100 (100)	99.9 (100)	96.3 (74.5)
Redundancy	4	4	4	4	4	4

Phasing Power and FOM (Bijvoet Differences)[a]					
	Se 1 λ1	Se 1 λ2	Se 1 λ3	Se2 λ1	Se2 λ2
Acentric PP	1.4 (0.51)	1.6 (0.52)	1.5 (0.47)	1.5 (0.54)	1.5 (0.48)

Phasing Power and FOM (Dispersive Differences)[a]					
	Se 1 λ1 vs. Se 2 λ4	Se 1 λ2 vs. Se 2 λ4	Se 1 λ3 vs. Se 2 λ4	Se2 λ1 vs. Se 2 λ4	Se2 λ2 vs. Se 2 λ4
Acentric PP	1.3 (0.88)	1.3 (0.90)	0.48 (0.32)	1.7 (1.2)	1.9 (0.87)
Centric PP	1.04 (0.65)	1.04 (0.72)	0.4 (0.25)	1.2 (0.74)	1.4 (0.76)
Acentic FOM	0.59 (0.16)				
Centric FOM	0.53 (0.19)				

Refinement				
			RMS Deviations	
Resolution (Å)	R Factor (%)	R_{free}[c] (%)	Bond Length (Å)	Bond Angle
20-2.35	22.6	27.4	0.006	1.1°

[a]Values in parenthesis represent the highest resolution shell.

[b]$R_{sym} = \Sigma_h \Sigma_j |I_j(h) - \langle I(h) \rangle| / \Sigma_h \Sigma_j I_j(h)$.

[c]R-value for a 5% subset of reflections selected at random and omitted from refinement.

an N-terminal four helix bundle (HB) and the C-terminal Vps9 domain, which has a novel α helical fold that does not resemble other GEFs of known structure. Three of the helices in the HB (αHB2 - αHB4) are arranged in the form of a triple stranded anti-parallel coiled-coil with a right-handed super-helical pitch. The sub-domain corresponding to the SMART/PFAM Vps9 HMM adopts a layered fold in which a helical hairpin formed by αV2 and αV3 supports a middle layer consisting of αV4 and αV5. The N-terminus of αV6 packs against the central portion of αV3, allowing the C-terminus of αV6 to thread a V-shaped groove between αV4 and αV5.

Although the short helix following αV6 packs against the Vps9 domain and was required for soluble expression, two observations indicate that this helix (αC1) is not a generally conserved element of the Vps9 domain: (i) the sequence similarity following the Vps9 domain is weak; and (ii) the location of the Vps9 domain at the C-terminus of the ALS2 and Q9P207 proteins precludes the possibility of an analogous helix. In contrast, significant similarity is evident in the structure-based manual alignment of the sequences N-terminal to the SMART/PFAM HMM (Figure 29 and Appendix 1). This region encodes an amphipathic helix (αV1) followed by an ordered stretch of random coil. The non-polar surface of αV1 rests in a hydrophobic groove formed by the C-terminus of αV6 and the connecting segment between αV4 and αV5. Likewise, non-polar residues in the ordered random coil following αV1 pack against non-polar residues from αV4. Given the overall conservation of physiochemical similarity for the residues that lie within these interfaces, we conclude that the Vps9 homology domain begins with

αV1 and terminates with αV6. This structural definition of the Vps9 domain will be retained in subsequent discussion.

A striking feature of the Rabex-5 structure is an extensive interface between the helical bundle and the Vps9 domain. In this interface, which buries 1940 $Å^2$ of surface area, residues from one side of αHB4 pack in a shallow groove between αV5 and the connecting loop of the αV2/αV3 helical hairpin. Of particular note is a solvent excluded hydrogen bonding interaction between the buried carboxylate group of a highly conserved glutamic acid residue in αHB4 (Glu212) and the main chain NH group of Arg285 at the N-terminus of αV3. This interaction is enhanced by the favorable disposition of the negatively charged side chain of Glu212 with respect to the positive end of the αV3 helix dipole. An analogous solvent excluded hydrogen bonding interaction occurs between the carboxylate group of Asp235 in αV1 and the main chain NH and side chain hydroxyl groups of Ser333 at the N-terminus of αV5. Consistent with this observation, mutation of either acidic residue to alanine results in mutant proteins that do not express in a soluble form whereas alanine substitutions involving exposed acidic residues have no effect on solubility (see below).

The catalytically active region of Rabex-5 has high specificity for two Rab GTPases. To gain insight into the family-wide specificity of Rab effectors and regulatory factors, the Lambright lab has constructed a panel of 37 mammalian Rab GTPases, representing the known Rab family excluding highly homologous isoforms. The majority (32 of the 37 Rab GTPases) expressed in a soluble form with GST fused to the N-terminus. Three Rab GTPases (Rab8, Rab10, and Rab17) that were not soluble as GST-

fusions could be expressed in a soluble form with NusA as an N-terminal fusion partner. Two Rab GTPases (Rab36 and Rab37) could not be expressed in a soluble form as either GST or NusA fusions. The fusion-proteins were purified to >95% homogeneity as determined by SDS-PAGE (see Experimental Procedures). With the exception of Rab10, Rab17, Rab26 and Rab28, all of the soluble Rab GTPases exhibited an intrinsic ability to exchange nucleotide and hydrolyze GTP, indicative of a folded, functional state.

A well established approach for measuring the kinetics of nucleotide release from monomeric GTPases, including Rab GTPases, takes advantage of the fluorescent GDP analog mant-GDP, which has an N-methylanthraniloyl label attached to the 2'- and/or 3'-hydroxyl group of the ribose moiety (304,308,333,334). Release of GDP-mant is accompanied by a large decrease in emission intensity, as the environmentally-sensitive fluorescence of the mant label is efficiently quenched in bulk solution. A particular advantage of this assay in the present context is that it can be readily adapted to microplate format, allowing efficient kinetic screening over a broad range of GEF concentrations.

The ability of Rabex-$5_{132-391}$ to catalyze the release of GDP-mant from 34 Rab GTPases was determined using a quantitative microplate assay. Representative results for 33 Rab GTPases are shown in Figure 31. Under the conditions of these assays, in which the concentration of Rabex-$5_{132-391}$ is varied over a broad range from 20 nM to 1 μM, measurable GEF activity was observed for only three Rab GTPases: Rab5, Rab21, and Rab22. For the Rab GTPases analyzed (excluding Rab10, Rab17, Rab26, and Rab28), addition of 10 mM EDTA, which reduces the free Mg^{2+} concentration to sub-

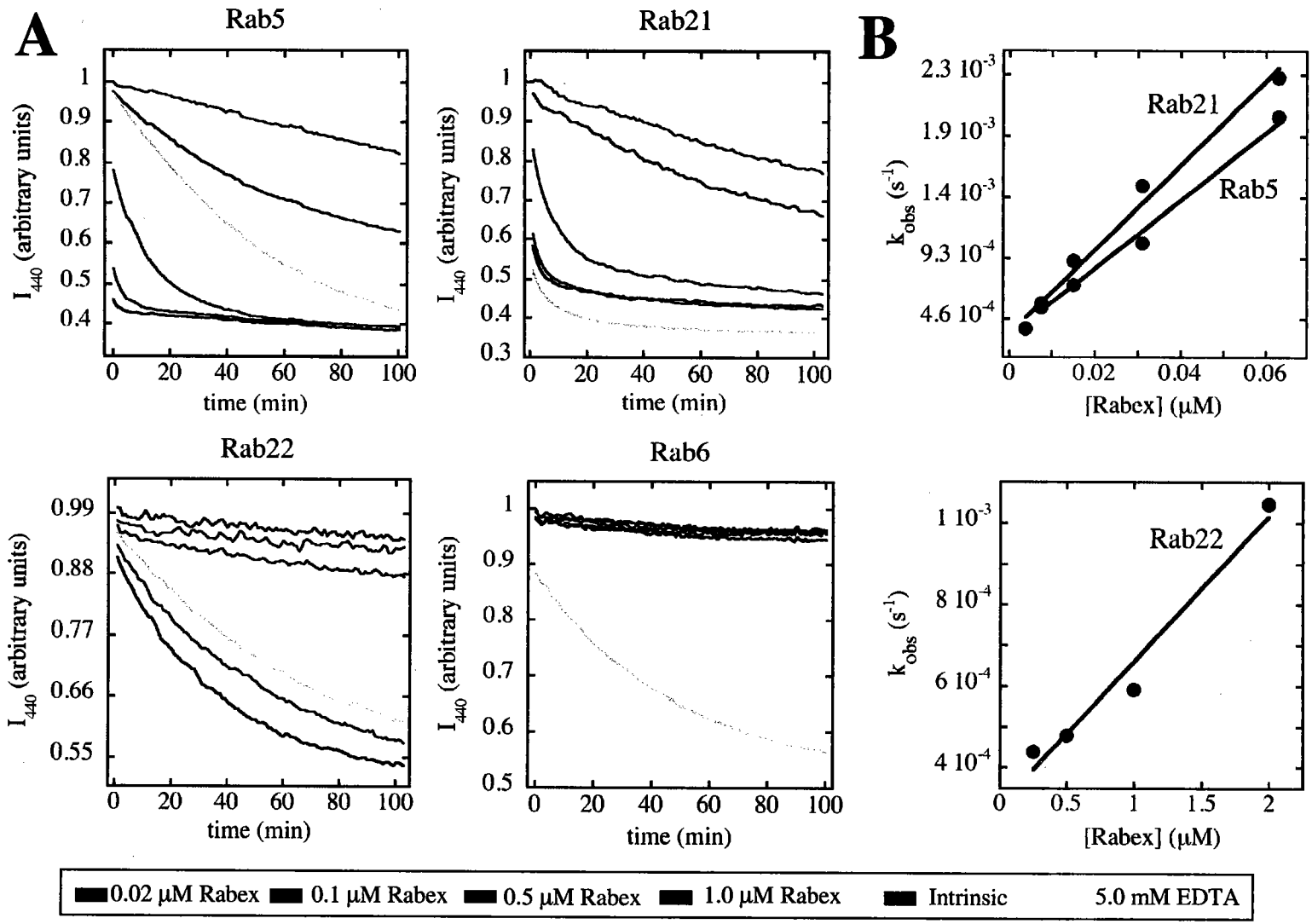

A
Rab5
Rab21
Rab22
Rab6
I_{440} (arbitrary units)
time (min)
B
Rab21
Rab5
Rab22
k_{obs} (s^{-1})
[Rabex] (μM)
0.02 μM Rabex
0.1 μM Rabex
0.5 μM Rabex
1.0 μM Rabex
Intrinsic
5.0 mM EDTA

Figure 31. Family-wide analysis of Rab specificity. (**A**) Kinetics of mant-GDP release from GST-fusions of Rab5, Rab21, Rab22, and Rab6 in the absence (light blue) and presence of 0.02 μM (red), 0.1 μM (dark blue), 0.5 μM (green), and 1 μM (black) Rabex-$5_{132\text{-}391}$ or following addition of 10 mM EDTA (yellow). (**B**) Dependence of the observed pseudo-first order rate constant for mant-GDP release on the concentration of Rabex-$5_{132\text{-}391}$.

micromolar levels, greatly stimulates the rate of nucleotide release, indicating that the absence of detectable GEF activity for the majority of Rab GTPases is not due to an inability to release GDP-mant but instead reflects high intrinsic specificity in the interaction with Rabex-$5_{132\text{-}391}$.

For the three Rab GTPases that exhibited GEF-stimulated nucleotide release in the initial screen, a more detailed set of kinetic experiments was conducted. Whereas the catalytic efficiency (k_{cat}/K_m) of Rabex-$5_{132\text{-}391}$ for Rab5 and Rab21 is indistinguishable within experimental error, it is ~100 fold lower for Rab22. Although a similar kinetic analysis is not possible for the other Rab GTPases examined in this study, the absence of detectable GEF activity implies limits on the relative catalytic efficiency. Accounting for the signal-to-noise of the measurements and the readily detectable though weak GEF activity for Rab22, we conservatively estimate that the k_{cat}/K_m for the majority of other Rab GTPases must be at least several hundred fold lower than that observed for Rab5. Thus, Rabex-$5_{132\text{-}391}$ exhibits equally robust GEF activity for Rab5 and Rab21, weak activity for Rab22, and insignificant or no activity for the 28 other Rab GTPases for which nucleotide release could be effectively stimulated by addition of EDTA.

Residues in a conserved surface of the Vps9 domain are critical for GEF activity. A number of residues that are exposed in the Rabex-$5_{132\text{-}391}$ structure exhibit a high degree of conservation in the two other mammalian proteins (RIN1 and ALS2) that contain Vps9 domains and have been shown to catalyze nucleotide exchange for Rab5. A similar overall pattern of conserved surface residues is evident in the larger alignment of 20 Vps9 domain proteins from diverse organisms. As shown in Figure 32, the

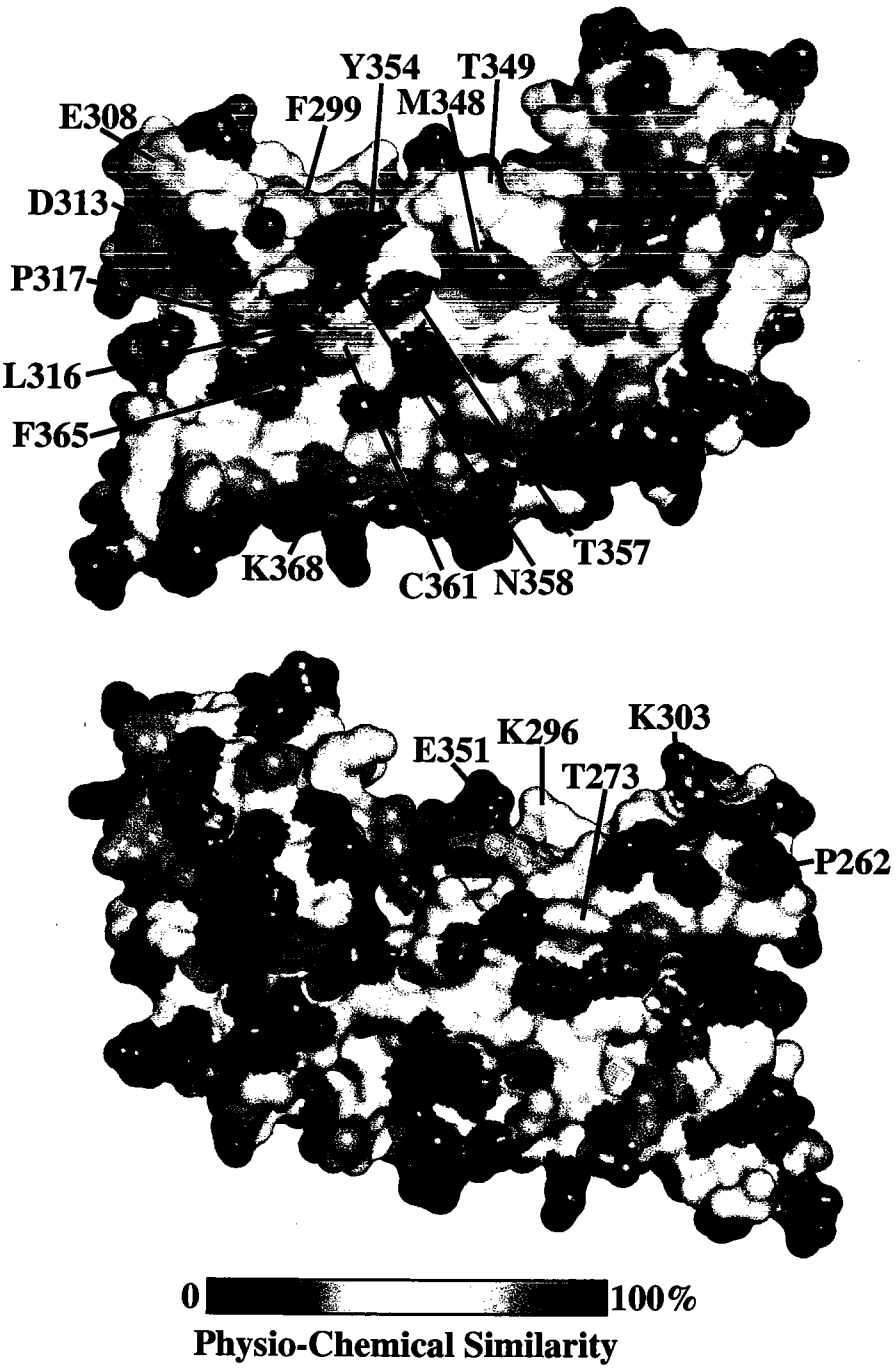


Figure 32. Correlation of solvent exposure with sequence conservation. The conservation of physio-chemical similarity in a sequence alignment of twenty Vps9 domain proteins (see Figure 29) was mapped onto the surface of the Rabex-5$_{132-394}$ structure. The orientation of the top view corresponds to that of Figure 2C. The orientation of the top and bottom views are related by a rotation of 180° about the vertical axis. Note that the only invariant residue in the Vps9 domain (Asp313) is entirely solvent exposed and lies adjacent to a conserved groove comprised primarily of exposed non-polar residues.

majority of conserved, solvent exposed residues cluster on a common surface of the Vps9 domain. Of particular note is an aspartic acid residue (Asp313) in the short, ordered αV4/αV5 loop. The side chain of Asp313, which represents the only invariant residue in the SMART/PFAM Vps9 HMM, is completely solvent exposed and does not engage in intramolecular stabilizing interactions. Intriguingly, Asp313 is situated immediately adjacent to a shallow hydrophobic groove formed by residues from αV4 and αV5. A small number of conserved residues are visible on the opposite surface; however, these residues are generally more buried and are therefore likely to be conserved for structural stability. To identify determinants of Rab GEF activity, conserved or partially conserved residues with substantial solvent exposure were substituted with alanine. All of the mutant proteins expressed in a soluble form at wild type levels and were well behaved in solution. This contrasts with the lack of soluble expression for alanine mutants of the two buried acidic residues, Glu212 and Asp235 that, as discussed above, are predicted to be critical for structural stability. The ability of Rabex-$5_{132\text{-}391}$ mutants to catalyze release of mant-GDP was determined for both Rab5 and Rab21. As shown in Figure 33, mutation of four exposed residues (Asp313, Pro317, Tyr354, and Thr357) that are clustered near the N-termini αV4 and αV6 resulted in severely impaired catalytic efficiency. Mutation of three residues flanking the critical cluster (Thr349, Glu351, and Asn358) resulted in moderate defects whereas mutation of other residues had minimal effect. Notably, the severity of the defects resulting from alanine substitution is strongly correlated with both the conservation of physiochemical similarity and the degree of solvent exposure. Consequently, the simplest interpretation is that the contiguous surface

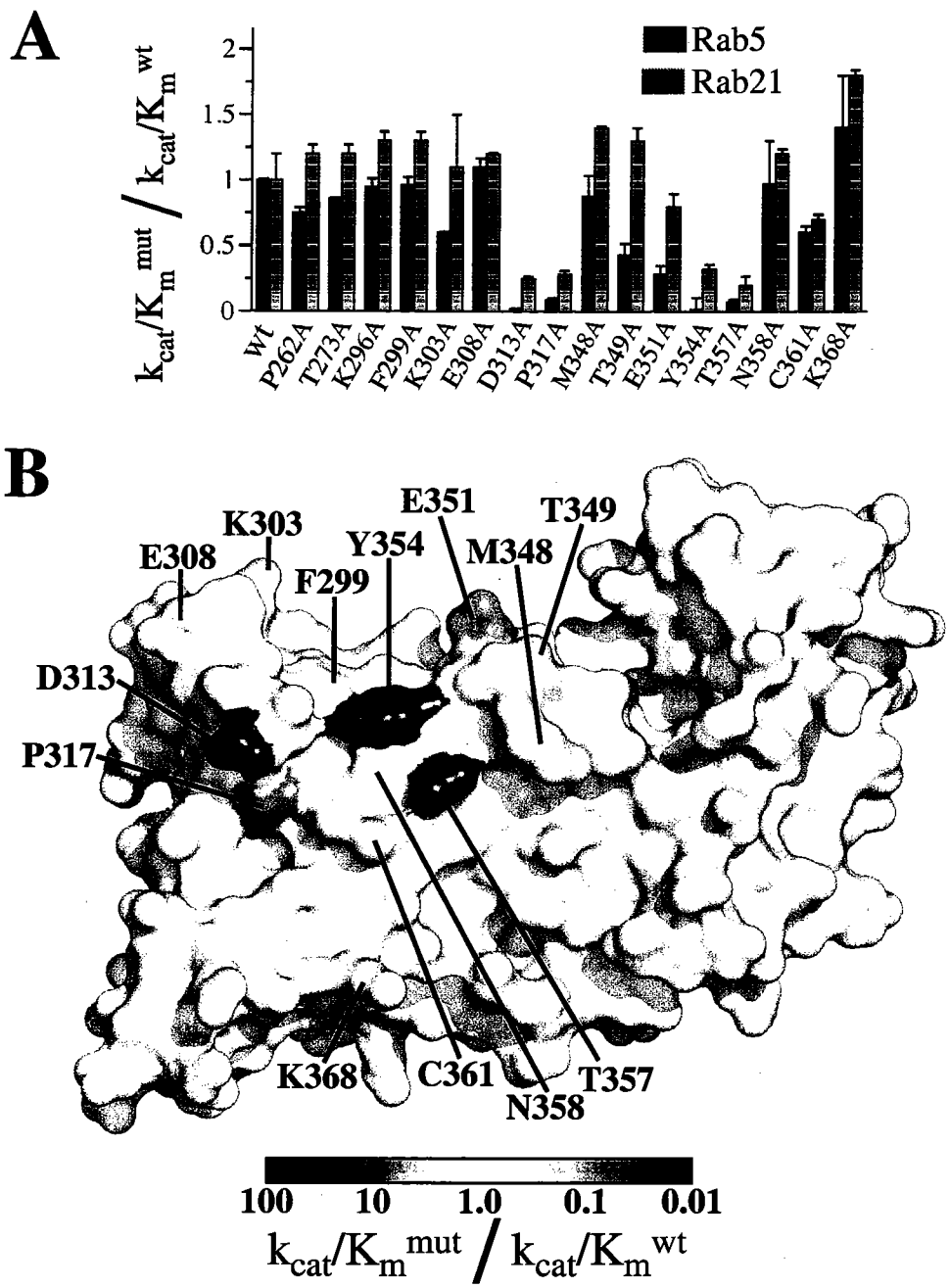


Figure 33. Mutational analysis of the predicted Rab binding surface. **(A)** The catalytic efficiency (k_{cat}/K_m) of mutant (mut) and wild type (wt) Rabex-5$_{132-391}$ was determined for GST fusions of Rab5 and Rab21. To simplify comparison, the results are expressed as the ratio of k_{cat}/K_m for mutant proteins relative to that of the wild type protein. **(B)** Mutated residues are colored according to the relative (k_{cat}/K_m) for mutant and wild type Rabex-5$_{132-391}$ with GST-Rab5 as the substrate. Note that the severity of the defects are strongly correlated with the conservation of physiochemical similarity as well as the degree of solvent exposure (compare with Figure 32).

formed by conserved residues from αV4 and αV6, as well as the loops N-terminal to these helices, corresponds to the overlapping interaction epitopes for Rab5 and Rab21.

Structural similarity in the GTPase binding sites of Vps9 and Sec7 domains. At the level of the overall fold, the Vps9 domain does not exhibit obvious structural similarity to other exchange factors as judged by fold recognition using the DALI server or by visual inspection (335). However, the functionally critical sub-structure of the Vps9 domain, consisting of the αV4 and αV6 helices, bears a striking resemblance to an analogous pair of helices in the Sec7 domain that account for the majority of the interactions with Arf GTPases (Figure 34). In the Sec7-Arf complexes, the αG and αH helices of the Sec7 domain engage the switch regions of Arf such that non-polar residues from the switch I region dock in a hydrophobic groove between the N-terminus of αG and the C-terminus of αH while residues from the switch II region interact primarily with residues from αG (234-236). The 'glutamic acid finger' in the loop preceding αG plays a critical role in promoting GDP release and, in the complex with nucleotide free Arf, intrudes into the nucleotide binding site to interact with the invariant lysine residue in the $GX_4GK(S/T)$ motif that encodes the phosphate binding loop. Like the Sec7 domain, the Vps9 domain also contains a hydrophobic groove at the N-termini of the αV4 and αV6 helices as well as an analogous 'aspartic acid finger' in the loop preceding αV4. As demonstrated by the mutational analysis, both the invariant aspartic acid finger (Asp313) and at least one highly conserved aromatic residue in the hydrophobic groove (Tyr354) are critical determinants of exchange activity. These observations suggest the intriguing possibility that the Vps9 and Sec7 domains might share a similar mode of interaction with their respective GTPases, despite dissimilar overall structures.

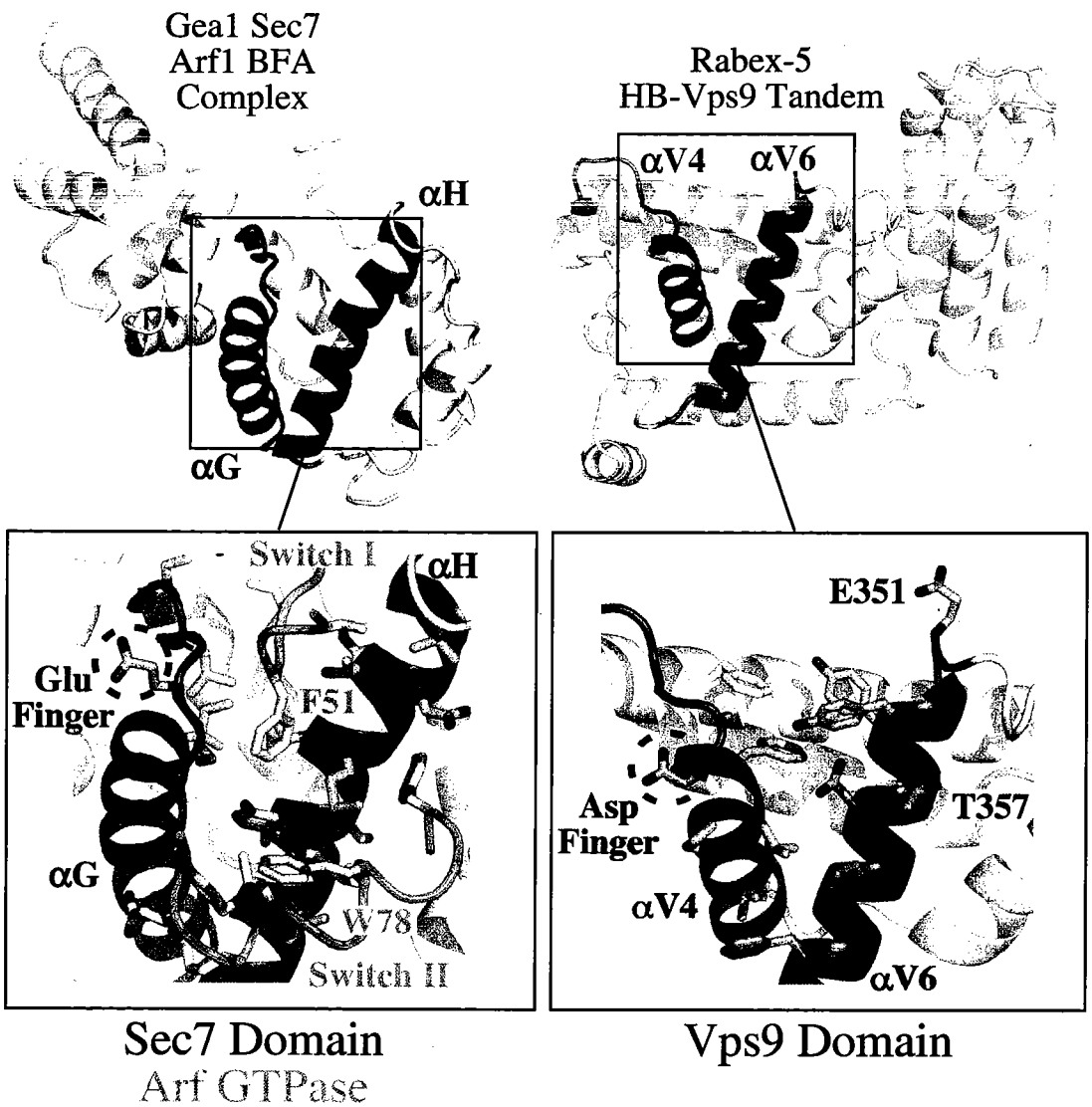


Figure 34. Comparison of the Vps9 and Sec7 domains. The HB-Vps9 tandem of Rabex-5 is compared with the Sec7 domain of Gea1 from the complex with Arf1 GDP and Brefeldin A (234). In the Sec7 domain, a pair of helices and adjacent loop regions (highlighted in blue) account for the majority of contacts with Arf1 (semitransparent orange). An analogous pair of helices and adjacent loops in the Vps9 domain contain all of the conserved residues with significant solvent exposure as well as the known determinants of GEF activity. Critical glutamate (Sec7) and aspartate (Vps9) residues are shown in red.

A SNP in ALS2 disrupts the structure and GEF function of the Vps9 domain. A high degree of physiochemical similarity for residues implicated in structural stability and exchange activity suggest that the crystal structure of the tandem HB-Vps9 domains of Rabex-5 should serve as a useful working model for the corresponding catalytic regions of other Vps9 domain GEFs including the ALS2 and RIN proteins. One of the known SNPs in the ALS2 gene (4844ΔT) corresponds to the deletion of a single base in an exon encoding part of the Vps9 domain of ALS2 (324). The 4844ΔT SNP, which generates a frame shift at the codon for a residue located near the N-terminus of αV3, replaces the C-terminal half of the Vps9 domain with 43 residues of non-native sequence (Figure 35). It is clear that the 4844ΔT SNP severely disrupts the structure of the Vps9 domain, eliminating most if not all of the residues implicated as critical for the interaction with Rab GTPases and exposing many non-polar residues that would otherwise be buried in the hydrophobic core. Furthermore, the frame-shifted sequence following 4844ΔT encodes a high percentage of non-polar residues, which likely exacerbates the severity of the structural defect. In RIN1, a splice variant that deletes the region from the N-terminus of αHB4 to the N-terminus of αV2 fails to interact with Rab5 in a two-hybrid assay. This region contains the two conserved/buried acidic residues which appear to be critical for structural stability and are essential for the soluble expression of the Rabex-5 HB-Vps9 tandem. Thus, the inability of the RIN1 splice variant to interact with Rab5 likely results from an indirect structural defect rather than the loss of critical exchange determinants.

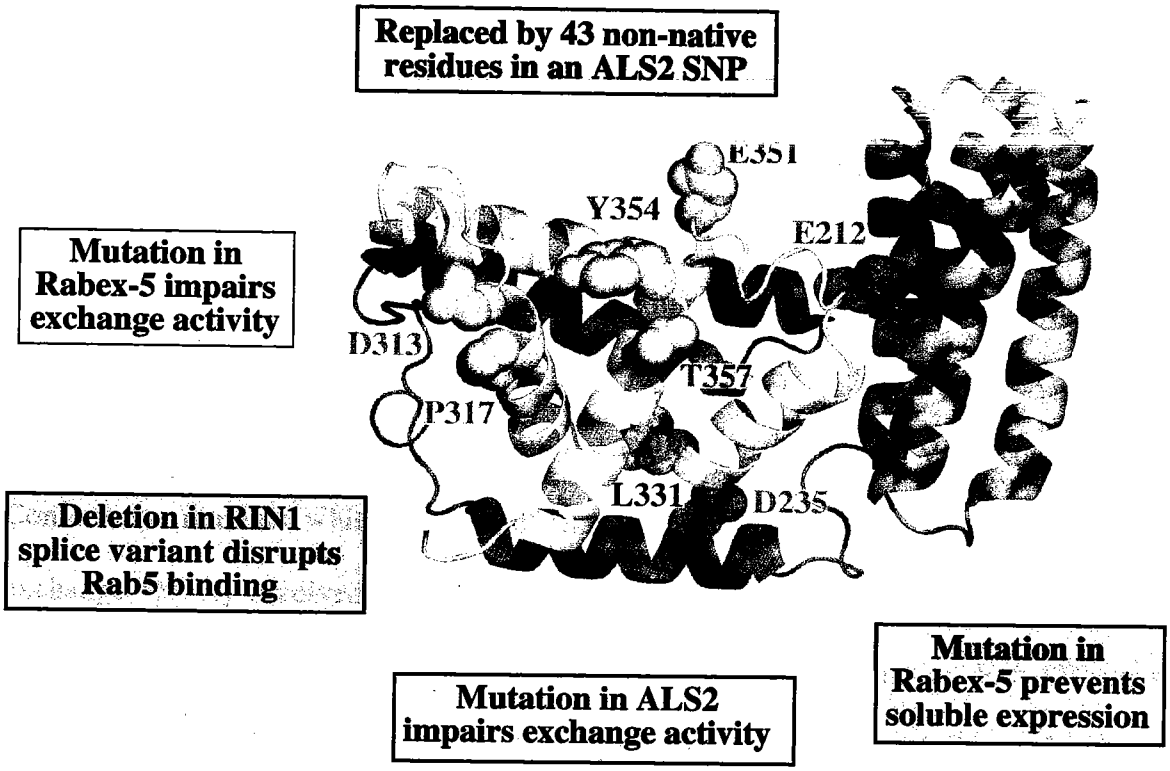


Figure 35. Correlation of structure and function in Vps9 domain proteins. The structure and mutational analysis of the HB-Vps9 tandem of Rabex-5 provides insight into functional observations from published work on the RIN1 and ALS2 proteins. A SNP in the Vps9 domain of the ALS2 protein (yellow) replaces the core of the Vps9 domain, including highly conserved residues that are critical for exchange activity (gray), with 43 residues of non-native sequence (324). A splice variant of RIN1, which deletes a region from αHB4 to the N-terminus of αV1, fails to interact with Rab5 in a two hybrid assay (317). Mutation of buried acidic residues in the corresponding region of the Rabex-5 HB-Vps9 tandem prevents soluble expression, presumably due to a folding defect. Likewise, the loss of exchange activity resulting from mutation of a conserved leucine residue in ALS2 is an indirect consequence of its buried location in the hydrophobic core (316). Conversely, loss of exchange activity resulting from mutation of a conserved proline residue in ALS2 most likely reflects its partially exposed location within the exchange epitope.

Discussion

A consensus GEF interaction epitope derived from crystal structures of distantly related nucleotide free GTPases in complex with non-homologous GEFs provides a convenient framework for assessing the general significance of the mutational data on GEF interactions with Rab GTPases. Two critical determinants for recognition of exocytic Rab GTPases by Mss4 (Phe51 and Val61) lie within the GEF consensus epitope as do three of the Ypt1 residues implicated in the interaction with Dss4 (corresponding to switch II residues Arg85, Ile87 and Thr89 in Rab3A) as well as four Rab3A residues found to be important for interaction with Rab3 GRF (Phe51, Thr54, Val55 and Phe59). Conversely, mutants which exhibited relatively small effects on Mss4 stimulated exchange activity involved residues outside the consensus epitope (Ser31, Val52, and Lys60). Thus, the available crystallographic and mutational data on Rab GTPases strongly supports a mode of GEF interactions similar to other GTPase families, suggesting that the GEF consensus epitope will serve as a meaningful likelihood predictor for the large family of Rab GTPases.

The crystal structure of the catalytically active core of Rabex-5 reveals an integrated architecture in which the Vps9 domain interacts extensively with an N-terminal helical bundle. In addition to the structural evidence, the lack of soluble expression for either the isolated Vps9 domain or the E212A mutant at the HB-Vps9 domain interface supports the conclusion that the helical bundle and Vps9 domain function together as a catalytic tandem. Several independent observations suggest that an N-terminal helical bundle, or an analogous helical domain, is likely to be a general

feature of Vps9 domain proteins. First, residues in the Rabex-5 Vps9 domain that contact the helical bundle are conserved. Second, the sequences of Vps9 proteins encode four predicted helices N-terminal to the Vps9 domain with significant overall similarity to the helical bundle of Rabex-5. Given that the region N-terminal to the Vps9 domain in some proteins contains more than four predicted helices, it is conceivable that larger helical domains could substitute for the helical bundle. Third, the level of sequence conservation within elements of predicted secondary structure is strongly correlated with the degree to which residues are buried in the Rabex-5 structure. Finally, a construct of RIN1 corresponding to the HB-Vps9 tandem of Rabex-5 can be expressed in a soluble form and has comparable catalytic activity and Rab specificity (data not shown). Moreover, a splice-variant of RIN1, which deletes the region from the N-terminus of αN4 to the N-terminus of αV2, fails to interact with a dominant negative mutant of Rab5 in a yeast two-hybrid assay, as does a construct that begins at the N-terminus of αV1 and extends through the C-terminus of the protein (317).

A tandem domain architecture has been observed in GEFs for other GTPases families. For example, the Sos (Son of sevenless) protein contains two distinct GEF activities, a DH-PH tandem that functions as a Rac GEF, and a Ras GEF tandem consisting of a N-terminal helical domain and the core catalytic domain (114,336,337). In the latter case, the GTP-bound form of Ras stimulates Ras exchange activity by binding at the interface between the two domains and inducing a conformational change that allows the N-terminal domain to contribute to the interaction with the nucleotide free form of Ras (336). The DH-PH tandem of Rho family GEFs exhibits considerable

structural variability with respect to the arrangement of the DH and PH domains (116,338-341). In at least one instance, the PH domain directly contacts the nucleotide free form of the GTPase, an interaction that is important for exchange activity (338,339). In Rabex-5, the surface of the helical bundle is poorly conserved, suggesting that the helical bundle is not directly involved as a general determinant of the interaction with Rab GTPases. Nevertheless, both structural and functional observations suggest that the helical bundle contributes indirectly to exchange activity by stabilizing the Vps9 domain. Beyond a structural role, it is possible that the helical bundle may play a role in allosteric regulation of GEF activity, which has been observed in RIN1 (317).

An earlier investigation of the kinetics of nucleotide exchange in full length Rabex-5, as well as the yeast Vps9 protein, concluded that these GEFs are relatively weak exchangers for Rab5 compared with other well characterized GEFs, in particular RCC1 and EF-Ts, which catalyze exchange for Ran and EF-Tu, respectively (314). I also observe equivalently weak Rab5 exchange activity for full length Rabex-5 (data not shown) even though the HB-Vps9 tandem, when expressed independently, has much more robust GEF activity. The reaction velocity of full length Rabex-5 saturates in the nM concentration range (K_m = 270 nM) with a k_{cat} of 0.007 s^{-1} (314). For the HB-Vps9 tandem, on the other hand, $K_m > 2$ μM and $k_{cat} > 0.1$ s^{-1}. One possible explanation is that the low catalytic activity of full length Rabex-5 reflects incomplete or improper folding of one or more domains in the bacterially expressed protein. In support of this view, Rabex-5 in complex with its cognate binding partner Rabaptin-5 has higher exchange activity for Rab5 compared with Rabex-5 alone (53). An alternative

hypothesis, motivated by the observation of allosteric regulation in RIN1 and Sos, is that the lower exchange activity of full length Rabex-5 reflects auto-inhibition by regulatory elements in the regions N- and/or C-terminal to the HB-Vps9 tandem. In the context of this model, it is possible that the interaction of GTP-bound Rab5 with Rabaptin-5, in addition to its role in endosomal targeting, might stimulate the GEF activity of the HB-Vps9 tandem. Additional experiments are necessary to distinguish between these possibilities.

We have shown that the HB-Vps9 tandem of Rabex-5 has equivalently high catalytic efficiency for Rab5 and Rab21, weak (100 fold lower) catalytic efficiency for Rab22, and no detectable activity for 28 other Rab GTPases for which nucleotide release can be stimulated by addition of EDTA. A phylogenetic analysis of Rab GTPases from diverse organisms identified eight distinct sub-families, one of which (sub-family V) consists of three Rab GTPases: Rab5, Rab21, and Rab22 (48,60). Thus, the HB-Vps9 tandem of Rabex-5 selectively recognizes sub-family V Rab GTPases, although it shows a strong preference for Rab5 and Rab21 over Rab22. Despite similarity within 'sub-family motifs', inspection of the full Rab-family alignment does not reveal a simple pattern of substitutions that could account for the observed specificity. It is therefore likely that the structural basis for highly selective recognition of the HB-Vps9 tandem by Rab GTPases involves multiple determinants.

Little is currently known about Rab21, although it is reported to have a relatively broad tissue distribution and to localize to the ER in unpolarized epithelial cells and to an apical pool of vesicles in polarized cells (342). Rab22 localizes to EEA1-positive

endosomes as well as the TGN and is proposed to regulate transport between the TGN and endosomes (285,343,344). Interestingly, the N-terminus of EEA1 binds Rab22 as well as Rab5, suggesting a possible molecular mechanism for regulating the fusion of endosomes with vesicles derived from the TGN (285). Furthermore, the Rabaptin-5/Rabex-5 complex has recently been shown to interact with GGA proteins, which associate with the cytosolic face of the TGN and function as Arf-dependent clathrin adaptors (345). One possibility is that GGA-dependent recruitment of the Rabaptin-5/Rabex-5 complex to the TGN leads to activation of Rab22, which may in turn be required for the fusion of TGN-derived vesicles with endosomes. The weak catalytic efficiency of the Rabex-5 HB-Vps9 tandem for Rab22 would seem to suggest that Rab22 may not be a physiological substrate; however, if the weak activity *in vitro* reflects a high K_m rather than a low k_{cat}, then the locally high concentrations resulting from co-localization might conceivably overcome a low affinity for the GDP-bound form of Rab22. Additional studies will be required to characterize the Rab specificity of the other proteins that contain Vps9 domains and determine whether these proteins and/or the Rabaptin-5/Rabex-5 complex function as Rab21 and/or Rab22 GEFs *in vivo*.

Since the ALS2 gene was cloned, nine SNPs from different families have been identified throughout the coding region (322-326). All of the SNPs involve frame shift or non-sense mutations that result in C-terminal truncations, the clinical severity of which does not correlate with the extent of the truncation (322). Given that the truncated protein products lack part or all of the C-terminal Vps9 domain, it has been suggested that the motor neuron degeneration associated with the ALS2 SNPs might result from

loss of Rab5 GEF activity (316). However, it has recently been demonstrated that the various ALS2 SNPs, including that in the Vps9 domain, give rise to unstable proteins that are targeted for degradation by the proteosome (327). From the structural and functional data presented here, it is clear that the SNP in the Vps9 domain necessarily results in a major structural defect in addition to eliminating critical determinants of GEF activity. The identification of point mutants that severely impair GEF activity without apparent effect on the structure of the Rabex-5 HB-Vps9 domain tandem, should provide useful reagents for exploring the *in vivo* consequences resulting from impaired GEF activity in ALS2 as well as other proteins which contain Vps9 domains.

CHAPTER V

DISCUSSION

Determinants of Effector Recognition

As critical regulators of membrane trafficking, the function of Rab GTPases is contingent upon specific interactions with effector proteins in the active state. In the inactive state, however, Rab proteins maintain the ability to interact with general accessory factors, such as REP and RabGDI. A model for Rab effector binding put forth by Ostermeier and Brunger suggests that Rab protein specificity is generated through interactions with the hypervariable CDRs whereas the nucleotide state is detected by interactions with the relatively conserved switch regions (Figure 5B). Consistent with this model, structural data for the active conformation of heterotrimeric G-proteins, the second largest GTPase family, demonstrates an identical conformation for the switch regions (265-267). Comparison of the Rab5C and Rab3A structures demonstrates that the active conformation of the switch regions provides information concerning Rab identity as well as the nucleotide state, and therefore serves as a determinant of Rab specificity. Of particular note is the conformational variability of invariant hydrophobic residues at the switch interface, which make critical contributions to effector binding in the Rab3A/Rabphilin-3A complex.

The switch regions of activated Rab proteins are composed primarily of residues that are highly conserved in Rab family members. The unexpected conformational

variability observed in the switch regions can be explained by nonconservative substitutions in flanking regions that vary between or even within Rab subfamilies. It is likely that the marginal conformational stability of the switch regions in the active state makes them sensitive to the variability of the flanking regions against which they pack. These features of Rab GTPases provide a mechanism for generating specificity in the active state while preserving the ability to recognize general accessory factors in the inactive state. They may have also facilitated expansion of the Rab GTPase family to accommodate the large number of functionally distinct members.

The interaction of Rab5C with EEA1 and Rabenosyn-5 shows no significant difference in affinity for the full length protein compared with the GTPase domain, which partially truncates CDRs 1 and 3. This truncation does not, however, affect the α3/β5 loop corresponding to CDR 2 and leaves part of CDR 1 and 3 intact. Also relevant is the size of the interaction epitope and affinity of the effector interactions. In chapter III, the compact (40 aa) C_2H_2 Zn^{2+} fingers of EEA1 and Rabenosyn-5 were shown to be effectors of Rab5C. Furthermore, the C_2H_2 Zn^{2+} finger in EEA1 is both necessary and sufficient for Rab5C binding to the N-terminus of EEA1. Structural homology models suggest that it is unlikely the C_2H_2 Zn^{2+} finger of EEA1 or Rabenosyn-5 could contact both the switch regions, by definition an essential requirement for an effector interaction, as well as the CDR regions of Rab5C. The low micromolar affinity of the Rab5C/EEA1 interaction also contrasts with the 10-20nM affinity of the Rab3A/Rabphilin-3A complex (346). While both Rab/effector interactions appear to involve buried hydrophobic surface area, the 100 fold difference in affinity is likely a reflection of the 2800Å^2 of buried

surface area in the Rab3A/Rabphilin-3A complex in comparison to the relatively small hydrophobic surface area of the predicted Rab5 binding epitope in the EEA1 homology model (90). These observations are consistent with the hypothesis that the hydrophobic switch interface may comprise the entire interaction epitope for some if not the majority of Rab effectors. Viewed from this perspective, the large interaction epitopes, observed in complexes such as Rab3A/Rabphilin-3A or Rab27A/Melanophilin, serve to enhance affinity of Rab/effector interactions where necessary, without loss in specificity (90,347,348).

Further analysis of both Rab GTPase structure and Rab/effector interactions is necessary to test this alternative hypothesis for Rab/effector specificity determination. So far, each Rab GTPase structure determined has shown a unique conformation with respect to the invariant hydrophobic triad highlighted in the Rab5C structure, indicative of combinatorial variability in the conformation of the switch regions (Figure 36). The determination of additional Rab structures will test the extent of structural variability within the Rab family and establish whether Rab subfamilies have similar or distinct active conformations. With respect to the switch interface, the structures of subfamily V members Rab5C and Rab22 are highly related, whereas subfamily VII members Ypt7 and Rab9 show little similarity (Sudharshan Eathiraj, personal communication) (349). This is consistent with the observation that Rab5 and Rab22 both interact with the C_2H_2 Zn^{2+} finger at the N-terminus of EEA1 (162,285).

A more thorough analysis, by mutagenesis or chimeric studies, should provide a better understanding of the role of the hypervariable CDRs in membrane targeting and

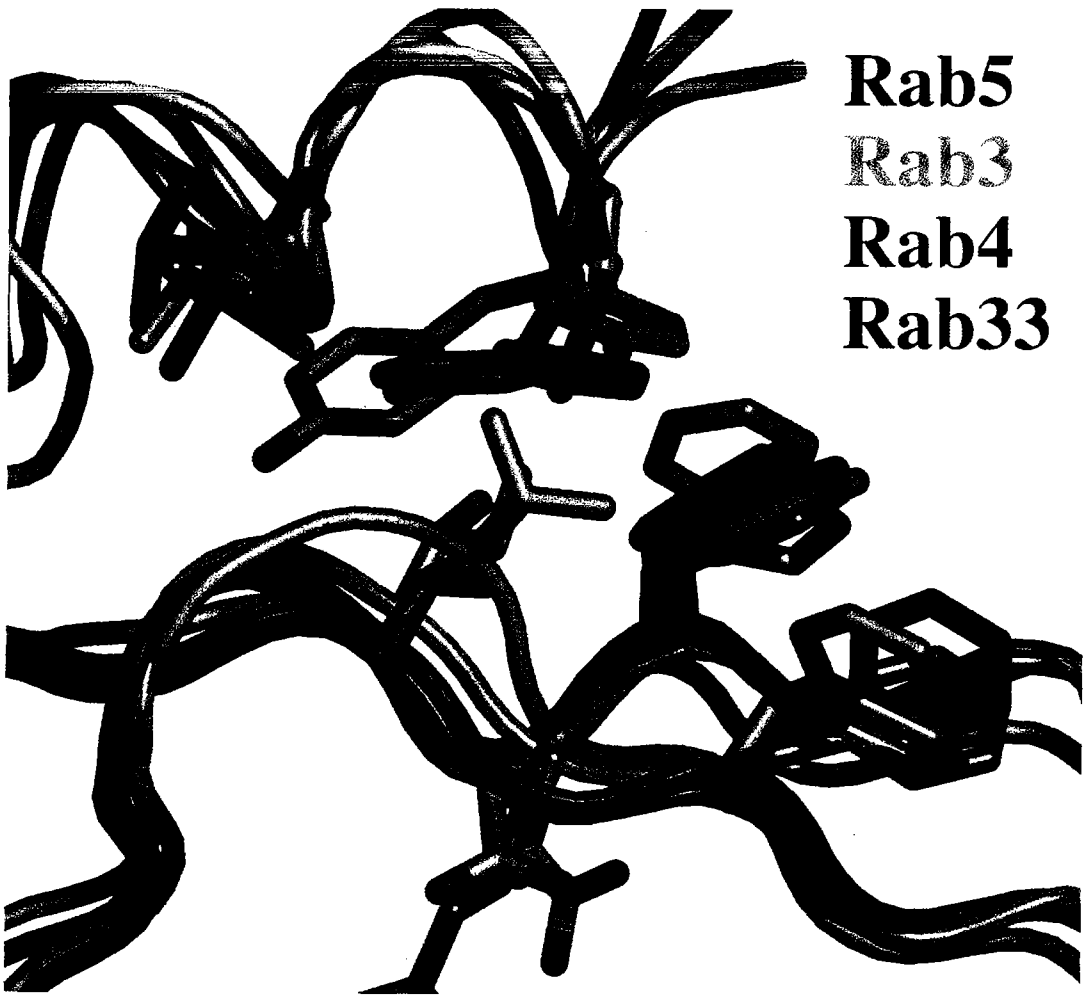


Figure 36. Combinatorial Variation in the Rab Switch Interface. Comparison of multiple Rab GTPase structures demonstrates a unique conformation at the switch interface consistent with a potential epitope for both Rab and nucleotide specificity. Rab5 shown in blue, Rab3 in orange, Rab4 in green and Rab33 in magenta.

effector recognition. The identification of more Rab effectors coupled with the analysis of their affinity and interaction epitopes would provide insight into the role of the CDRs and whether the mode of interaction observed in the Rab3A/Raphilin-3A complex represents an exception or the rule (90). The most critical data necessary for a definitive understanding of Rab/effector interactions is additional structures of Rab/effector complexes by either NMR or crystallographic methods. The characterized interaction epitope for the C_2H_2 Zn^{2+} finger of both EEA1 and Rabenosyn-5 should provide an excellent opportunity to obtain structural information for a complex with Rab5 or Rab22.

Characteristics of Rab GTPases

The structure of Rab5C provided additional evidence of how Rab GTPases differ with respect to Rab3A as well as other GTPase families. Overall, the conclusions derived from the structure of Rab3A in comparison to heterotrimeric G-proteins, Ras and other GTPases families are consistent with the Rab5C structure and have subsequently been corroborated by additional Rab structures (18,88,90,239,256,349,350). For example, variations in nucleotide contacts appear to account for differences in intrinsic nucleotide exchange and hydrolysis rates. Careful analysis of the growing number of Rab GTPase structures and comparison with measured rates for both intrinsic exchange and hydrolysis could provide insight into the structural bases for enzymatic variability.

EEA1 Mediated Endosomal Tethering

Through regulated effector recruitment, Rab proteins appear to support the formation of membrane subdomains and may be integral in defining organelle identity. It has been suggested that small pools of activated Rab GTPases may recruit additional factors and/or complexes that tether incoming vesicles, promote membrane association and activate SNARE mediated fusion. Chapter III provided additional evidence that the endosomal associated EEA1 may function in tethering of CCVs with the endosome.

The structure of the C-terminal region of EEA1 bound to the headgroup of PtdIns(3)P reveals an organized quaternary assembly consisting of a parallel coiled-coil and a dyad symmetric FYVE domain homodimer. The unique mode of membrane engagement deduced from the quaternary structure of the C-terminal region combined with the characterization of Rab5 binding at the N-terminal C_2H_2 Zn^{2+} finger provides insight into the structural basis of endosome tethering. As depicted in our hypothetical model (Figure 26), the predicted coiled-coil from residues 281-1346 could project a significant distance (~170 nm) into the surrounding cytoplasm. CCVs from the plasma membrane containing activated Rab5 could then be tethered through the N-terminal interaction and recruitment of additional Rab5 effectors including hVps34 would increase PtdIns(3)P and EEA1 on the vesicular membrane. The low micromolar affinity between Rab5C and the N-terminus of EEA1 would likely be amplified by simultaneous interaction of an EEA1 dimer with two membrane bound Rab5-GTP molecules, allowing for stable tethering with sufficient dynamic flexibility to promote fusion.

The elongated structure depicted in Figure 26 is based on the prediction of coiled-coil over the majority of the EEA1 protein. However, there is no experimental data available concerning the structure within most of this region. Electron microscopy (EM) studies with immunogold-labeled EEA1 locate the N-terminus around 50 nm from the endosome membrane, suggesting that the coiled-coil region may contain additional bends or kinks (282,351). By EM, the predicted heptad repeat of the yeast tethering factor Uso1p has been shown to form an extended coiled-coil of ~45 nm (156). A similar analysis could be performed for full-length EEA1 to test our hypothesis of an elongated structure and identify potential hinge regions. Purification of additional full-length proteins including prenylated Rab5 may allow in vitro reconstitution of tethered intermediates. Although reconstitution assays have been conducted for partially purified systems, assays using highly purified protein will allow for the identification of additional endosomal requirements for tethering and fusion events. If sufficiently stable, isolated tethered intermediates could be further analyzed by light or electron microscopic techniques.

Recent observations on the complex of Arl-1 and the GRIP domain of the effector Golgin-245 serves as an interesting corollary with studies on EEA1 (Figure 37) (352,353). The GRIP domain homodimer is ideally configured to interact with two molecules of Arl-1, providing the driving force for stable membrane anchoring. This organization is reminiscent of the organized quaternary structure of the EEA1 C-terminal region that ideally positions the FYVE domain dimer for membrane engagement (Figure 17 & 18). The Arl-1/GRIP complex is further analogous to the proposed model for how

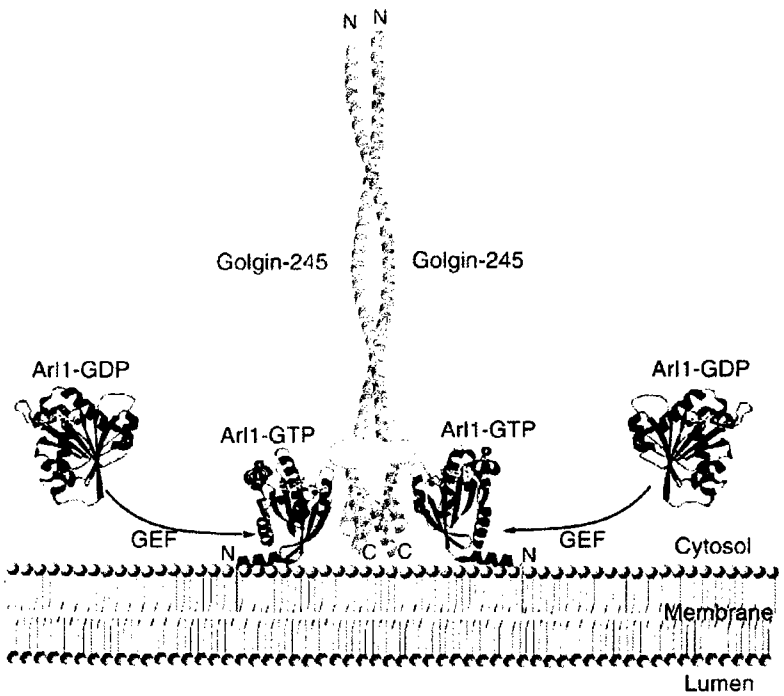


Figure 37. Schematic model of recruitment of Golgin-245 by Arl-1. The Arl-1 GTP could recruit Golgin-245 to the Golgi membrane by interacting with its GRIP domain. Golgin-245 forms a parallel dimer in which its coiled-coil region and GRIP domain are independently dimerized. Each subunit of the GRIP dimer interacts separately with one Arl-1-GTP molecule, thus efficiently anchoring Golgin-245 to the membrane. Switch I, interswitch, switch II regions and the N-terminal region of Arl-1 are black, blue, magenta and red respectively. GDP and GTP molecules are shown in stick models and Mg^{2+} as cyan sphere. Figure taken from Wu et al., 2003.

the dimeric C_2H_2 Zn^{2+} finger region of EEA1 interacts with two molecules of activated Rab5 on endosomal membranes. In addition, the size of the interaction epitope and the engagement of the switch regions of Arl-1, but not other regions, is consistent with the predicted mode of Rab5 binding to the C_2H_2 Zn^{2+} finger of EEA1. In both cases, the predicted coiled-coil region proximal to the protein-protein or protein-lipid interaction domains appears to facilitate the tethering function.

Structural and biochemical analysis of additional Rab/effector complexes will provide information on which, if any, interaction determinants are conserved and which are unique. By definition, effectors must recognize the active conformation of one or both switch regions. The extent of the interaction epitope beyond the switch regions is likely to vary considerably. Further attention must also be given to any dependence on quaternary structure, including oligomeric state or heterocomplex formation, when assessing requirements for Rab binding. The high proportion of protein complexes and self associating proteins among known Rab effectors may be significant for presentation of a functional interaction epitope or enhancement of affinity through avidity.

Currently the majority of published reports on Rab interactions test a small subset of Rab GTPases for interaction with effector proteins, which are often identified in two-hybrid screens. To avoid bias, known and putative Rab effectors should be systematically tested for interaction with all members of the Rab GTPase family. An understanding of functional Rab specificity requires an analysis of the entire spectrum of Rab/effector interactions. This information must be combined with other data, such as *in vivo* localization and characterization of additional domains within the effector to

understand its role in membrane trafficking. In some cases, a systematic analysis of partially characterized effector binding domains may reveal higher affinity interactions with other Rab GTPases, which would suggest alternative functional hypotheses.

Structural Mechanisms of Rab/GEF Interactions

The concept of a consensus GEF interaction epitope derived from known GTPase/GEF structures provides a framework for the analysis of Rab/GEF interactions. The consensus GEF interaction epitope is consistent with the limited available mutational and biochemical data concerning Rab/GEF recognition. However, additional Rab/GEF structural and biochemical data are necessary to identify determinants for recognition of Rab family proteins and identify differences with other GTPase families.

Of equal importance are recognition determinants in the exchange factor. As previously discussed, the known exchange factors for Rab GTPases show limited sequence similarity, making overall structural homology unlikely. This reasoning is consistent with the apparent lack of structural homology among GEFs for the various GTPase families (114-117). The structure of the Rabex-5 HB-Vps9 tandem provides the first insight into the interaction determinants of a highly specific Rab GEF. The most striking feature of the Vps9 domain is the critical stabilizing interaction with an N-terminal helical bundle, which is not accurately represented in sequence homology databases. Careful analysis of sequence conservation and predicted secondary structure suggests that the helical bundle is present in other Vps9 domain proteins. This similarity in primary and secondary structure should be confirmed by the characterization of

additional Vps9 domains to determine whether the helical region is required for structural stability and/or nucleotide exchange activity. The corresponding HB-Vps9 tandem of RIN1 has already been expressed and exhibits high exchange activity for Rab5 (Anna Delprato, personal communication.

The increased exchange activity of the Rabex-5 HB-Vps9 construct in comparison to published reports for full-length Rabex-5 may be due to poor behavior of the full-length protein or autoinhibitory effects. Our experience suggests that full-length Rabex-5 purified from *E. coli* is poorly behaved and therefore may be misfolded. However, this does not preclude the possibility of an inhibitory domain within the full-length protein that is not present in the HB-Vps9 domain construct. The full-length protein contains additional recognizable domains outside of the HB-Vps9 region, including a C_4 Zn^{2+} finger, a coiled-coil region and a proline rich region, the functions of which are unknown. Any of these domains may be capable of inhibiting Rabex-5 exchange activity in the context of a "closed" conformation for the full-length protein. There is some evidence in other Vps9 domain containing proteins that allosteric regulation may affect Rab exchange activity. Ras associates with the Rab binding domain (RBD) juxtaposed to the Vps9 domain of RIN1 and has been shown to affect RIN1 exchange activity (317,354). The observed differences in exchange activity may also be a reflection of protein stability. It has been shown that Rabex-5 forms an obligate heterodimer with Rabaptin-5 *in vivo,* and the activity in the complex is higher (53). It is uncertain whether Rabaptin-5 stabilizes the full-length Rabex-5 protein or relieves an inhibitory state.

The Rab family wide analysis of GEF specificity demonstrates that Rabex-5 stimulates nucleotide exchange for members of subfamily V (Rab5, Rab21 and Rab22). An analysis of conserved residues within subfamily V that differ in other Rab subfamilies identifies motifs in both switch I (^{56}AAFLT^{60}) and switch II (^{83}YHSLAPMY^{90}) as putative specificity determinants. It is likely that the structural basis for highly selective recognition of Rab GTPases by the HB-Vps9 tandem involves multiple determinants in the switch regions. A complex between Rabex-5 and one of the Rab GTPases of Subfamily V would provide the structural information necessary to identify these determinants and allow for a re-evaluation of the consensus GEF interaction epitope to include Rab/GEF determinants.

The range of interactions between Rab5 effectors and other Rab GTPases shows how cross-talk between membrane compartments may occur. EEA1 associates with both Rab22 and Rab5, whereas Rabenosyn-5 and Rabaptin-5 interact with Rab5 as well as Rab4. Here we show that a Rab accessory factor, Rabex-5, can interact with both Rab21 and Rab22 in addition to Rab5. Rab22 localizes to EEA1-positive endosomes as well as the TGN and is proposed to regulate transport between the TGN and endosomes (285,343,344). This lends support to the interplay of vesicular trafficking steps that must exist between the Golgi, early endosomes and recycling endosomes for proper protein sorting to occur. Additional studies will be required to characterize the Rab specificity of the other proteins that contain Vps9 domains and determine whether these proteins and/or the Rabaptin-5/Rabex-5 complex function as Rab21 and/or Rab22 GEFs *in vivo*.

The critical point mutants identified in this study may also be useful for understanding the role of Rabex-5 and other Vps9 domain containing proteins *in vivo*. The significant number of proteins present in the human genome that contain a Vps9 domain underscores the importance of Rab5 activation and the necessity of multiple pathways for regulating endocytic vesicle trafficking. Rab5 is essential for endocytosis and inhibition of the Vps9 domain proteins Rabex-5 and RIN1 perturb endosome fusion and EGFR internalization, respectively. The yeasts *S. cerevisiae* and *S. pombe* both contain two Vps9 domain proteins; however, only the *S. cerevisiae* Vps9 protein has been characterized. Deletion of the other three Vps9 domain proteins should be analyzed for differential stimulation of endocytosis. Other Vps9 domain proteins, such as RIN1 and ALS2, also require further characterization to understand their potential roles in endocytosis and disease. Recently, the ALS2 gene has been shown to interact with mutant forms of the superoxide dismutase (SOD1) gene product through its RhoGEF domain (355). Both proteins are associated with the onset of Amyotrophic Lateral Sclerosis (ALS). The potential role for the Vps9 domain in ALS2 and its role in both dominant and recessive forms of ALS have not been tested. A similar analysis for the role of RIN family members in Ras mediated signaling is at its early stages.

Family-wide Rab GTPase Characterization

Using Rab5 mediated endocytic transport as a model system, we have gained insight into the structural basis for the interaction of effectors and regulatory factors with Rab GTPases. In addition, we have used structural and biochemical approaches to define

how specific Rab5 interacting proteins function in the endocytic and recycling pathways. Insights from the analysis of the Rab5 structure will be investigated further in a family wide structural analysis of Rab proteins. In addition, the development of high-throughput approaches to characterize interactions with effectors and accessory factors was motivated by concepts described in this work. This unbiased proteomic approach will result in a more sophisticated structural and functional understanding of the role of Rab family proteins in the cell.

REFERENCES

1. Zerial, M., Suomalainen, M., Zanetti-Schneider, M., Schneider, C., and Garoff, H. (1987) *Embo J* **6**, 2661-2667
2. Segev, N. (2001) *Sci STKE* **2001**, RE11
3. Deneka, M., Neeft, M., and van der Sluijs, P. (2003) *Crit Rev Biochem Mol Biol* **38**, 121-142
4. Pfeffer, S. R. (2001) *Trends Cell Biol* **11**, 487-491
5. Pfeffer, S. R. (1994) *Curr Opin Cell Biol* **6**, 522-526
6. Rothman, J. E., and Wieland, F. T. (1996) *Science* **272**, 227-234
7. Chen, Y. A., and Scheller, R. H. (2001) *Nat Rev Mol Cell Biol* **2**, 98-106
8. Sollner, T. H. (2003) *Mol Membr Biol* **20**, 209-220
9. Rothman, J. E. (1994) *Nature* **372**, 55-63
10. Sollner, T. H., and Rothman, J. E. (1996) *Experientia* **52**, 1021-1025
11. Zerial, M., and McBride, H. (2001) *Nat Rev Mol Cell Biol* **2**, 107-117
12. Lazar, T., Gotte, M., and Gallwitz, D. (1997) *Trends Biochem Sci* **22**, 468-472
13. Schimmoller, F., Simon, I., and Pfeffer, S. R. (1998) *J Biol Chem* **273**, 22161-22164
14. Bock, J. B., Matern, H. T., Peden, A. A., and Scheller, R. H. (2001) *Nature* **409**, 839-841
15. Walworth, N. C., Brennwald, P., Kabcenell, A. K., Garrett, M., and Novick, P. (1992) *Mol Cell Biol* **12**, 2017-2028
16. Stenmark, H., Parton, R. G., Steele-Mortimer, O., Lutcke, A., Gruenberg, J., and Zerial, M. (1994) *Embo J* **13**, 1287-1296
17. Sprang, S. R. (1997) *Annu Rev Biochem* **66**, 639-678
18. Dumas, J. J., Zhu, Z., Connolly, J. L., and Lambright, D. G. (1999) *Structure Fold Des* **7**, 413-423
19. Bourne, H. R., Sanders, D. A., and McCormick, F. (1990) *Nature* **348**, 125-132
20. Macara, I. G., Lounsbury, K. M., Richards, S. A., McKiernan, C., and Bar-Sagi, D. (1996) *Faseb J* **10**, 625-630
21. Segev, N., and Botstein, D. (1987) *Mol Cell Biol* **7**, 2367-2377
22. Salminen, A., and Novick, P. J. (1987) *Cell* **49**, 527-538
23. Balch, W. E. (1990) *Trends Biochem Sci* **15**, 473-477
24. Gotte, M., Lazar, T., Yoo, J. S., Scheglmann, D., and Gallwitz, D. (2000) *Subcell Biochem* **34**, 133-173
25. Louvet, O., Roumanie, O., Barthe, C., Peypouquet, M. F., Schaeffer, J., Doignon, F., and Crouzet, M. (1999) *Mol Gen Genet* **261**, 589-600
26. Itoh, T., Watabe, A., Toh, E. A., and Matsui, Y. (2002) *Mol Cell Biol* **22**, 7744-7757
27. Rak, A., Fedorov, R., Alexandrov, K., Albert, S., Goody, R. S., Gallwitz, D., and Scheidig, A. J. (2000) *Embo J* **19**, 5105-5113
28. Jones, S., Newman, C., Liu, F., and Segev, N. (2000) *Mol Biol Cell* **11**, 4403-4411

29. Wang, W., Sacher, M., and Ferro-Novick, S. (2000) *J Cell Biol* **151**, 289-296
30. Barrowman, J., Sacher, M., and Ferro-Novick, S. (2000) *Embo J* **19**, 862-869
31. Morsomme, P., and Riezman, H. (2002) *Dev Cell* **2**, 307-317
32. Segev, N., Mulholland, J., and Botstein, D. (1988) *Cell* **52**, 915-924
33. Walch-Solimena, C., Collins, R. N., and Novick, P. J. (1997) *J Cell Biol* **137**, 1495-1509
34. Burton, J., Roberts, D., Montaldi, M., Novick, P., and De Camilli, P. (1993) *Nature* **361**, 464-467
35. Lipschutz, J. H., and Mostov, K. E. (2002) *Curr Biol* **12**, R212-214
36. Siniossoglou, S., Peak-Chew, S. Y., and Pelham, H. R. (2000) *Embo J* **19**, 4885-4894
37. Li, B., and Warner, J. R. (1998) *Yeast* **14**, 915-922
38. Luo, Z., and Gallwitz, D. (2003) *J Biol Chem* **278**, 791-799
39. Wurmser, A. E., Sato, T. K., and Emr, S. D. (2000) *J Cell Biol* **151**, 551-562
40. Wichmann, H., Hengst, L., and Gallwitz, D. (1992) *Cell* **71**, 1131-1142
41. Schimmoller, F., and Riezman, H. (1993) *J Cell Sci* **106 (Pt 3)**, 823-830
42. Benli, M., Doring, F., Robinson, D. G., Yang, X., and Gallwitz, D. (1996) *Embo J* **15**, 6460-6475
43. Jedd, G., Mulholland, J., and Segev, N. (1997) *J Cell Biol* **137**, 563-580
44. Burd, C. G., Mustol, P. A., Schu, P. V., and Emr, S. D. (1996) *Mol Cell Biol* **16**, 2369-2377
45. Hama, H., Tall, G. G., and Horazdovsky, B. F. (1999) *J Biol Chem* **274**, 15284-15291
46. Singer-Kruger, B., Stenmark, H., Dusterhoft, A., Philippsen, P., Yoo, J. S., Gallwitz, D., and Zerial, M. (1994) *J Cell Biol* **125**, 283-298
47. Touchot, N., Chardin, P., and Tavitian, A. (1987) *Proc Natl Acad Sci U S A* **84**, 8210-8214
48. Pereira-Leal, J. B., and Seabra, M. C. (2000) *J Mol Biol* **301**, 1077-1087
49. Echard, A., Opdam, F. J., de Leeuw, H. J., Jollivet, F., Savelkoul, P., Hendriks, W., Voorberg, J., Goud, B., and Fransen, J. A. (2000) *Mol Biol Cell* **11**, 3819-3833
50. Bucci, C., Parton, R. G., Mather, I. H., Stunnenberg, H., Simons, K., Hoflack, B., and Zerial, M. (1992) *Cell* **70**, 715-728
51. Barbieri, M. A., Roberts, R. L., Mukhopadhyay, A., and Stahl, P. D. (1996) *Biocell* **20**, 331-338
52. Nielsen, E., Severin, F., Backer, J. M., Hyman, A. A., and Zerial, M. (1999) *Nat Cell Biol* **1**, 376-382
53. Lippe, R., Miaczynska, M., Rybin, V., Runge, A., and Zerial, M. (2001) *Mol Biol Cell* **12**, 2219-2228
54. Zahraoui, A., Joberty, G., Arpin, M., Fontaine, J. J., Hellio, R., Tavitian, A., and Louvard, D. (1994) *J Cell Biol* **124**, 101-115
55. Marzesco, A. M., Dunia, I., Pandjaitan, R., Recouvreur, M., Dauzonne, D., Benedetti, E. L., Louvard, D., and Zahraoui, A. (2002) *Mol Biol Cell* **13**, 1819-1831

56. Lutcke, A., Jansson, S., Parton, R. G., Chavrier, P., Valencia, A., Huber, L. A., Lehtonen, E., and Zerial, M. (1993) *J Cell Biol* **121**, 553-564
57. Sonnichsen, B., De Renzis, S., Nielsen, E., Rietdorf, J., and Zerial, M. (2000) *J Cell Biol* **149**, 901-914
58. Pfeffer, S. (2003) *Cell* **112**, 507-517
59. Segev, N. (2001) *Curr Opin Cell Biol* **13**, 500-511
60. Pereira-Leal, J. B., and Seabra, M. C. (2001) *J Mol Biol* **313**, 889-901
61. Nielsen, E., Christoforidis, S., Uttenweiler-Joseph, S., Miaczynska, M., Dewitte, F., Wilm, M., Hoflack, B., and Zerial, M. (2000) *J Cell Biol* **151**, 601-612
62. Haubruck, H., Prange, R., Vorgias, C., and Gallwitz, D. (1989) *Embo J* **8**, 1427-1432
63. Armstrong, J., Craighead, M. W., Watson, R., Ponnambalam, S., and Bowden, S. (1993) *Mol Biol Cell* **4**, 583-592
64. Geppert, M., Bolshakov, V. Y., Siegelbaum, S. A., Takei, K., De Camilli, P., Hammer, R. E., and Sudhof, T. C. (1994) *Nature* **369**, 493-497
65. Sudhof, T. C. (2004) in *Traffic Control: Rab GTPases in Vesicular Transport*, Breckenridge, Colorado
66. Seabra, M. C. (1998) *Cell Signal* **10**, 167-172
67. Gomes, A. Q., Ali, B. R., Ramalho, J. S., Godfrey, R. F., Barral, D. C., Hume, A. N., and Seabra, M. C. (2003) *Mol Biol Cell* **14**, 1882-1899
68. Chavrier, P., Gorvel, J. P., Stelzer, E., Simons, K., Gruenberg, J., and Zerial, M. (1991) *Nature* **353**, 769-772
69. Stenmark, H., Valencia, A., Martinez, O., Ullrich, O., Goud, B., and Zerial, M. (1994) *Embo J* **13**, 575-583
70. Beranger, F., Paterson, H., Powers, S., de Gunzburg, J., and Hancock, J. F. (1994) *Mol Cell Biol* **14**, 744-758
71. Diaz, E., and Pfeffer, S. R. (1998) *Cell* **93**, 433-443
72. Carroll, K. S., Hanna, J., Simon, I., Krise, J., Barbero, P., and Pfeffer, S. R. (2001) *Science* **292**, 1373-1376
73. Olkkonen, V. M., and Ikonen, E. (2000) *N Engl J Med* **343**, 1095-1104
74. Deacon, S. W., and Gelfand, V. I. (2001) *J Cell Biol* **152**, F21-24
75. Seabra, M. C., Mules, E. H., and Hume, A. N. (2002) *Trends Mol Med* **8**, 23-30
76. Menasche, G., Pastural, E., Feldmann, J., Certain, S., Ersoy, F., Dupuis, S., Wulffraat, N., Bianchi, D., Fischer, A., Le Deist, F., and de Saint Basile, G. (2000) *Nat Genet* **25**, 173-176
77. Wilson, S. M., Yip, R., Swing, D. A., O'Sullivan, T. N., Zhang, Y., Novak, E. K., Swank, R. T., Russell, L. B., Copeland, N. G., and Jenkins, N. A. (2000) *Proc Natl Acad Sci U S A* **97**, 7933-7938
78. Hume, A. N., Collinson, L. M., Rapak, A., Gomes, A. Q., Hopkins, C. R., and Seabra, M. C. (2001) *J Cell Biol* **152**, 795-808
79. Seabra, M. C. (1996) *Ophthalmic Genet* **17**, 43-46
80. D'Adamo, P., Menegon, A., Lo Nigro, C., Grasso, M., Gulisano, M., Tamanini, F., Bienvenu, T., Gedeon, A. K., Oostra, B., Wu, S. K., Tandon, A., Valtorta, F., Balch, W. E., Chelly, J., and Toniolo, D. (1998) *Nat Genet* **19**, 134-139

81. D'Adamo, P., Welzl, H., Papadimitriou, S., Raffaele di Barletta, M., Tiveron, C., Tatangelo, L., Pozzi, L., Chapman, P. F., Knevett, S. G., Ramsay, M. F., Valtorta, F., Leoni, C., Menegon, A., Wolfer, D. P., Lipp, H. P., and Toniolo, D. (2002) *Hum Mol Genet* **11**, 2567-2580
82. Shamoon, H. (1992) *Drugs* **44 Suppl 3**, 1-12
83. Cunningham, S. A., Frizzell, R. A., and Morris, A. P. (1995) *J Physiol* **482**, 27S-30S
84. Dugan, J. M., deWit, C., McConlogue, L., and Maltese, W. A. (1995) *J Biol Chem* **270**, 10982-10989
85. Via, L. E., Deretic, D., Ulmer, R. J., Hibler, N. S., Huber, L. A., and Deretic, V. (1997) *J Biol Chem* **272**, 13326-13331
86. Scianimanico, S., Desrosiers, M., Dermine, J. F., Meresse, S., Descoteaux, A., and Desjardins, M. (1999) *Cell Microbiol* **1**, 19-32
87. Bourne, H. R., Sanders, D. A., and McCormick, F. (1991) *Nature* **349**, 117-127
88. Stroupe, C., and Brunger, A. T. (2000) *J Mol Biol* **304**, 585-598
89. Dever, T. E., Glynias, M. J., and Merrick, W. C. (1987) *Proc Natl Acad Sci U S A* **84**, 1814-1818
90. Ostermeier, C., and Brunger, A. T. (1999) *Cell* **96**, 363-374
91. Valencia, A., Chardin, P., Wittinghofer, A., and Sander, C. (1991) *Biochemistry* **30**, 4637-4648
92. Brennwald, P., and Novick, P. (1993) *Nature* **362**, 560-563
93. Ullrich, O., Stenmark, H., Alexandrov, K., Huber, L. A., Kaibuchi, K., Sasaki, T., Takai, Y., and Zerial, M. (1993) *J Biol Chem* **268**, 18143-18150
94. Burton, J. L., Burns, M. E., Gatti, E., Augustine, G. J., and De Camilli, P. (1994) *Embo J* **13**, 5547-5558
95. Yang, X., Matern, H. T., and Gallwitz, D. (1998) *Embo J* **17**, 4954-4963
96. Vitale, G., Rybin, V., Christoforidis, S., Thornqvist, P., McCaffrey, M., Stenmark, H., and Zerial, M. (1998) *Embo J* **17**, 1941-1951
97. Alexandrov, K., Simon, I., Yurchenko, V., Iakovenko, A., Rostkova, E., Scheidig, A. J., and Goody, R. S. (1999) *Eur J Biochem* **265**, 160-170
98. Anant, J. S., Desnoyers, L., Machius, M., Demeler, B., Hansen, J. C., Westover, K. D., Deisenhofer, J., and Seabra, M. C. (1998) *Biochemistry* **37**, 12559-12568
99. Andres, D. A., Seabra, M. C., Brown, M. S., Armstrong, S. A., Smeland, T. E., Cremers, F. P., and Goldstein, J. L. (1993) *Cell* **73**, 1091-1099
100. Araki, S., Kikuchi, A., Hata, Y., Isomura, M., and Takai, Y. (1990) *J Biol Chem* **265**, 13007-13015
101. Peter, F., Nuoffer, C., Pind, S. N., and Balch, W. E. (1994) *J Cell Biol* **126**, 1393-1406
102. Pfeffer, S. R., Dirac-Svejstrup, A. B., and Soldati, T. (1995) *J Biol Chem* **270**, 17057-17059
103. Soldati, T., Riederer, M. A., and Pfeffer, S. R. (1993) *Mol Biol Cell* **4**, 425-434
104. Dirac-Svejstrup, A. B., Sumizawa, T., and Pfeffer, S. R. (1997) *Embo J* **16**, 465-472

105. Garrett, M. D., Zahner, J. E., Cheney, C. M., and Novick, P. J. (1994) *Embo J* **13**, 1718-1728
106. Calero, M., and Collins, R. N. (2002) *Biochem Biophys Res Commun* **290**, 676-681
107. Calero, M., Winand, N. J., and Collins, R. N. (2002) *FEBS Lett* **515**, 89-98
108. Hutt, D. M., Da-Silva, L. F., Chang, L. H., Prosser, D. C., and Ngsee, J. K. (2000) *J Biol Chem* **275**, 18511-18519
109. Martincic, I., Peralta, M. E., and Ngsee, J. K. (1997) *J Biol Chem* **272**, 26991-26998
110. Abdul-Ghani, M., Gougeon, P. Y., Prosser, D. C., Da-Silva, L. F., and Ngsee, J. K. (2001) *J Biol Chem* **276**, 6225-6233
111. Sivars, U., Aivazian, D., and Pfeffer, S. R. (2003) *Nature* **425**, 856-859
112. Cherfils, J., and Chardin, P. (1999) *Trends Biochem Sci* **24**, 306-311
113. Sprang, S. R., and Coleman, D. E. (1998) *Cell* **95**, 155-158
114. Boriack-Sjodin, P. A., Margarit, S. M., Bar-Sagi, D., and Kuriyan, J. (1998) *Nature* **394**, 337-343
115. Goldberg, J. (1998) *Cell* **95**, 237-248
116. Worthylake, D. K., Rossman, K. L., and Sondek, J. (2000) *Nature* **408**, 682-688
117. Renault, L., Kuhlmann, J., Henkel, A., and Wittinghofer, A. (2001) *Cell* **105**, 245-255
118. Yu, H., and Schreiber, S. L. (1995) *Nature* **376**, 788-791
119. Zhu, Z., Dumas, J. J., Lietzke, S. E., and Lambright, D. G. (2001) *Biochemistry* **40**, 3027-3036
120. Horiuchi, H., Lippe, R., McBride, H. M., Rubino, M., Woodman, P., Stenmark, H., Rybin, V., Wilm, M., Ashman, K., Mann, M., and Zerial, M. (1997) *Cell* **90**, 1149-1159
121. Armstrong, J. (2000) *Int J Biochem Cell Biol* **32**, 303-307
122. Novick, P., and Zerial, M. (1997) *Curr Opin Cell Biol* **9**, 496-504
123. Albert, S., and Gallwitz, D. (1999) *J Biol Chem* **274**, 33186-33189
124. Du, L. L., Collins, R. N., and Novick, P. J. (1998) *J Biol Chem* **273**, 3253-3256
125. Strom, M., Vollmer, P., Tan, T. J., and Gallwitz, D. (1993) *Nature* **361**, 736-739
126. Vollmer, P., Will, E., Scheglmann, D., Strom, M., and Gallwitz, D. (1999) *Eur J Biochem* **260**, 284-290
127. Cuif, M. H., Possmayer, F., Zander, H., Bordes, N., Jollivet, F., Couedel-Courteille, A., Janoueix-Lerosey, I., Langsley, G., Bornens, M., and Goud, B. (1999) *Embo J* **18**, 1772-1782
128. Richardson, C. J., Jones, S., Litt, R. J., and Segev, N. (1998) *Mol Cell Biol* **18**, 827-838
129. Barbieri, M. A., Li, G., Mayorga, L. S., and Stahl, P. D. (1996) *Arch Biochem Biophys* **326**, 64-72
130. Nassar, N., Hoffman, G. R., Manor, D., Clardy, J. C., and Cerione, R. A. (1998) *Nat Struct Biol* **5**, 1047-1052
131. Rittinger, K., Walker, P. A., Eccleston, J. F., Smerdon, S. J., and Gamblin, S. J. (1997) *Nature* **389**, 758-762

132. Scheffzek, K., Ahmadian, M. R., Kabsch, W., Wiesmuller, L., Lautwein, A., Schmitz, F., and Wittinghofer, A. (1997) *Science* **277**, 333-338
133. Scheffzek, K., Lautwein, A., Kabsch, W., Ahmadian, M. R., and Wittinghofer, A. (1996) *Nature* **384**, 591-596
134. Ahmadian, M. R., Stege, P., Scheffzek, K., and Wittinghofer, A. (1997) *Nat Struct Biol* **4**, 686-689
135. McLauchlan, H., Newell, J., Morrice, N., Osborne, A., West, M., and Smythe, E. (1998) *Curr Biol* **8**, 34-45
136. Allan, B. B., Moyer, B. D., and Balch, W. E. (2000) *Science* **289**, 444-448
137. Mammoto, A., Sasaki, T., Kim, Y., and Takai, Y. (2000) *J Biol Chem* **275**, 13167-13170
138. Peranen, J., and Furuhjelm, J. (2001) *Methods Enzymol* **329**, 188-196
139. Hammer, J. A., 3rd, and Wu, X. S. (2002) *Curr Opin Cell Biol* **14**, 69-75
140. Echard, A., Jollivet, F., Martinez, O., Lacapere, J. J., Rousselet, A., Janoueix-Lerosey, I., and Goud, B. (1998) *Science* **279**, 580-585
141. Cantalupo, G., Alifano, P., Roberti, V., Bruni, C. B., and Bucci, C. (2001) *Embo J* **20**, 683-693
142. Jordens, I., Fernandez-Borja, M., Marsman, M., Dusseljee, S., Janssen, L., Calafat, J., Janssen, H., Wubbolts, R., and Neefjes, J. (2001) *Curr Biol* **11**, 1680-1685
143. Bielli, A., Thornqvist, P. O., Hendrick, A. G., Finn, R., Fitzgerald, K., and McCaffrey, M. W. (2001) *Biochem Biophys Res Commun* **281**, 1141-1153
144. Bahadoran, P., Aberdam, E., Mantoux, F., Busca, R., Bille, K., Yalman, N., de Saint-Basile, G., Casaroli-Marano, R., Ortonne, J. P., and Ballotti, R. (2001) *J Cell Biol* **152**, 843-850
145. Wu, X., Rao, K., Bowers, M. B., Copeland, N. G., Jenkins, N. A., and Hammer, J. A., 3rd. (2001) *J Cell Sci* **114**, 1091-1100
146. Wu, X., Wang, F., Rao, K., Sellers, J. R., and Hammer, J. A., 3rd. (2002) *Mol Biol Cell* **13**, 1735-1749
147. Lapierre, L. A., Kumar, R., Hales, C. M., Navarre, J., Bhartur, S. G., Burnette, J. O., Provance, D. W., Jr., Mercer, J. A., Bahler, M., and Goldenring, J. R. (2001) *Mol Biol Cell* **12**, 1843-1857
148. Hales, C. M., Vaerman, J. P., and Goldenring, J. R. (2002) *J Biol Chem* **277**, 50415-50421
149. Lindsay, A. J., and McCaffrey, M. W. (2002) *J Biol Chem* **277**, 27193-27199
150. Pruyne, D. W., Schott, D. H., and Bretscher, A. (1998) *J Cell Biol* **143**, 1931-1945
151. Schott, D., Ho, J., Pruyne, D., and Bretscher, A. (1999) *J Cell Biol* **147**, 791-808
152. Whyte, J. R., and Munro, S. (2002) *J Cell Sci* **115**, 2627-2637
153. Brennwald, P., Kearns, B., Champion, K., Keranen, S., Bankaitis, V., and Novick, P. (1994) *Cell* **79**, 245-258
154. McBride, H. M., Rybin, V., Murphy, C., Giner, A., Teasdale, R., and Zerial, M. (1999) *Cell* **98**, 377-386
155. de Renzis, S., Sonnichsen, B., and Zerial, M. (2002) *Nat Cell Biol* **4**, 124-133

156. Sapperstein, S. K., Lupashin, V. V., Schmitt, H. D., and Waters, M. G. (1996) *J Cell Biol* **132**, 755-767
157. Sonnichsen, B., Lowe, M., Levine, T., Jamsa, E., Dirac-Svejstrup, B., and Warren, G. (1998) *J Cell Biol* **140**, 1013-1021
158. Barr, F. A., and Short, B. (2003) *Curr Opin Cell Biol* **15**, 405-413
159. De Matteis, M. A., and Morrow, J. S. (2000) *J Cell Sci* **113 (Pt 13)**, 2331-2343
160. Fasshauer, D., Antonin, W., Subramaniam, V., and Jahn, R. (2002) *Nat Struct Biol* **9**, 144-151
161. Lippe, R., Horiuchi, H., Runge, A., and Zerial, M. (2001) *Methods Enzymol* **329**, 132-145
162. Simonsen, A., Lippe, R., Christoforidis, S., Gaullier, J. M., Brech, A., Callaghan, J., Toh, B. H., Murphy, C., Zerial, M., and Stenmark, H. (1998) *Nature* **394**, 494-498
163. Rubino, M., Miaczynska, M., Lippe, R., and Zerial, M. (2000) *J Biol Chem* **275**, 3745-3748
164. Rothman, J. E. (1994) *Adv Second Messenger Phosphoprotein Res* **29**, 81-96
165. Brunger, A. T. (2001) *Curr Opin Struct Biol* **11**, 163-173
166. Jahn, R., and Sudhof, T. C. (1999) *Annu Rev Biochem* **68**, 863-911
167. Kee, Y., Lin, R. C., Hsu, S. C., and Scheller, R. H. (1995) *Neuron* **14**, 991-998
168. Nichols, B. J., Holthuis, J. C., and Pelham, H. R. (1998) *Eur J Cell Biol* **77**, 263-268
169. Pevsner, J., Hsu, S. C., and Scheller, R. H. (1994) *Proc Natl Acad Sci U S A* **91**, 1445-1449
170. Munson, M., and Hughson, F. M. (2002) *J Biol Chem* **277**, 9375-9381
171. Munson, M., Chen, X., Cocina, A. E., Schultz, S. M., and Hughson, F. M. (2000) *Nat Struct Biol* **7**, 894-902
172. Waters, M. G., and Hughson, F. M. (2000) *Traffic* **1**, 588-597
173. Pfeffer, S. R. (1999) *Nat Cell Biol* **1**, E17-22
174. Simonsen, A., Gaullier, J. M., D'Arrigo, A., and Stenmark, H. (1999) *J Biol Chem* **274**, 28857-28860
175. Seals, D. F., Eitzen, G., Margolis, N., Wickner, W. T., and Price, A. (2000) *Proc Natl Acad Sci U S A* **97**, 9402-9407
176. Price, A., Seals, D., Wickner, W., and Ungermann, C. (2000) *J Cell Biol* **148**, 1231-1238
177. Laage, R., and Ungermann, C. (2001) *Mol Biol Cell* **12**, 3375-3385
178. Sato, T. K., Rehling, P., Peterson, M. R., and Emr, S. D. (2000) *Mol Cell* **6**, 661-671
179. Shorter, J., Beard, M. B., Seemann, J., Dirac-Svejstrup, A. B., and Warren, G. (2002) *J Cell Biol* **157**, 45-62
180. Siniossoglou, S., and Pelham, H. R. (2002) *J Biol Chem* **277**, 48318-48324
181. Conibear, E., Cleck, J. N., and Stevens, T. H. (2003) *Mol Biol Cell* **14**, 1610-1623
182. Gaullier, J. M., Simonsen, A., D'Arrigo, A., Bremnes, B., Stenmark, H., and Aasland, R. (1998) *Nature* **394**, 432-433

183. Corvera, S., D'Arrigo, A., and Stenmark, H. (1999) *Curr Opin Cell Biol* **11**, 460-465
184. Stenmark, H., and Aasland, R. (1999) *J Cell Sci* **112 (Pt 23)**, 4175-4183
185. Stenmark, H., Aasland, R., Toh, B. H., and D'Arrigo, A. (1996) *J Biol Chem* **271**, 24048-24054
186. Raiborg, C., Bache, K. G., Gillooly, D. J., Madshus, I. H., Stang, E., and Stenmark, H. (2002) *Nat Cell Biol* **4**, 394-398
187. Katzmann, D. J., Stefan, C. J., Babst, M., and Emr, S. D. (2003) *J Cell Biol* **162**, 413-423
188. Hicke, L., and Dunn, R. (2003) *Annu Rev Cell Dev Biol* **19**, 141-172
189. Teis, D., and Huber, L. A. (2003) *Cell Mol Life Sci* **60**, 2020-2033
190. Vieira, A. V., Lamaze, C., and Schmid, S. L. (1996) *Science* **274**, 2086-2089
191. Di Guglielmo, G. M., Baass, P. C., Ou, W. J., Posner, B. I., and Bergeron, J. J. (1994) *Embo J* **13**, 4269-4277
192. Oksvold, M. P., Skarpen, E., Lindeman, B., Roos, N., and Huitfeldt, H. S. (2000) *J Histochem Cytochem* **48**, 21-33
193. Howe, C. L., Valletta, J. S., Rusnak, A. S., and Mobley, W. C. (2001) *Neuron* **32**, 801-814
194. Pol, A., Calvo, M., and Enrich, C. (1998) *FEBS Lett* **441**, 34-38
195. York, R. D., Yao, H., Dillon, T., Ellig, C. L., Eckert, S. P., McCleskey, E. W., and Stork, P. J. (1998) *Nature* **392**, 622-626
196. Hayes, S., Chawla, A., and Corvera, S. (2002) *J Cell Biol* **158**, 1239-1249
197. Tsukazaki, T., Chiang, T. A., Davison, A. F., Attisano, L., and Wrana, J. L. (1998) *Cell* **95**, 779-791
198. Itoh, F., Divecha, N., Brocks, L., Oomen, L., Janssen, H., Calafat, J., Itoh, S., and Dijke Pt, P. (2002) *Genes Cells* **7**, 321-331
199. Miura, S., Takeshita, T., Asao, H., Kimura, Y., Murata, K., Sasaki, Y., Hanai, J. I., Beppu, H., Tsukazaki, T., Wrana, J. L., Miyazono, K., and Sugamura, K. (2000) *Mol Cell Biol* **20**, 9346-9355
200. Goodman, O. B., Jr., Krupnick, J. G., Santini, F., Gurevich, V. V., Penn, R. B., Gagnon, A. W., Keen, J. H., and Benovic, J. L. (1996) *Nature* **383**, 447-450
201. Laporte, S. A., Oakley, R. H., Zhang, J., Holt, J. A., Ferguson, S. S., Caron, M. G., and Barak, L. S. (1999) *Proc Natl Acad Sci U S A* **96**, 3712-3717
202. McDonald, P. H., Chow, C. W., Miller, W. E., Laporte, S. A., Field, M. E., Lin, F. T., Davis, R. J., and Lefkowitz, R. J. (2000) *Science* **290**, 1574-1577
203. Clague, M. J., and Urbe, S. (2001) *J Cell Sci* **114**, 3075-3081
204. Barbieri, M. A., Hoffenberg, S., Roberts, R., Mukhopadhyay, A., Pomrehn, A., Dickey, B. F., and Stahl, P. D. (1998) *J Biol Chem* **273**, 25850-25855
205. Christoforidis, S., and Zerial, M. (2000) *Methods* **20**, 403-410
206. Roberts, R. L., Barbieri, M. A., Pryse, K. M., Chua, M., Morisaki, J. H., and Stahl, P. D. (1999) *J Cell Sci* **112 (Pt 21)**, 3667-3675
207. Yarden, Y. (2001) *Eur J Cancer* **37 Suppl 4**, S3-8
208. Yarden, Y., and Sliwkowski, M. X. (2001) *Nat Rev Mol Cell Biol* **2**, 127-137

209. Barbieri, M. A., Roberts, R. L., Gumusboga, A., Highfield, H., Alvarez-Dominguez, C., Wells, A., and Stahl, P. D. (2000) *J Cell Biol* **151**, 539-550
210. Sorkina, T., Huang, F., Beguinot, L., and Sorkin, A. (2002) *J Biol Chem* **277**, 27433-27441
211. Wilde, A., Beattie, E. C., Lem, L., Riethof, D. A., Liu, S. H., Mobley, W. C., Soriano, P., and Brodsky, F. M. (1999) *Cell* **96**, 677-687
212. Benmerah, A., Gagnon, J., Begue, B., Megarbane, B., Dautry-Varsat, A., and Cerf-Bensussan, N. (1995) *J Cell Biol* **131**, 1831-1838
213. van Delft, S., Schumacher, C., Hage, W., Verkleij, A. J., and van Bergen en Henegouwen, P. M. (1997) *J Cell Biol* **136**, 811-821
214. Galisteo, M. L., Dikic, I., Batzer, A. G., Langdon, W. Y., and Schlessinger, J. (1995) *J Biol Chem* **270**, 20242-20245
215. Fazioli, F., Minichiello, L., Matoska, V., Castagnino, P., Miki, T., Wong, W. T., and Di Fiore, P. P. (1993) *Embo J* **12**, 3799-3808
216. Danielsen, A. J., and Maihle, N. J. (2002) *Growth Factors* **20**, 1-15
217. Carpenter, G., Soler, C., Baulida, J., Beguinot, L., and Sorkin, A. (1995) *Ann N Y Acad Sci* **766**, 44-51
218. Lanzetti, L., Rybin, V., Malabarba, M. G., Christoforidis, S., Scita, G., Zerial, M., and Di Fiore, P. P. (2000) *Nature* **408**, 374-377
219. Matoskova, B., Wong, W. T., Nomura, N., Robbins, K. C., and Di Fiore, P. P. (1996) *Oncogene* **12**, 2679-2688
220. Cavalli, V., Vilbois, F., Corti, M., Marcote, M. J., Tamura, K., Karin, M., Arkinstall, S., and Gruenberg, J. (2001) *Mol Cell* **7**, 421-432
221. Murray, J. T., Panaretou, C., Stenmark, H., Miaczynska, M., and Backer, J. M. (2002) *Traffic* **3**, 416-427
222. Hayakawa, A., Hayes, S. J., Lawe, D. C., Sudharshan, E., Tuft, R., Fogarty, K., Lambright, D., and Corvera, S. (2003) *J Biol Chem*
223. Tall, G. G., Hama, H., DeWald, D. B., and Horazdovsky, B. F. (1999) *Mol Biol Cell* **10**, 1873-1889
224. Burd, C. G., and Emr, S. D. (1998) *Mol Cell* **2**, 157-162
225. Patki, V., Lawe, D. C., Corvera, S., Virbasius, J. V., and Chawla, A. (1998) *Nature* **394**, 433-434
226. Mills, I. G., Jones, A. T., and Clague, M. J. (1998) *Curr Biol* **8**, 881-884
227. Christoforidis, S., McBride, H. M., Burgoyne, R. D., and Zerial, M. (1999) *Nature* **397**, 621-625
228. Longva, K. E., Blystad, F. D., Stang, E., Larsen, A. M., Johannessen, L. E., and Madshus, I. H. (2002) *J Cell Biol* **156**, 843-854
229. Polo, S., Sigismund, S., Faretta, M., Guidi, M., Capua, M. R., Bossi, G., Chen, H., De Camilli, P., and Di Fiore, P. P. (2002) *Nature* **416**, 451-455
230. Komada, M., and Soriano, P. (1999) *Genes Dev* **13**, 1475-1485
231. Katzmann, D. J., Babst, M., and Emr, S. D. (2001) *Cell* **106**, 145-155
232. Babst, M., Katzmann, D. J., Snyder, W. B., Wendland, B., and Emr, S. D. (2002) *Dev Cell* **3**, 283-289
233. Babst, M., Odorizzi, G., Estepa, E. J., and Emr, S. D. (2000) *Traffic* **1**, 248-258

234. Mossessova, E., Corpina, R. A., and Goldberg, J. (2003) *Mol Cell* **12**, 1403-1411
235. Mossessova, E., Gulbis, J. M., and Goldberg, J. (1998) *Cell* **92**, 415-423
236. Renault, L., Guibert, B., and Cherfils, J. (2003) *Nature* **426**, 525-530
237. Merithew, E., Hatherly, S., Dumas, J. J., Lawe, D. C., Heller-Harrison, R., and Lambright, D. G. (2001) *J Biol Chem* **276**, 13982-13988
238. Martinez, O., and Goud, B. (1998) *Biochim Biophys Acta* **1404**, 101-112
239. Esters, H., Alexandrov, K., Constantinescu, A. T., Goody, R. S., and Scheidig, A. J. (2000) *J Mol Biol* **298**, 111-121
240. Gorvel, J. P., Chavrier, P., Zerial, M., and Gruenberg, J. (1991) *Cell* **64**, 915-925
241. Stenmark, H., Vitale, G., Ullrich, O., and Zerial, M. (1995) *Cell* **83**, 423-432
242. Gournier, H., Stenmark, H., Rybin, V., Lippe, R., and Zerial, M. (1998) *Embo J* **17**, 1930-1940
243. Christoforidis, S., Miaczynska, M., Ashman, K., Wilm, M., Zhao, L., Yip, S. C., Waterfield, M. D., Backer, J. M., and Zerial, M. (1999) *Nat Cell Biol* **1**, 249-252
244. Misra, S., and Hurley, J. H. (1999) *Cell* **97**, 657-666
245. Lawe, D. C., Patki, V., Heller-Harrison, R., Lambright, D., and Corvera, S. (2000) *J Biol Chem* **275**, 3699-3705
246. Gonzalez, L., Jr., and Scheller, R. H. (1999) *Cell* **96**, 755-758
247. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol* **276**, 307-326
248. Navaza, J. (1994) *Acta Crystallogr Sect. A* **50**, 147-163
249. Collaborative Computational Project, N. (1994) *Acta Crystallogr D Biol Crystallogr* **50**, 760-763
250. Brunger, A. T. (1992), 3.1 Ed., Yale University Press, New Haven, CT
251. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard. (1991) *Acta Crystallogr A* **47 (Pt 2)**, 110-119
252. Kraulis, P. J. (1991) *J Appl. Crystallogr.* **24**, 946-950
253. Merritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol* **277**, 505-524
254. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins Struct. Funct. Gen.* **11**, 282
255. Sharp, K., Fine, R., and Honig, B. (1987) *Science* **236**, 1460-1463
256. Chattopadhyay, D., Langsley, G., Carson, M., Recacha, R., DeLucas, L., and Smith, C. (2000) *Acta Crystallogr D Biol Crystallogr* **56 (Pt 8)**, 937-944
257. Burstein, E. S., Linko-Stentz, K., Lu, Z. J., and Macara, I. G. (1991) *J Biol Chem* **266**, 2689-2692
258. Fukui, K., Sasaki, T., Imazumi, K., Matsuura, Y., Nakanishi, H., and Takai, Y. (1997) *J Biol Chem* **272**, 4655-4658
259. Jones, S., Richardson, C. J., Litt, R. J., and Segev, N. (1998) *Mol Biol Cell* **9**, 2819-2837
260. Sasaki, T., Kikuchi, A., Araki, S., Hata, Y., Isomura, M., Kuroda, S., and Takai, Y. (1990) *J Biol Chem* **265**, 2333-2337
261. Bucci, C., Chiariello, M., Lattero, D., Maiorano, M., and Bruni, C. B. (1999) *Biochem Biophys Res Commun* **258**, 657-662
262. Chou, J. H., and Jahn, R. (2000) *J Biol Chem* **275**, 9433-9440
263. Burstein, E. S., and Macara, I. G. (1992) *Proc Natl Acad Sci U S A* **89**, 1154-1158

264. Wada, M., Nakanishi, H., Satoh, A., Hirano, H., Obaishi, H., Matsuura, Y., and Takai, Y. (1997) *J Biol Chem* **272**, 3875-3878
265. Lambright, D. G., Noel, J. P., Hamm, H. E., and Sigler, P. B. (1994) *Nature* **369**, 621-628
266. Coleman, D. E., Berghuis, A. M., Lee, E., Linder, M. E., Gilman, A. G., and Sprang, S. R. (1994) *Science* **265**, 1405-1412
267. Sunahara, R. K., Tesmer, J. J., Gilman, A. G., and Sprang, S. R. (1997) *Science* **278**, 1943-1947
268. Dumas, J. J., Merithew, E., Sudharshan, E., Rajamani, D., Hayes, S., Lawe, D., Corvera, S., and Lambright, D. G. (2001) *Mol Cell* **8**, 947-958
269. Lawe, D. C., Chawla, A., Merithew, E., Dumas, J., Carrington, W., Fogarty, K., Lifshitz, L., Tuft, R., Lambright, D., and Corvera, S. (2002) *J Biol Chem* **277**, 8611-8617
270. Merithew, E., Stone, C., Eathiraj, S., and Lambright, D. G. (2003) *J Biol Chem* **278**, 8494-8500
271. Ayad, N., Hull, M., and Mellman, I. (1997) *Embo J* **16**, 4497-4507
272. Bailly, E., McCaffrey, M., Touchot, N., Zahraoui, A., Goud, B., and Bornens, M. (1991) *Nature* **350**, 715-718
273. Prekeris, R., Klumperman, J., and Scheller, R. H. (2000) *Mol Cell* **6**, 1437-1448
274. Shisheva, A., Chinni, S. R., and DeMarco, C. (1999) *Biochemistry* **38**, 11711-11721
275. Somsel Rodman, J., and Wandinger-Ness, A. (2000) *J Cell Sci* **113 Pt 2**, 183-192
276. Cheever, M. L., Sato, T. K., de Beer, T., Kutateladze, T. G., Emr, S. D., and Overduin, M. (2001) *Nat Cell Biol* **3**, 613-618
277. Bravo, J., Karathanassis, D., Pacold, C. M., Pacold, M. E., Ellson, C. D., Anderson, K. E., Butler, P. J., Lavenir, I., Perisic, O., Hawkins, P. T., Stephens, L., and Williams, R. L. (2001) *Mol Cell* **8**, 829-839
278. Ellson, C. D., Gobert-Gosse, S., Anderson, K. E., Davidson, K., Erdjument-Bromage, H., Tempst, P., Thuring, J. W., Cooper, M. A., Lim, Z. Y., Holmes, A. B., Gaffney, P. R., Coadwell, J., Chilvers, E. R., Hawkins, P. T., and Stephens, L. R. (2001) *Nat Cell Biol* **3**, 679-682
279. Kanai, F., Liu, H., Field, S. J., Akbary, H., Matsuo, T., Brown, G. E., Cantley, L. C., and Yaffe, M. B. (2001) *Nat Cell Biol* **3**, 675-678
280. Xu, Y., Hortsman, H., Seet, L., Wong, S. H., and Hong, W. (2001) *Nat Cell Biol* **3**, 658-666
281. Song, X., Xu, W., Zhang, A., Huang, G., Liang, X., Virbasius, J. V., Czech, M. P., and Zhou, G. W. (2001) *Biochemistry* **40**, 8940-8944
282. Mu, F. T., Callaghan, J. M., Steele-Mortimer, O., Stenmark, H., Parton, R. G., Campbell, P. L., McCluskey, J., Yeo, J. P., Tock, E. P., and Toh, B. H. (1995) *J Biol Chem* **270**, 13503-13511
283. Patki, V., Virbasius, J., Lane, W. S., Toh, B. H., Shpetner, H. S., and Corvera, S. (1997) *Proc Natl Acad Sci U S A* **94**, 7326-7330
284. Gaullier, J. M., Ronning, E., Gillooly, D. J., and Stenmark, H. (2000) *J Biol Chem* **275**, 24595-24600

285. Kauppi, M., Simonsen, A., Bremnes, B., Vieira, A., Callaghan, J., Stenmark, H., and Olkkonen, V. M. (2002) *J Cell Sci* **115**, 899-911
286. Weisman, L. S., Emr, S. D., and Wickner, W. T. (1990) *Proc Natl Acad Sci U S A* **87**, 1076-1080
287. Weisman, L. S., and Wickner, W. (1992) *J Biol Chem* **267**, 618-623
288. Callaghan, J., Nixon, S., Bucci, C., Toh, B. H., and Stenmark, H. (1999) *Eur J Biochem* **265**, 361-366
289. Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S., and Rowe, A., eds), pp. 90-125, Royal Society of Chemistry, Cambridge, UK
290. La Fortelle, E. D., and Bricogne, G. (1997) *Methods Enzymol* **276**, 472-494
291. CCP4. (1994) *Acta Crystallogr D Biol Crystallogr* **50**, 760-763
292. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr A* **47**, 110-119
293. Kutateladze, T. G., Ogburn, K. D., Watson, W. T., de Beer, T., Emr, S. D., Burd, C. G., and Overduin, M. (1999) *Mol Cell* **3**, 805-811
294. Gillooly, D. J., Morrow, I. C., Lindsay, M., Gould, R., Bryant, N. J., Gaullier, J. M., Parton, R. G., and Stenmark, H. (2000) *Embo J* **19**, 4577-4588
295. Callaghan, J., Simonsen, A., Gaullier, J. M., Toh, B. H., and Stenmark, H. (1999) *Biochem J* **338 (Pt 2)**, 539-543
296. Bernstein, B. E., Hoffman, R. C., Horvath, S., Herriott, J. R., and Klevit, R. E. (1994) *Biochemistry* **33**, 4460-4470
297. Kelley, L. A., MacCallum, R. M., and Sternberg, M. J. (2000) *J Mol Biol* **299**, 499-520
298. Kutateladze, T., and Overduin, M. (2001) *Science* **291**, 1793-1796
299. Kochoyan, M., Keutmann, H. T., and Weiss, M. A. (1991) *Proc Natl Acad Sci U S A* **88**, 8455-8459
300. Lachenmann, M. J., Ladbury, J. E., Phillips, N. B., Narayana, N., Qian, X., and Weiss, M. A. (2002) *J Mol Biol* **316**, 969-989
301. Thukral, S. K., Morrison, M. L., and Young, E. T. (1991) *Proc Natl Acad Sci U S A* **88**, 9188-9192
302. Mills, I. G., Jones, A. T., and Clague, M. J. (1999) *Mol Membr Biol* **16**, 73-79
303. Mills, I. G., Urbe, S., and Clague, M. J. (2001) *J Cell Sci* **114**, 1959-1965
304. Zhu, Z., Delprato, A., Merithew, E., and Lambright, D. G. (2001) *Biochemistry* **40**, 15699-15706
305. Burton, J. L., Slepnev, V., and De Camilli, P. V. (1997) *J Biol Chem* **272**, 3663-3668
306. Moya, M., Roberts, D., and Novick, P. (1993) *Nature* **361**, 460-463
307. Collins, R. N., Brennwald, P., Garrett, M., Lauring, A., and Novick, P. (1997) *J Biol Chem* **272**, 18281-18289
308. Nuoffer, C., Wu, S. K., Dascher, C., and Balch, W. E. (1997) *Mol Biol Cell* **8**, 1305-1316

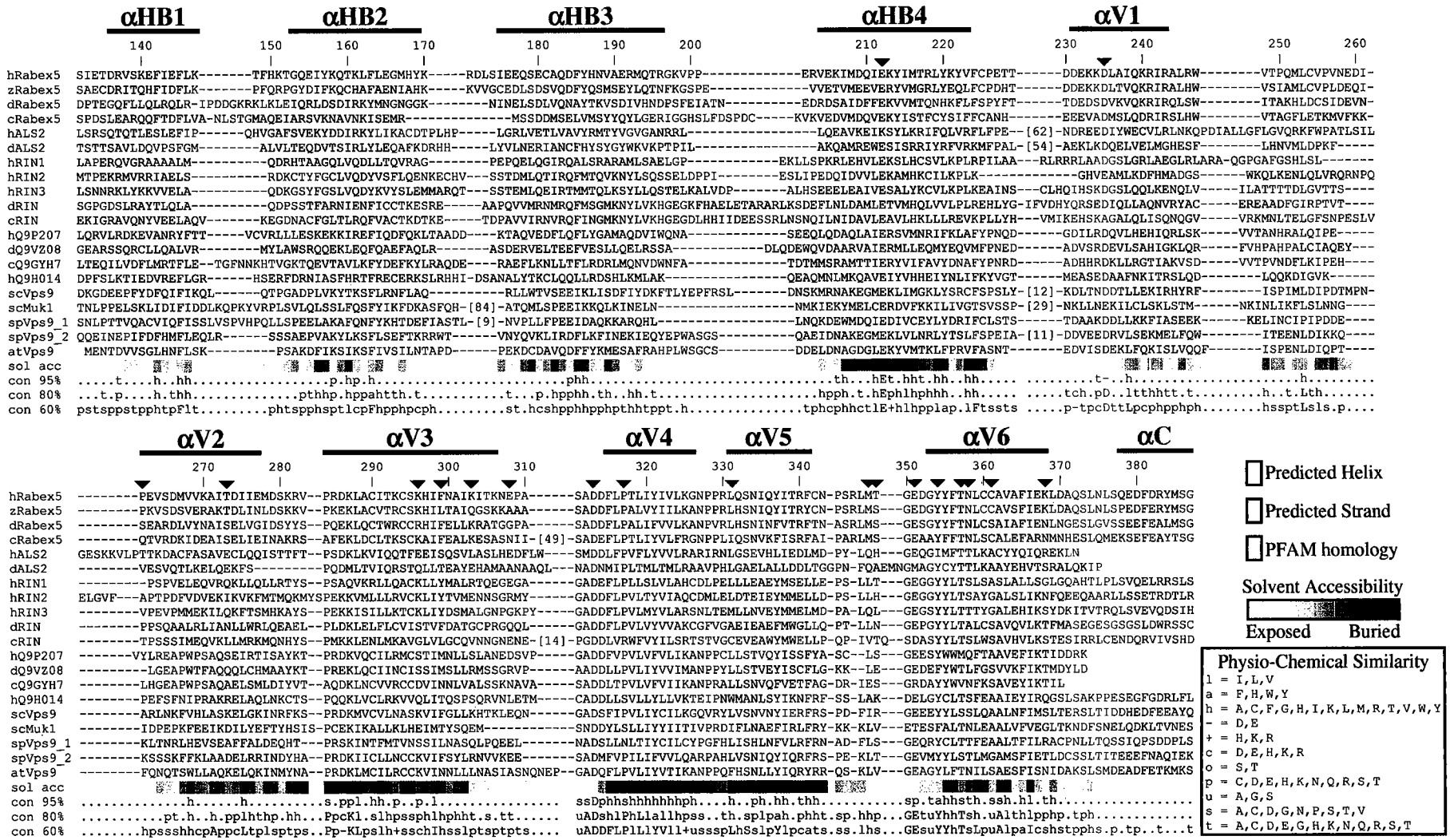
309. Muller-Pillasch, F., Zimmerhackl, F., Lacher, U., Schultz, N., Hameister, H., Varga, G., Friess, H., Buchler, M., Adler, G., and Gress, T. M. (1997) *Genomics* **46**, 389-396
310. Burstein, E. S., Brondyk, W. H., and Macara, I. G. (1992) *J Biol Chem* **267**, 22715-22718
311. Brondyk, W. H., McKiernan, C. J., Burstein, E. S., and Macara, I. G. (1993) *J Biol Chem* **268**, 9410-9415
312. Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E. L., Studholme, D. J., Yeats, C., and Eddy, S. R. (2004) *Nucleic Acids Res* **32**, D138-141
313. Letunic, I., Copley, R. R., Schmidt, S., Ciccarelli, F. D., Doerks, T., Schultz, J., Ponting, C. P., and Bork, P. (2004) *Nucleic Acids Res* **32**, D142-144
314. Esters, H., Alexandrov, K., Iakovenko, A., Ivanova, T., Thoma, N., Rybin, V., Zerial, M., Scheidig, A. J., and Goody, R. S. (2001) *J Mol Biol* **310**, 141-156
315. Kajiho, H., Saito, K., Tsujita, K., Kontani, K., Araki, Y., Kurosu, H., and Katada, T. (2003) *J Cell Sci* **116**, 4159-4168
316. Otomo, A., Hadano, S., Okada, T., Mizumura, H., Kunita, R., Nishijima, H., Showguchi-Miyata, J., Yanagisawa, Y., Kohiki, E., Suga, E., Yasuda, M., Osuga, H., Nishimoto, T., Narumiya, S., and Ikeda, J. E. (2003) *Hum Mol Genet* **12**, 1671-1687
317. Tall, G. G., Barbieri, M. A., Stahl, P. D., and Horazdovsky, B. F. (2001) *Dev Cell* **1**, 73-82
318. Han, L., Wong, D., Dhaka, A., Afar, D., White, M., Xie, W., Herschman, H., Witte, O., and Colicelli, J. (1997) *Proc Natl Acad Sci U S A* **94**, 4954-4959
319. Wang, Y., Waldron, R. T., Dhaka, A., Patel, A., Riley, M. M., Rozengurt, E., and Colicelli, J. (2002) *Mol Cell Biol* **22**, 916-926
320. Dhaka, A., Costa, R. M., Hu, H., Irvin, D. K., Patel, A., Kornblum, H. I., Silva, A. J., O'Dell, T. J., and Colicelli, J. (2003) *J Neurosci* **23**, 748-757
321. Afar, D. E., Han, L., McLaughlin, J., Wong, S., Dhaka, A., Parmar, K., Rosenberg, N., Witte, O. N., and Colicelli, J. (1997) *Immunity* **6**, 773-782
322. Devon, R. S., Helm, J. R., Rouleau, G. A., Leitner, Y., Lerman-Sagie, T., Lev, D., and Hayden, M. R. (2003) *Clin Genet* **64**, 210-215
323. Eymard-Pierre, E., Lesca, G., Dollet, S., Santorelli, F. M., di Capua, M., Bertini, E., and Boespflug-Tanguy, O. (2002) *Am J Hum Genet* **71**, 518-527
324. Gros-Louis, F., Meijer, I. A., Hand, C. K., Dube, M. P., MacGregor, D. L., Seni, M. H., Devon, R. S., Hayden, M. R., Andermann, F., Andermann, E., and Rouleau, G. A. (2003) *Ann Neurol* **53**, 144-145
325. Hadano, S., Hand, C. K., Osuga, H., Yanagisawa, Y., Otomo, A., Devon, R. S., Miyamoto, N., Showguchi-Miyata, J., Okada, Y., Singaraja, R., Figlewicz, D. A., Kwiatkowski, T., Hosler, B. A., Sagie, T., Skaug, J., Nasir, J., Brown, R. H., Jr., Scherer, S. W., Rouleau, G. A., Hayden, M. R., and Ikeda, J. E. (2001) *Nat Genet* **29**, 166-173
326. Yang, Y., Hentati, A., Deng, H. X., Dabbagh, O., Sasaki, T., Hirano, M., Hung, W. Y., Ouahchi, K., Yan, J., Azim, A. C., Cole, N., Gascon, G., Yagmour, A.,

Ben-Hamida, M., Pericak-Vance, M., Hentati, F., and Siddique, T. (2001) *Nat Genet* **29**, 160-165
327. Yamanaka, K., Vande Velde, C., Eymard-Pierre, E., Bertini, E., Boespflug-Tanguy, O., and Cleveland, D. W. (2003) *Proc Natl Acad Sci U S A* **100**, 16041-16046
328. Cerione, R. A., and Zheng, Y. (1996) *Curr Opin Cell Biol* **8**, 216-222
329. Macia, E., Chabre, M., and Franco, M. (2001) *J Biol Chem* **276**, 24925-24930
330. Whitehead, I. P., Campbell, S., Rossman, K. L., and Der, C. J. (1997) *Biochim Biophys Acta* **1332**, F1-23
331. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr D Biol Crystallogr* **54 (Pt 5)**, 905-921
332. Kawashima, T., Berthet-Colominas, C., Wulff, M., Cusack, S., and Leberman, R. (1996) *Nature* **379**, 511-518
333. Alexandrov, K., Scheidig, A. J., and Goody, R. S. (2001) *Methods Enzymol* **329**, 14-31
334. Simon, I., Zerial, M., and Goody, R. S. (1996) *J Biol Chem* **271**, 20470-20478
335. Holm, L., and Sander, C. (1993) *J. Mol. Biol.* **233**, 123-138
336. Margarit, S. M., Sondermann, H., Hall, B. E., Nagar, B., Hoelz, A., Pirruccello, M., Bar-Sagi, D., and Kuriyan, J. (2003) *Cell* **112**, 685-695
337. Nimnual, A. S., Yatsula, B. A., and Bar-Sagi, D. (1998) *Science* **279**, 560-563
338. Rossman, K. L., Worthylake, D. K., Snyder, J. T., Cheng, L., Whitehead, I. P., and Sondek, J. (2002) *J Biol Chem* **277**, 50893-50898
339. Rossman, K. L., Worthylake, D. K., Snyder, J. T., Siderovski, D. P., Campbell, S. L., and Sondek, J. (2002) *Embo J* **21**, 1315-1326
340. Snyder, J. T., Worthylake, D. K., Rossman, K. L., Betts, L., Pruitt, W. M., Siderovski, D. P., Der, C. J., and Sondek, J. (2002) *Nat Struct Biol* **9**, 468-475
341. Soisson, S. M., Nimnual, A. S., Uy, M., Bar-Sagi, D., and Kuriyan, J. (1998) *Cell* **95**, 259-268
342. Opdam, F. J., Kamps, G., Croes, H., van Bokhoven, H., Ginsel, L. A., and Fransen, J. A. (2000) *Eur J Cell Biol* **79**, 308-316
343. Mesa, R., Salomon, C., Roggero, M., Stahl, P. D., and Mayorga, L. S. (2001) *J Cell Sci* **114**, 4041-4049
344. Rodriguez-Gabin, A. G., Cammer, M., Almazan, G., Charron, M., and Larocca, J. N. (2001) *J Neurosci Res* **66**, 1149-1160
345. Mattera, R., Arighi, C. N., Lodge, R., Zerial, M., and Bonifacino, J. S. (2003) *Embo J* **22**, 78-88
346. Wang, X., Hu, B., Zimmermann, B., and Kilimann, M. W. (2001) *J Biol Chem* **276**, 32480-32488
347. Fukuda, M. (2002) *J Biol Chem* **277**, 40118-40124
348. Fukuda, M., Kuroda, T. S., and Mikoshiba, K. (2002) *J Biol Chem* **277**, 12432-12436

349. Constantinescu, A. T., Rak, A., Alexandrov, K., Esters, H., Goody, R. S., and Scheidig, A. J. (2002) *Structure (Camb)* **10**, 569-579
350. Pasqualato, S., Senic-Matuglia, F., Renault, L., Goud, B., Salamero, J., and Cherfils, J. (2003) *J Biol Chem*
351. Wilson, J. M., de Hoop, M., Zorzi, N., Toh, B. H., Dotti, C. G., and Parton, R. G. (2000) *Mol Biol Cell* **11**, 2657-2671
352. Panic, B., Perisic, O., Veprintsev, D. B., Williams, R. L., and Munro, S. (2003) *Mol Cell* **12**, 863-874
353. Wu, M., Lu, L., Hong, W., and Song, H. (2004) *Nat Struct Mol Biol* **11**, 86-94
354. Barbieri, M. A., Kong, C., Chen, P. I., Horazdovsky, B. F., and Stahl, P. D. (2003) *J Biol Chem* **278**, 32027-32036
355. Kanekura, K., Hashimoto, Y., Niikura, T., Aiso, S., Matsuoka, M., and Nishimoto, I. (2004) *J Biol Chem*
356. Clague, M.J. (1999) *Curr Biol* **9**, 258-260

APPENDIX

Appendix 1. Structure-based sequence alignment of 20 Vps9 domain proteins. Annotation and color coding are as described in the legend for Figure 29.


αHB1
αHB2
αHB3
αHB4
αV1
αV2
αV3
αV4
αV5
αV6
αC
hRabex5
zRabex5
dRabex5
cRabex5
hALS2
dALS2
hRIN1
hRIN2
hRIN3
dRIN
cRIN
hQ9P207
dQ9VZ08
cQ9GYH7
hQ9H014
scVps9
scMuk1
spVps9_1
spVps9_2
atVps9
sol acc
con 95%
con 80%
con 60%
Predicted Helix
Predicted Strand
PFAM homology
Solvent Accessibility
Exposed
Buried
Physio-Chemical Similarity
l = I,L,V
a = F,H,W,Y
h = A,C,F,G,H,I,K,L,M,R,T,V,W,Y
- = D,E
+ = H,K,R
c = D,E,H,K,R
o = S,T
p = C,D,E,H,K,N,Q,R,S,T
u = A,G,S
s = A,C,D,G,N,P,S,T,V
t = A,C,D,E,G,H,K,N,Q,R,S,T