ATYPICAL GUANINE NUCLEOTIDE EXCHANGE FACTORS FOR RHO FAMILY GTPASES: DOCK9 ACTIVATION OF CDC42 AND SMGGDS ACTIVATION OF RHOA

Brant L. Hamel

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biochemistry and Biophysics.

Chapel Hill
2010

Approved by:

Henrik Dohlman, Ph.D.

Brian Kuhlman, Ph.D.

Matthew Redinbo, Ph.D.

David Siderovski, Ph.D.

John Sondek, Ph.D.

ABSTRACT

BRANT L. HAMEL: Atypical Guanine Nucleotide Exchange Factors for Rho Family GTPases: DOCK9 Activation of Cdc42 and SmgGDS activation of RhoA
(Under the direction of John Sondek)

Rho GTPases regulate diverse cellular processes ranging from cell morphology and motility to mitosis. The activation of Rho GTPases is tightly controlled by the actions of guanine nucleotide exchange factors (GEFs). While the mechanism of canonical Dbl family exchange factors is established, both DOCK proteins and SmgGDS catalyze nucleotide exchange by distinct mechanisms. The structure of the DOCK9 GEF domain bound to Cdc42 was recently described, while no structural information on SmgGDS is available. Here, we describe a C-terminal DOCK9 fragment, soluble in bacteria, that is sufficient to catalyze nucleotide exchange on Cdc42. We also provide evidence that full-length DOCK9 is significantly more active than the minimal GEF domain, implicating the ability of other domains to contribute to the DOCK9 exchange mechanism. In contrast to the reported ability of SmgGDS to activate both Rho and Ras family GTPases, we find exclusive activation of RhoA and RhoC both *in vitro* and *in vivo*. The mechanism of SmgGDS nucleotide exchange is shown to be distinct from Dbl family GEFs and to require the presence of an intact C-terminal polybasic region on the GTPase. Using a homology model of SmgGDS, an electronegative surface patch and a highly conserved binding groove are identified that are required for the ability of SmgGDS to interact with RhoA. Our results illustrate that further structural characterization is necessary for a fuller understanding of DOCK9 exchange and that SmgGDS is able to function as a *bona fide* GEF solely for RhoA and RhoC and does so through a unique interface distinct from other known Rho family exchange factors.

This work is dedicated to my wife Sarah and my daughter Jillian, who have given me the strength and hope necessary for its completion.

ACKNOWLEGEMENTS

My sincere thanks go out to the many people who have made this work possible. The current and past members of the Sondek lab have created an exceptional work environment, not only providing technical assistance and advice, but providing genuine camaraderie and support throughout my tenure. I would like to particularly thank Cynthia Holley for many useful discussions on the DOCK family proteins and Rafael Rojas for his insights into SmgGDS function. Aurelie Gresset and Stephanie Hicks were invaluable for both their scientific enthusiasm and their scientific insights, as well as for their emotional support.

I am also indebted to a number of collaborators whose work has enhanced this project. Elizabeth Monaghan-Benson from the Burridge laboratory at UNC-Chapel Hill was indispensible in performing assays for activated RhoA in cells. Tracy Berg and her mentor Carol Williams from the Medical University of Wisconsin gave us a number of insights into SmgGDS function and helped examine the *in vivo* effects of SmgGDS mutations.

I would thank the members of my committee for their support, encouragement, and scientific advice. I am grateful to the staff of the Department of Biochemistry and Biophysics, the Biophysics program, and the Department of Pharmacology for all the behind the scenes work they do that allowed me to focus on science. I would also like to acknowledge the National Science Foundation for awarding me a graduate research fellowship.

The support, advice, training, and mentorship of John Sondek has been very influential in shaping me as a scientist, so I thank him for his unwavering commitment to conducting excellent science. Finally I wish to acknowledge the love and support of my family: my parents, grandparents, and in-laws, but most importantly that of my wife Sarah who has stood by me through all the ups and downs of life and has given me more joy then I ever thought possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3D-PSSM	Three-dimensional position-specific scoring matrix
A	Alanine
A_{600}	Absorbance at 600 nm
ACK	Activated Cdc42-associated kinase
Akt	Protein kinase B
ANKRD28	Ankryin repeat domain 28
Arf	ADP-ribosylation factor
Arp	Actin-related protein
ARM	Armadillo
ARNO	Arf nucleotide binding site opener
BAI1	Brain-specific angiogenesis inhibitor 1
Bis-Tris	Bis-(2-hydroxy-ethyl)- tris-hydroxymethylaminomethane
BLAST	Basic local alignment search tool
β-Pix	Pak-interacting exchange factor
C	Cysteine
C2	Ca^{2+} binding domain
Cdc42	Cell division control protein 42
cDNA	complementary DNA
Ced-5	Cell death abnormality 5
COS-7	African green monkey SV40 kidney fibroblast cell line
CRK	chicken tumor virus regulator of kinase

CrkL	CRK-like
CT	Carboxy-terminal construct
C-terminus	Carboxy-terminus
CV	column volume
CZH1	CDM-zizimin homology 1
CZH2	CDM-zizimin homology 2
D	Aspartic acid
DB1	DNA-binding protein 1
Dbl	Diffuse B-cell lymphoma
Dbs	Dbl's big sister
DCC	Deleted in colorectal cancer
DH	Dbl-homology
DHR-1	Dock homology region 1
DHR-2	Dock homology region 2
DNA	Deoxyribonucleic acid
DOCK	Dedicator of cytokinesis
DOCK180	180 kDa protein downstream of Crk
DTT	Dithiothreitol
E	Glutamic acid
EDTA	Ethylenediamine tetra-acetic acid
EGFR	Epidermal growth factor receptor
ELMO	Engulfment and cell motility
ErbB2	Erythroblastic viral oncogene homolog 2

ERK	Extracellular regulated kinase
ERM	Ezrin, radiixin, moesin family
F	Phenylalanine
FL	Full length
G	Glycine
G protein	Guanine nucleotide binding protein
GAP	GTPase accelerating protein
GBD	GTPase binding domain
GDI	Guanine dissociation inhibitor
GDP	Guanosine 5’-diphosphate
GEF	Guanine nucleotide exchange factor
GPCR	G-protein coupled receptor
Grb2	Growth factor receptor-bound protein 2
GST	Glutathione S-transferase
GTP	Guanosine 5’-triphosphate
GTPase	Guanosine triphosphatase
GTPγS	Guanosine 5'-(γ-thio)triphosphate
H	Histidine
His	Histidine
HA	Hemagglutinin
Hck	Hemopoietic cell kinase
HEK293T	Human embryonic kidney 293T cell line
HeLa	Henrietta Lacks cervical cancer cell line

HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIV	Human immunodeficiency virus
IPTG	Isopropyl-β-D-thiogalactopyranoside
ITSN	Intersectin
JNK	C-Jun N-terminal kinase
K	Lysine
KAP3	Kinesin superfamily associated protein 3
K_{cat}	Enzyme turnover number
kDa	Kilodalton
K_i	Dissociation constant for inhibitor binding
KIF	Kinesin family member
K_m	Michaelis constant
I	Isoleucine
IL4	Interleukin-4
L	Leucine
LB	Luria broth
LIM-kinase	LIN-11, Isl1, and MEC-3 related domain containing kinase
λ_{em}	Emission wavelength
λ_{ex}	Excitation wavelength
LIC	Ligation-independent cloning
MANT	N-methylanthraniloyl
MAPK	Mitogen activated protein kinase
mDia	Mammalian homologue of *Drosophila* diaphonous

MEK	MAPK/ERK kinase
Mg^{2+}	Magnesium ion
$MgCl_2$	Magnesium chloride
MOCA	Modifier of cell adehesion
mRNA	Messenger ribonucleic acid
MSA	Multiple sequence alignment
Mss4	Mammalian suppressor of Sec4
N	Asparagine
NaCl	Sodium cholride
Nck	Noncatalytic region of tyrosine kinase
Nef	Negative regulatory factor
NF-κβ	Nuclear transcription factor κβ
NIH-3T3	National Institutes of Health 3T3 cell line
NMR	Nuclear magnetic resonance
N-terminus	Amino-terminus
P	Proline
Pak	p21 activated kinase
PAR	Partitioning defective protein
PBR	Polybasic region
PBEQ	Poisson-Boltzman equation
PC12	Rat pheochromocytoma cell line 12
PCR	Polymerase chain reaction
PDB	Protein databank

PDZ	PSD95, DlgA, zo-1 homologous domain
PEG	Polyethelyne glycol
PI3K	Phosphoinositide 3-kinase
PH	Pleckstrin homology
P loop	Phosphate binding loop
PVDF	Polyvinyldene flouride
Q	Glutamine
Rab	Ras homologue from rat brain
RBD	Rho binding domain
Ral	Ras-like
RalGDS	Guanine nucleotide dissociation stimulator for Ral
Ran	Ras-related nuclear protein
Rap	Ras-related protein
Ras	Rat sarcoma virus protein
RCC1	Regulator of chromosome condensation 1
Rho	Ras homologous
ROCK	Rho-associated coiled-coil containing protein kinase
S	Serine
SDS-PAGE	Sodium dodecylsulfate-polyacrylamide gel electrophoresis
SERp	Surface entropy reduction program
SH2	Src-homology 2
SH3	Src-homology 3
siRNA	Small interfering ribonucleic acid

SMAP	SmgGDS associated protein
SmgGDS	Guanine nucleotide dissociation stimulator for smg p21
SopE	Salmonella outer protein-E
Sos	Son of sevenless
Src	Sarcoma viral oncogene homologue
T	Threonine
TEV	Tobacco etch virus
Tiam1	T-cell invasion and metastasis factor
TNFAIP1	TNF-α induced protein 1
Tris	Tris-hydroxymethylaminomethane
V	Valine
VPS9	Vacuolar protein sorting defect 9
W	Tryptophan
WASP	Wiskott-Aldrich syndrome protein
Wnt	Wingless-type
XPLN	Exchange factor found in platelets and in leukemic and neuronal libraries
Y	Tyrosine

CHAPTER 1: INTRODUCTION

Part 1: G proteins and their activation by GEFs

G proteins have diverse biological roles

Guanine nucleotide binding proteins, or G proteins, are defined by their abilities to bind the nucleotides guanosine 5'- diphosphate (GDP) and guanosine 5'-triphospate (GTP) (*1*). G proteins are found in diverse biological settings ranging from elongation factors which regulate proper translation of protein from mRNA (*2*) to Gα-subunits of the heterotrimeric G proteins which transduce downstream signals after activation by G protein-coupled receptors at the cell membrane (*3*). The present work will focus on the members of the Ras superfamily of small GTPases which act as molecular switches.

Members of the Ras superfamily of small GTPases are monomeric proteins that tightly bind to guanine nucleotides. GTPases have the intrinsic capacity to hydrolyze the gamma phosphate of GTP to yield GDP, but typically do so inefficiently without the help of accessory proteins known as GTPase accelerating proteins (GAPs). Their ability to act as switches derives from the fact that these proteins adopt distinct conformations when the GTPase is bound to GTP (the active conformation capable of interacting with downstream signaling partners) or GDP (the inactive conformation). These distinct states are driven by two conformationally flexible regions of the protein, switch 1 and switch 2 (*1*). GTPases are activated by proteins known as guanine nucleotide exchange factors (GEFs) that manipulate the nucleotide binding pocket to catalyze the removal of GDP prior to binding of GTP which activates signaling.

The Ras superfamily can be subdivided into five distinct families of GTPases: Ran, Rab, Arf, Ras, and Rho, all with unique biological roles (*4*) (**Fig. 1**). The Ran family consists of a single member, Ran, which is involved in regulating nucleocytoplasmic transport. The GAP and GEF regulating Ran are localized in the cytoplasm and nucleus respectively (*5*). Therefore, the active pool of Ran is found only in the nucleus. In the nucleus, active Ran causes dissociation of incoming nuclear cargo bound to importins and causes the association of cargo with exportins to promote nuclear export. In the cytoplasm, cargo is released due to the presence of inactive Ran, which is recycled back to the nucleus to be reactivated (*6*).

The Rab family is the largest member of the Ras superfamily consisting of over 60 different proteins (*4*). Rab proteins control the intracellular transport of proteins between different organelles and vesicular compartments. Rab proteins play roles in vesicle budding, association with molecular motors, vesicle fusion, and vesicle uncoating, thus regulating almost every aspect of membrane trafficking inside cells (*7*). The Arf family GTPases are also involved in membrane trafficking through their ability to initiate the formation of coat protein complexes during vesicular formation (*8*). Additionally, Arf GTPases have cellular roles related to controlling the actin cytoskeleton through their ability to attract proteins that regulate phosphoinositide levels as well the ability to signal to downstream Rho family GTPases (*9*).

The Ras family of GTPases is the second largest monomeric GTPase family with 36 members and is also one of the most intensely studied families (*4*). Ras GTPases are activated by a wide variety of extracellular stimuli and lead to the downstream activation of proteins that control cell growth, differentiation, and survival. The most well studied

Ras family GTPases are the Ras GTPases (N-Ras, H-Ras, and K-Ras), the Ral GTPases (RalA and RalB) and the Rap GTPases (Rap1A and Rap1B) (*10*). The biology of the Ras and Rap GTPases are discussed below.

Ras GTPases can be activated by receptor tyrosine kinases such as the epidermal growth factor receptor (EGFR). Upon stimulation of EGFR, autophosphorylation occurs and leads to the recruitment of the Grb2 adaptor protein through interaction of phosphotyrosine with the Grb2 SH2 domain. The SH3 domain of Grb2 is then able to localize the Ras exchange factor Sos1 to the plasma membrane where it can activate Ras to signal to its downstream effectors (*11*).

Ras is best known for its activation of the mitogen-activated protein kinase (MAPK) cascade (*11*). Activated Ras brings the Raf serine/threonine kinase to the membrane where it is converted to an active state. Raf then phosphorylates MEK kinase, which phosphorylates ERK kinase, which phosphorylates a variety of proteins that control cell growth. Ras can also activate a number of additional effectors including RalGDS, Tiam1, and phosphoinositide 3-kinases (PI3Ks) that lead to downstream activation of Akt, Rac, and NF-κB (*10*).

The role of the Ras proteins in oncogenesis was appreciated from an early date. The first family members, H-Ras and K-Ras, were identified as homologues of transforming sequences from Harvery and Kirsten sarcoma viruses (*12*). Mutations that activate Ras are found in 30% of all human cancers and up to 90% in some particular cancers such as pancreatic (*13*). The most commonly found mutations are G12V and Q61L which render the GTPase insensitive to the action of GAPs leaving it constitutively

activated. Downstream targets of Ras such as Raf and other kinases are popular targets for anticancer drug development (*13*).

Originally the Ras proteins were thought to be functionally redundant, but important differences in the localization and signaling of the Ras proteins have been found. The Ras proteins are composed of a conserved nucleotide binding domain and a C-terminal hypervariable region that is highly divergent between Ras isoforms. The hypervariable region and its associated lipid modifications regulate the localization and membrane binding of Ras proteins and thus the identity of their downstream targets. All Ras isoforms are farnesylated in the endoplasmic reticulum at a C-terminal CAAX motif (C, cysteine; A, aliphatic amino acid; X, serine or methionine), but H-Ras and N-Ras are additionally palmitoylated in the Golgi apparatus at a cysteine upstream of the CAAX motif (*10-11*). In contrast the major K-Ras isoform (K-Ras4B, henceforth referred to as K-Ras) is not palmitoylated, but contains a stretch of polybasic residues in its C-terminus upstream of the CAAX motif. It is thought that N-Ras and H-Ras are recruited to caveolin-rich lipid raft microdomains in the cell membrane while K-Ras associates more broadly with the plasma membrane (*10*).

The Rap proteins were originally identified as both showing homology to Ras proteins and reversing the phenotype of K-Ras transformed cells (*14*). Rap proteins are able to bind to Raf kinase (and most other Ras effectors) as their amino acid sequence is identical to the Ras proteins in the effector binding region (*14*). However Rap does not activate Raf, likely because of differences in subcellular localization of Rap and Ras proteins (*15*). Rap proteins contain a strong polybasic region similar to K-Ras, but instead of being farnesylated at their CAAX motif, a geryanylgeranyl group is added

(*16*). In addition to their role opposing Ras function, Rap proteins have key roles in controlling cell adhesion (*17*) and signaling through integrins (*14*). Rap has also been shown to localize to the nucleus where it can regulate the expression of other proteins (*18*).

Finally, the Rho GTPase family contains 22 members (*19*). The Rho GTPases are most known for their ability to modulate the actin cytoskeleton. The most well characterized members of the family are RhoA, Cdc42, and Rac1. RhoA was initially found to cause the formation of actin stress fibers while Rac1 and Cdc42 promoted the formation of lamellipodia and filopodia respectively (*20*). Since that time the Rho GTPases have been shown to be vital in numerous additional cellular processes including: cell cycle regulation (*21*), modulation of gene expression (*22*), exocytosis (*23*), neurite outgrowth (*24*), proliferation (*20*), and establishing cell polarity (*20*). The specific biological roles of Cdc42 and Rho are explored below.

Cdc42 exists in two splice variants, a ubiquitously expressed isoform known as placental Cdc42 and a brain specific isoform (*19*). The isoforms differ in their C-terminal polybasic region with the brain isoform having less basic residues. The brain specific isoform has been implicated in regulating neuronal differentiation during embryonic development of the rat neocortex (25). However, most studies have focused on the ubiquitously expressed placental isoform. Cdc42 has multiple binding partners and plays a pivotal role in cell signaling, regulating diverse cellular processes important not only for cell morphology and migration, but also endocytosis and establishment of cellular polarity. Activated Cdc42 binds to WASP proteins to activate the Arp2/3 complex. Arp2/3 activation leads to actin polymerization and the formation of filopodia

(26). Cdc42 is also important in endocytosis through its ability to bind ACK. Cdc42 binding to ACK inhibits ACK interactions with clathrins, thus allowing clathrins to interact with adaptor proteins that promote endocytosis (26). Cdc42 also interacts with PAR6 and induces cell polarity by localizing the PAR6/PAR3 complex to the apical domain of polarized epithelial cells (27). Cdc42 additionally binds to Paks, activates their ability to phosphorylate downstream targets, thus activating multiple pathways including MAPK activation and cytoskeleton rearrangements (28).

RhoA, B, and, C contain high sequence homology (85%) with the most divergence in the C-terminal polybasic region (*19*). RhoA and RhoC contain numerous basic residues in this region; in contrast RhoB has minimal basic character with only two basic residues. While RhoA and RhoC are both geranylgeranylated on their C-terminal CAAX motif, RhoB can be farnesylated or geranylgeranylated (*19*). In addition, RhoB is also palmitoylated on two cysteines located directly proceeding the CAAX box. Palmitoylation is known to be critical for the tumor-suppressive and apoptotic functions of RhoB. (29). In addition, differences in the post-translational processing also affect the subcellular localization of the Rho isoforms. For example, RhoA and RhoC are known to associate with the plasma membrane, while RhoB localizes to endosomal membranes (30). Consistent with its localization to endosomes, RhoB is important for regulation of endocytic trafficking (27).

Activation of RhoA leads to the binding and activation of two major effectors, ROCK (Rho kinase) and mDia (mammalian homologue of *Drosophilia* diaphanous). Binding of activated RhoA to mDia relieves its autoinhibition and induces actin polymerization. RhoA induced activation leads ROCK to phosphorylate multiple

downstream targets including myosin light chain and LIM-kinase, resulting in the cross-linking of actin by myosin. The combined effect of ROCK and mDia activation is the induction of actomyosin bundles that comprise stress fibers (31).

RhoB and RhoC also activate ROCK and diaphanous proteins and the Rho isoforms were originally thought to be functionally redundant. Thus, many early studies only looked at RhoA or used pharmacological inhibitors that affected all Rho isoforms equally. However, differences in activation and effector binding between the Rho isoforms have begun to emerge. RhoB has been shown to interact directly with TNFAIP1 (TNF-α induced protein 1) to promote apoptosis in HeLa cells (32). In addition, RhoB, but not RhoA, binds to the DB1 transcription factor to inhibit its activity (33). A Dbl family GEF, XPLN, has been shown to activate RhoA and RhoB, but not RhoC (34). RhoC, but not RhoA or RhoB, has been shown to specifically activate formin-like 2 to cause amoeboid invasive cell motility in cancer cells (35). Thus, it is clear that there are important functional differences between the Rho isoforms.

Unlike the Ras family, activating mutations in Rho family members have not been observed in cancers. However, upregulated expression of both Rho GTPases and Rho GEFs have been observed in multiple tumors implicating Rho family signaling as important for cancer pathophysiology (36). Cdc42 overexpression has been shown to correlate with tumor progression in breast and testicular cancers, but the importance of Cdc42 in cancer progression has not been clearly illustrated (36). In contrast, RhoA and RhoC have been demonstrated to have oncogenic roles in a number of different cancers. Both RhoA and RhoC mRNA expression is significantly upregulated in ovarian (37), breast (38), esophageal squamous cell (39), and colorectal (40) cancers. The amount of

active RhoA has also been shown to be increased in cancers (41). Interestingly, the expression of RhoA and RhoC appear to be inversely correlated in a number of studies suggesting they have distinct roles in different stages of tumor progression (42-43). RhoC contributes to the metastatic potential of lung cancer by upregulating matrix metalloproteinases and increasing cell motility (44). In esophageal squamous cell carcinoma cells, RhoA appears linked to tumor growth while RhoC is associated with tumor metastasis (45). However, the role of RhoA may vary by cell type as RhoA has been linked to metastasis in colorectal cancers (41) and cell growth and proliferation in other cancers (45-46). Unlike RhoA and RhoC, RhoB has been shown to function as a tumor suppressor and its expression is often decreased in cancer cell lines (36).

The structure of a GTPase allows it to act as a molecular switch during a nucleotide exchange cycle

The ability of a GTPase to bind to guanine nucleotides derives from conserved structural elements found in all GTPases (**Fig. 2**). The β and γ phosphates of bound nucleotide are stabilized by a phosphate binding loop (P loop) motif, GxxxxGKS/T while the nucleotide base interacts with a N/TK/QxD motif after the β5 sheet. Specificity for guanine results from two motifs, the DxxG motif at the beginning of switch 1 and the SAK/LT motif after the β6 sheet (*1*). The backbone of a conserved threonine (T37 in RhoA) contacts the γ phosphate when GTP is bound. A Mg^{2+} cofactor is also essential for high affinity nucleotide binding. While, the core GTPase structure consists of five α-helices and six β-sheets, different family members possess additional regions (*1*). For example, Arf proteins have an extended N-terminus that mediates membrane interactions

(*8*) while Ran has an extended C-terminus that plays roles in nuclear transport (*5*). We will focus on the Rho GTPases, which have an inserted helical region, α-insert, between the β5 sheet and α4 helix.

GTPase signaling originates from two regions, switch 1 and switch 2, that alter their conformation depending upon the bound nucleotide. The hydrolysis of the γ-phosphate from GTP by the GTPase releases interactions with highly conserved residues T37 of switch 1 and G62 of switch 2 (*1*). This leads to a "relaxation" of the switch regions that alters their conformation, while the rest of the protein remains stable (**Fig. 2B**). Signaling to downstream effectors only occurs in the GTP-bound state. Thus, effector molecules make extensive interactions with the switch regions of the Rho GTPases to discriminate between the active and inactive GTPase conformations (47).

A suite of proteins are responsible for controlling the activity of Rho GTPases inside of the cell (**Fig. 3**). GTPases typically hydrolyze the γ-phosphate of bound GTP very slowly. The rate of hydrolysis is increased dramatically by GTPase accelerating proteins (GAPs). GAPs provide an "arginine finger" that stabilizes the transition state of the hydrolysis reaction, thus increasing the rate of γ-phosphate hydrolysis (48). GAPs interact with switch 1, switch 2, and the P-loop on the GTPase (47). Guanine nucleotide dissociation inhibitors (GDIs) extract prenylated GTPases from membranes and sequester them in an inactive GDP-bound state in the cytosol. Thus, they inhibit Rho GTPases both by directly stabilizing the switch regions to prevent nucleotide release, but more importantly by altering their localization to prevent interactions with GEFs or downstream effectors (49). GDIs interact with switch 1, switch 2, α2, α3, β4, and the prenylated C-terminal tail of the GTPase (47). Guanine nucleotide exchange factors

(GEFs) are responsible for the activation of GTPases. GEFs function to stabilize the nucleotide-free state of the GTPase that leads to exchange of GDP for GTP. The mechanism of nucleotide exchange catalyzed by GEFs is reviewed in detail below.

The mechanisms of GTPase activation by GEFs

GEFs work primarily by stabilizing a nucleotide-free GTPase state. Initial binding of the GEF to the nucleotide-bound GTPase is thought to result in a low affinity ternary complex. Conformational rearrangements of the GTPase lead to the expulsion of nucleotide and a high affinity binary GTPase/GEF complex. The reaction is reversed by the association of free nucleotide with the GTPase/GEF complex and the dissociation of the GEF (*1*). The intracellular concentration of GTP is approximately 10-fold higher than that of GDP, so GEFs drive the formation of the activated GTP-bound GTPase state. Typically, the overall equilibrium constant for the reaction is close to 1 so that binding of the GEF is readily reversible (50). In general, GEFs interact with the switch regions, residues important for phosphate binding such as the P-loop, and areas involved in Mg^{2+} coordination (*1*). Usually, the binding sites of the nucleotide base and the ribose group are not disturbed by GEF interaction so that incoming nucleotide may still partially bind before the GEF is dissociated (51). Molecular details of GEF activation have been elucidated for many members of the Ras superfamily.

The first mechanism of Ras superfamily activation by a GEF was revealed by the structure of an active Sos fragment bound to nucleotide-free H-Ras (51). There is a large interface between Sos and Ras including interactions with both switch regions as well as the phosphate binding loop. Compared to a structure of nucleotide-bound Ras, switch 1

is reorientated in a conformation which opens the nucleotide binding site. An α-helix from Sos is inserted into the nucleotide binding site causing L938 of Sos to occlude the site of Mg^{2+} binding and E942 to occlude the binding of the α-phosphate. Switch 2 is also reorientated such that A59 would occlude Mg^{2+} binding and E62 is able to form an electrostatic interaction with K16, a residue that coordinates phosphate in nucleotide-bound Ras. The biological activation of Ras by Sos is more complicated as active Ras binds to an allosteric site on Sos to increase its GEF activity (52). Other Ras family GEFs have the same basic mechanism of exchange as they have homologous nucleotide exchange domains.

The structure of Arf1 bound to a Sec7 Arf GEF domain illustrated another mechanism of nucleotide exchange (53). Sec7 interacts with Arf1 through switch 1 and switch 2. The major contribution to nucleotide exchange comes from the insertion of a "glutamic acid finger" of Sec7 into the nucleotide binding site. Residue E97 of the Sec7 domain is positioned in a manner which not only sterically occludes Mg^{2+} binding, but also forms an ion pair with K30 of Arf, a residue that normally binds to the β phosphate of a bound nucleotide. Additional changes in the switch regions of Arf also contribute to the destabilization of bound nucleotide (53).

The structure of Ran bound to its GEF, Rcc1, was found to contain a polyanion in the phosphate binding site and thus may model an intermediate GTPase/GEF/nucleotide ternary complex more than the GTPase/GEF binary complex (54). Nonetheless, several important facets of the nucleotide exchange mechanism can be derived. Ran binding to its GEF, Rcc1, involves residues in the phosphate binding loop, switch 2, and the α3 and α4 helices. Unlike most other GTPase/GEF interactions, switch 1 is not contacted. The

major structural change observed was a shift in the phosphate binding loop that would result in the guanine base of a bound nucleotide clashing with residues normally used for base coordination. It is believed that additional disorganization of the phosphate binding loop would be observed in the absence of polyanion stabilization (54).

The Rab GTPase family is activated by a number of structurally and mechanistically distinct GEFs. The structure of Mss4 bound to Rab8 illustrated an unconventional mechanism of nucleotide exchange that relied on the partial unfolding of the GTPase to disrupt the nucleotide binding pocket (55). Mss4 forms an intermolecular β-sheet with the β2 sheet of Rab8 stabilized by additional interactions with switch 1. The interaction of Mss4 with Rab8 results in complete unfolding of the nucleotide binding pocket, in contrast to other exchange factors that typically do not alter the binding site of the guanine base or ribose group. One consequence of this dramatic change is the very slow binding of GTP to the Rab8/Mss4 binary complex, which limits the rate of the overall exchange reaction significantly compared to other Rab GEFs (55).

VPS9 domain-containing exchange factors activate Rab GTPases through a mechanism similar to the Sec7 domain activation of Arf, the insertion of an "aspartic acid finger" into the nucleotide binding site, which both occludes Mg^{2+} binding and provides electrostatic repulsion to the β phosphate (56). The structure of the catalytic coiled-coil domain of the Sec2p GEF bound to the Sec4p Rab GTPase did not show the GEF to directly interfere with nucleotide binding, but to cause reorientation of the switch regions and phosphate binding loop in manner inconsistent with nucleotide binding (57). Sec2p was seen to interact with both switch 1 and switch 2 of the GTPase. The interactions alter the GTPase conformation of I50 such that it occludes Mg^{2+} binding and the

conformation of F45 so that it is not able to stabilize the nucleotide binding pocket. In addition, the phosphate binding loop is reorientated in a position that would occlude nucleotide binding. Overall, the Rab GEFs illustrate multiple mechanisms to catalyze exchange including altering the conformation of the GTPase in a manner incompatible with nucleotide binding, directly inserting a residue that interferes with nucleotide binding, or disrupting the G protein fold.

Dbl family GEFs activate Rho GTPases the Dbl-homology (DH) and pleckstrin-homology (PH) domains which are almost invariably found in tandem as a DH/PH cassette. Numerous additional domains found in Rho GEFs are thought to control subcellular localization, regulation of GEF activity, and partnering to specific downstream effector pathways (58). The mechanism of Dbl family GEF activation of Rho GTPases has been elucidated by numerous structural studies. The structure of Rac1 bound the DH/PH domain of Tiam1 showed extensive interactions of both switch 1 and switch 2 with Tiam1 (59). The binding of Mg^{2+} is occluded by the altered position of A59 similar to that observed in the Sos/Ras complex and E62 alters conformation to interact with K16 of the phosphate binding loop. Switch 1 conformation is changed such that F28 can no longer stabilize the guanine base and I33 would sterically interfere with binding of the ribose base of bound nucleotide. For Rac1/Tiam1 only the DH domain contacted the GTPase (59). In contrast, structures of Cdc42/Dbs (60) and RhoA/PDZ-RhoGEF (61) showed contacts between both the DH and PH domains and the GTPase. Thus, while the DH domain invariantly makes essential contacts to reorient the nucleotide binding site, the PH domain may also contribute to selectivity and nucleotide exchange in some instances.

Additional proteins that do not contain the traditional DH/PH cassette act as Rho GEFs. One example of atypical Rho GEFs are bacterial toxins, such as SopE, that activate Rho GTPases during the process of invading host cells. The structure of the SopE toxin bound to Cdc42 established its role as a GEF (62). Despite no structural homology to DH/PH domains, SopE stabilizes a nucleotide-free conformation of Cdc42 that is remarkably similar to the nucleotide-free conformation of GTPases in complex with Dbl family proteins. The binding of Mg^{2+} is occluded by the switch 2 residue A59 and F28 is flipped out of the nucleotide binding pocket so that it can no longer stabilize nucleotide binding. Equivalent changes were seen for Rac1/Tiam1 leading to the hypothesis that diverse Rho GEFs have converged to stabilize a common nucleotide-free state (62).

Two additional families of GEFs that do not structurally resemble Dbl family proteins have been reported to activate Rho GTPases, and are the focus of this dissertation. SmgGDS is an armadillo (ARM) repeat containing protein that is reported to activate a diverse set of GTPases including Rho family members Rac1, RhoA, and Cdc42. The nucleotide exchange mechanism of SmgGDS is entirely unknown. The DOCK family GEFs activate Rac1 and Cdc42. Although undetermined at the start of this project, the mechanism by which DOCK9 activates Cdc42 was recently elucidated. DOCK family GEFs and SmgGDS are reviewed in detail below.

Part 2: The DOCK family of atypical exchange factors.

Identification of DOCK family proteins as Rho GEFs

The DOCK (dedicator of cytokinesis) proteins are a family of Rho GEFs that are structurally and mechanistically distinct from the more numerous Dbl family Rho GEFs. There are 11 DOCK members found in humans (**Fig. 4**). DOCKs 1-5 activate Rac, but require binding to an accessory protein, ELMO (engulfment and cell motility), for their GEF activity. DOCKs 9-11 activate Cdc42 in an ELMO-independent fashion, while DOCKs 6-8 may activate both Rac and Cdc42. To date, the biological roles, regulation, and exchange mechanism of DOCK proteins have been partially elucidated using multiple methods including overexpression, co-immunoprecipitation, biochemical, and structural analyses.

DOCK1 (also known as DOCK180) was originally identified as a 180 kDa protein that bound to the CRK adaptor protein. DOCK1 alters the morphology of NIH-3T3 cells when artificially localized to the plasma membrane, inducing a cell flattening phenotype that is similar to the constitutively active Rac1 phenotype (63). DOCK1 overexpression was subsequently found to increase the amount of active Rac1 in 293T cells and DOCK1 pulled-down Rac1 in the presence of EDTA, implicating it as a Rac1 specific GEF (64).

Soon after DOCK1 identification, ELMO was shown to be vital for DOCK1 catalyzed exchange. ELMO was first suggested to interact with DOCK1 due to genetic interactions of the homologous genes in *C. elegans* (65-66). The ELMO protein directly interacts with DOCK1 (67) and was subsequently shown to be required for the ability of DOCK1 to cause nucleotide exchange on Rac1 (68).

DOCK2 and DOCK3 were identified soon after DOCK1 as tissue-specific GEFs for Rac1. DOCK2 is expressed solely in hematopoietic cells (69) whereas DOCK3 (also

known as MOCA; modifier of cell adhesion) is expressed only in the brain. DOCK3 was originally identified as a presenilin binding protein, and subsequently found to activate Rac1 (70-71). The identification of three homologous proteins with the ability to activate Rac1 suggested that they comprised a new family of GEFs, but the size and the diversity of the family was not immediately appreciated.

The DOCK protein family was expanded with the identification of DOCK9 (also known as Zizimin1) as a Cdc42 activator (72). DOCK9 was isolated from cell lysates as a 220 kDa protein which bound to nucleotide-depleted, but not GTPγs-bound Cdc42. DOCK9 stimulated nucleotide exchange on Cdc42 and induced filopodia formation when overexpressed in NIH-3T3 cells. Sequence analysis of DOCK9 showed that it had two significant regions of homology with DOCK1. Using these regions of homology, a superfamily of related proteins was identified by BLAST analysis of sequence databases (72). Shortly thereafter, another group identified essentially the same two homologous regions, DOCK homology region 1 and 2 (DHR-1 and DHR-2). Analyzing sequence similarities among the homologous regions, the DOCK family was subdivided into four subfamilies: **DOCK-A**, consisting of DOCK1 (DOCK180), DOCK2, and DOCK5; **DOCK-B**, consisting of DOCK3 (MOCA) and DOCK4; **DOCK-C** consisting of DOCK6, DOCK7, and DOCK8; and **DOCK-D** consisting of DOCK9 (zizimin), DOCK10, and DOCK11 (73).

Recent structural studies have identified the domains that comprise the conserved DOCK homology regions (**Fig. 5**). All DOCK proteins contain a DHR-1 implicated in membrane association (74-75), a DHR-2 essential for GEF activity (76-77), and a less conserved intervening sequence. The structure of the DOCK1 DHR-1 revealed a C2

lipid binding domain that can specifically interact with phosphatidylinositol-(3,4,5)-triphosphate (78). The structure of the DOCK9 DHR-2 revealed a domain which mediates homodimerization and the catalytic GEF domain (79). Bioinformatic analyses predict that the intervening sequence between the C2 and dimerization domains is an array of armadillo (ARM) repeats. In addition to the central conserved domains shared across the whole family, the DOCK-A and DOCK-B subfamilies have SH3 domains in their N-terminus, while the DOCK-D subfamily contains a PH domain. No additional domains have been identified in the DOCK-C subfamily.

Biological functions of DOCKs 1-5

The DOCK1/ELMO/CRK complex is involved in the phagocytosis of apoptotic cells through a process involving Rac1 (80-82). Integrin activation leads to phosphorylation of DOCK1 and increases its association with CRK (83). DOCK1 has also been linked to the binding of other protein scaffolds such as Nck-2 (84). The BAI1 GPCR is also an upstream activator of DOCK1. BAI1 promotes engulfment of apoptotic cells and directly interacts with DOCK1 (85). A separate pathway implicates the Arf family GEF ARNO and Arf-6 as upstream activators of DOCK1. Expression of ARNO increases active Rac and causes lamellipodia formation, but when ARNO is co-expressed with a catalytically dead DOCK1 no changes are observed (86). ARNO is also found in multiprotein complexes with DOCK1 making it a point of convergence for Arf family and Rho family GTPase signaling (87).

The downstream effects of DOCK1 signaling are diverse and appear to be both cell-type and stimulus dependent. DOCK1 and ELMO have been shown to be

overexpressed in invasive glioma cells and inhibition of DOCK1/ELMO significantly reduced the invasive potential of these cells suggesting that the DOCK1 and ELMO may be oncogenes (88). DOCK1 plays a critical role in muscle development through its ability to regulate myoblast fusion and a DOCK1-null mouse is severely impaired in skeletal muscle formation (89). DOCK1 has been shown to be involved in axon guidance through its activation by the netrin receptor deleted in colorectal carcinoma (DCC) (90). DOCK1 plays a critical role in the phagocytosis of apoptotic cells (80-82). All of the diverse functions of DOCK1 are believed to result from its ability to activate Rac1.

DOCK1 has been reported to bind to phosphatidylinositol-(3,4,5)-triphosphate that would target it to membranes upon activation of PI3Ks. This activity has been suggested to require either the C-terminus of the protein (91) or the conserved DHR-1 (74). DOCK2 translocation to cell membranes has also been shown to require phosphatidyl inositol (3,4,5) triphosphate and to be further stabilized by interactions with phosphatidic acid (92). However, DOCK1 does not solely signal at the plasma membrane. Experiments measuring endogenous protein levels show significant nuclear localization of both DOCK1 and ELMO in HeLa cells (93). It is likely that DOCK proteins can signal in multiple cellular compartments to control a variety of Rho family dependent pathways.

DOCK1 is autoinhibited in its basal state through interactions of the SH3 domain and the adjoining region with the DHR-2 (94). Autoinhibition is relieved by truncation of the SH3 domain or by ELMO binding to DOCK1. Interestingly, DOCK1 has also been shown to bind the ankryin repeat protein ANKRD28. Interaction of ANKRD28 with DOCK1 promotes the stability of focal adhesions (95). The DOCK1/ANKRD28

interaction required the SH3 domain of DOCK1 and was competitive with ELMO binding. DOCK1 clearly has multiple binding partners which govern the physiological outcome of DOCK1 signaling.

DOCK2 expression is limited to hematopoetic cells where it activates Rac1 and binds to the CrkL (CRK-like) adaptor protein (69,96). Consistent with its expression profile, DOCK2 is important in immune function and therefore is implicated in a variety of disease states. For example, DOCK2 is essential for proper chemotaxis of lymphocytes and a DOCK2 knockout mouse shows multiple immunological defects (97). DOCK2 knockout mice also show markedly reduced allograft rejection with mismatched tissues which suggests DOCK2 could be a target for drugs to alleviate transplant rejection (98). DOCK2 is also required for proper leading edge formation and regulates the motility of neutrophils during chemotaxis (99). DOCK2 has been linked to HIV through the ability of the HIV virulence factor Nef to activate Rac through the DOCK2/ELMO complex (100). Like DOCK1, DOCK2 requires ELMO for proper activation of Rac (101).

DOCK3, expressed exclusively in the brain, was originally identified as a protein able to interact with the presenilin protein in a yeast two-hybrid assay (70). Presenilin is a transmembrane protein that functions as part of a protease complex. Mutations in presenilin are associated with an increased risk for Alzheimer's disease. Consistent with its interactions with presenilin, DOCK3 was found to aggregate in regions of the brain affected by Alzhimer's disease (102). Similarly, DOCK3 expression was found to decrease the amount of β-amyloid precursor protein by increasing its degradation rate (102). DOCK3 knockout mice show behavioral defects relating to motor impairment.

The mice had multiple defects in their axons including impaired axonal transport and an altered cytoskeleton (103). DOCK3 also plays a role in cell-cell adhesion through its ability to stabilize N-cadherin and β-catenin (104-105). DOCK3 binding to β-catenin reduces its nuclear localization resulting in decreased transcription of β-catenin regulated genes (105). Like DOCKs 1-2, DOCK3 binds and activates Rac1 in an ELMO-dependent fashion (71).

DOCK4 was originally identified as a Rap1 GEF that was deleted in an osteosarcoma cancer cell line and mutated in both prostate and ovarian cancer cell lines (106). DOCK4 specificity was reported to switch from Rap1 with wild-type DOCK4 to Rac1 and Cdc42 with the P1718L mutation found in the cancer cell lines. Subsequent studies cast doubt on the ability of DOCK4 to activate Rap1, finding that it only activated Rac1 (94,107). A splice variant of DOCK4, which also activates Rac1, has been shown to be expressed in the brain, eye, and inner ear tissue (108). Like DOCK3, DOCK4 has been implicated in the Wnt/ β-catenin signaling pathway through the ability of DOCK4 to interact with the β-catenin degradation complex in a manner which enhances the stability of β-catenin (109).

Very little is known about DOCK5, which was the last DOCK family member to be cloned. It has been observed to be upregulated during osteoclast differentiation and knockdown of its expression in osteoclasts led to apoptosis (110). DOCK5 is highly expressed in the eye and a mutation in DOCK5 which markedly decreases protein stability is associated with the rupture of the eye lens in mice (111). In zebrafish, DOCK5 is required for fusion of myoblasts (112).

The biological functions of DOCK homologues have been studied in a number of model organisms such as *Drosophilia melanogaster* and *Caenorhabditis elegans*. DOCK1 is known as Myoblast city in *D. melanogaster* and mutations in Myoblast city lead to defects in cytoskeleton organization and dorsal closure of the epidermis during embryogenesis (113). These defects were noted to be similar to defects caused by mutations in the Rac1 homologue giving an early indication that DOCK1 and Rac might operate in the same signaling pathway. Mutations in the *C. elegans* homologue of DOCK1, Ced-5, cause defects in cellular migration as well as the loss of engulfment of cell corpses (114-115). Similar defects were found in mutations of the Rac homologue. DOCK family proteins have also been shown to activate Rho GTPases in *Dictyostelium discoideum* (116) and *Arabadopsis thaliana* (117).

The role of ELMO in regulating DOCKs 1-4 function

ELMO binds to DOCKs 1-4 and promotes migration of cells when co-expressed with DOCK1 (107,118). Based on sequence homology, DOCK5 is also expected to interact with ELMO, but no interaction is predicted with DOCKs 6-11. There are two main sites of ELMO interaction with DOCK1. ELMO has a PH domain in its C-terminus which contains a long α-helical extension that interacts with a helical region in the N-terminus of DOCK1 located after the SH3 domain (119). ELMO also contains polyproline motifs in its C-terminus that interact directly with the SH3 domain of DOCK1 (119-120). DOCK1 has been reported to dimerize (121) and likely forms a heterotetramer with ELMO in cells (93).

Binding to ELMO has been reported to relieve the autoinhibition of DOCK proteins (94). The PH domain of ELMO has been proposed to stabalize a nucleotide-free conformation of Rac1 in combination with DOCK1 leading to the suggestion that DOCK and ELMO may function as a bipartite GEF (120). However, other studies have seen no effect of ELMO on *in vivo* DOCK1 activity suggesting that ELMO plays no direct role in the nucleotide exchange reaction (74,119). ELMO also prevents the degradation of DOCK1 by inhibiting it from being ubiquitinated (122).

There are a number of proteins which can interact with ELMO to affect DOCK protein activity. RhoG has been identified as being able to interact with ELMO and the ELMO/DOCK1 complex to activate Rac1 (123). The pathogenic bacterium *Shigella* injects the IpgB1 protein into host cells where it can mimic the function of RhoG in activating the DOCK/ELMO complex to promote membrane ruffling necessary for invasion (124). In addition, ELMO has been shown to bind to the Src family kinase Hck, leading to phosphorylation of ELMO (125). Phosphorylation of Y511 has been implicated as critical for the ability of ELMO to promote both phagocytosis and migration in cells. Interestingly, a Y511F mutation does not decrease the ability of ELMO to interact with DOCK1 suggesting that ELMO phosphorylation may regulate its binding to other proteins (126). ELMO can interact with ERM family proteins, which help to cross-link the plasma membrane and actin cytoskeleton, but these interactions do not appear to alter the ability of the DOCK/ELMO complex to exchange on Rac1 (127). While there are clearly many proteins that can modulate DOCK1 activity through effects on ELMO, the mechanisms by which they do so remain to be elucidated.

The biological roles of DOCKs 6-11

DOCK6 has been characterized as a GEF capable of catalyzing nucleotide exchange on both Rac1 and Cdc42 (128). The isolated DHR-2 of DOCK6 was able to cause exchange on Rac1 and Cdc42 both *in vitro* and *in vivo* as well as to preferentially bind the nucleotide-depleted states of Rac1 and Cdc42. Transfection of the DOCK6 DHR-2 into COS-7 cells resulted in the formation of both lamellipodia and filopodia, characteristic of Rac1 and Cdc42 activation respectively. Also, DOCK6 has been shown to regulate neurite outgrowth in N1E-115 cells (128). Interestingly, full-length DOCK6 was observed to neither stimulate nucleotide exchange, nor alter the cytoskeleton upon transfection, suggesting that it may be autoinhibited in its basal state as observed for DOCK1. However, since DOCK6 lacks the SH3 domain implicated in DOCK1 autoinhibition, it would require an alternative mechanism of autoinhibition. Similarly, the DOCK6 mechanism of activation is most likely dramatically different then DOCKs 1-5 since DOCK6 has no interactions with either CRK or ELMO.

DOCK7 was originally identified as a Rac specific exchange factor critical for the induction of neuronal polarity and axon formation (129). DOCK7 simulates nucleotide exchange on Rac1 and Rac3, but not Cdc42 or RhoA. However, a subsequent study found that DOCK7 activates both Rac1 and Cdc42 leaving the exact GTPase specificity of DOCK7 in dispute (130). DOCK7 is expressed in multiple tissues, but is highly enriched in the brain (129). In unpolarized neurons, DOCK7 localizes to the neurite which will become the future axon. Knockdown of DOCK7 abolished the ability of neurons to form an axon while overexpression of DOCK7 gave a multi-axon phenotype (129). Phosphorylation of DOCK7 on Y1118 by the ErbB2 tyrosine kinase receptor

enhances its GEF activity and in turn promotes the migration of Schwann cells (130). Despite the characterization of DOCK7 as essential for neuronal development and maintenance, a DOCK7 knockout mouse showed no neurobehavioral defects (131). It is possible that other GEFs such as DOCK6 or DOCK8 may compensate for the loss of DOCK7.

DOCK8 was initially identified for its ability to interact with Cdc42 in a yeast two-hybrid assay (132). Transfection of DOCK8 resulted in the formation of vesicular structures containing filamentous actin implicating DOCK8 in regulating cytoskeletal structure. Disruption of DOCK8 function has been observed in multiple disease phenotypes. DOCK8 has been shown to be deleted or to have its expression down-regulated in almost 90% of lung cancers (133) and 45% of hepatocellular carcinomas (134). DOCK8 has also been found to be deleted in multiple patients with mental retardation (135) and with a novel heritable immunodeficiency disease (136). The basis for DOCK8 immunological defects results from the inability DOCK8 null B-cells to undergo affinity maturation and produce high affinity antibodies when mounting an immune response (137).

DOCK9 (also known as zizimin1) was identified from cell lysates as a protein that specifically bound nucleotide-depleted Cdc42 (72). DOCK9 was shown to act as a GEF for Cdc42 *in vitro* and to induce filopodia formation when over-expressed. DOCK9 mRNA expression was found in almost all cell types with the highest levels in the heart and lowest levels in hematopoetic cells (72). A separate study showed high DOCK9 mRNA expression in the placenta, lungs, kidney, and brain (138). DOCK9 regulates dendrite growth in neuronal cells as knockdown of DOCK9 inhibited the length of

dendrites while its overexpression increased the dendritic length (139). Furthermore, dendrites were not lengthened when a DOCK9 mutant incapable of activating Cdc42 was overexpressed, implicating DOCK9 activation of Cdc42 as important for neuronal physiology (139). Indeed, a genetic study has linked DOCK9 with an increased risk of bipolar disorder (140).

The ability of DOCK9 to homodimerize was first recognized when FLAG-tagged DOCK9 co-immunoprecipitated with HA-tagged DOCK9. The region of DOCK9 required for dimerization was mapped to amino acids 1696-1878, which reside in the DHR-2 (121). The structure of the DOCK9 DHR-2 bound to Cdc42 revealed the dimerization interface consisted of two α-helices located between amino acids 1716-1746 (79). The biological significance of DOCK9 dimerization is not fully understood. C-terminal DOCK9 fragments lacking the dimerization interface are shown to be monomeric, but retain full the ability to catalyze nucleotide exchange on Cdc42 (121).

Three separate regions in the N-terminus of DOCK9 have been shown to interact with the DOCK9 DHR-2 in co-immunoprecipitation experiments suggesting DOCK9 may be autoinhibited in the basal state (121). Constructs spanning amino acids 1-175, 288-539 as well as the DHR-1 (640-885) bound to a portion of the DHR-2 (1693-2069). The DOCK9 PH domain (172-282) did not interact with the DHR-2. In transfection experiments, the C-terminal half of DOCK9 was able to bind four times the amount of Cdc42 as the full-length DOCK9 construct. In addition, incubation of the N-terminal half of DOCK9 to the C-terminal half inhibited the binding of Cdc42 by 70% compared to the C-terminal half alone. Thus, it appears DOCK9 is autoinhibited in the full-length state as

observed for DOCK1. However, the molecular details of this autoinhibition remain ill-defined, as well as relief of DOCK9 autoinhibition.

The PH domain of DOCK9 has been implicated in membrane binding and truncation of the PH domain dramatically reduced the membrane association of DOCK9 in NIH-3T3 cells (121). Furthermore, expression of the PH domain alone resulted in significant membrane association in cells. Lipid dot blots showed both full-length DOCK9 and the isolated DOCK9 PH domain bound to a variety of lipids with highest affinity to mono-phosphorylated inositides and phosphatidylserine. Interestingly, full-length DOCK9 bound more efficiently than the isolated PH domain suggesting that other domains may also contribute to membrane association (121). A subsequent study found that both the PH domain and the DHR-1 were critical for the membrane association of DOCK9 (139).

DOCK10 (also known as zizimin3) activates Cdc42 as well as the closely related GTPases TCL (141). DOCK10 mRNA is highly expressed in leukocytes, but is also found at lower levels in hematopoetic and neuronal tissues (138). DOCK10 has been shown to regulate the ameboid-like motility of invading melanoma cells through activation of Cdc42 and its downstream effector Pak2 (142). The expression of DOCK10 in B-cells is upregulated by interleukin-4 (IL4) suggesting that DOCK10 mediates the IL4-dependent activation of RhoGTPases in lymphocytes (138).

DOCK11 (also known as zizimin2) is a Cdc42 specific GEF that both binds preferentially to nucleotide-depleted Cdc42 and causes activation of Cdc42 when transfected into COS-7 cells (141). DOCK11 mRNA is highly expressed in leukocytes and the placenta, although it is detected ubiquitously (138). Unexpectedly, full-length

DOCK11 was identified as binding to the constitutively active form of Cdc42 harboring the Q61L mutation, which is GTPase deficient and expected to bind effector proteins (143). In contrast, the isolated DHR-2 of DOCK11 selectively bound Cdc42 T17N, which is unable to bind nucleotides and expected to interact strongly with GEFs. The binding site for activated Cdc42 was mapped to amino acids 66-126 in the N-terminus of DOCK11 and the binding of activated Cdc42 was shown to enhance the GEF activity of DOCK11 (143). This suggests a positive feedback model where initial activation of Cdc42 increases the GEF activity of DOCK11to lead to greater activation of Cdc42.

DOCK family GEF activity and GTPase specificity

Due to the large size of the DOCK proteins, a number of studies utilized truncation studies of full-length protein to identify the regions necessary for GEF activity. The first region identified was the Docker domain which spans amino acids 1111-1657 of DOCK1 (68). This region was sufficient to bind nucleotide-free Rac upon over-expression in 293T cells and to cause loading of nucleotide onto Rac. In addition, the same study identified a triple mutation within DOCK1, I1487A + S1488A + P1489A, which abolished the ability of DOCK1 to exchange on Rac1 (68). The corresponding mutation in DOCK9 also abolishes its ability to activate Cdc42 (139). However, these residues showed no contact to Cdc42 in the DOCK9/Cdc42 structure suggesting they may affect the proper folding of DOCK9 (79).

The identification of DOCK9 and subsequent sequence alignments with DOCK1 led to the identification of two regions of high sequence homology, initially termed CZH1 and CZH2. The CZH1 region was shown to be dispensable for nucleotide

exchange on Cdc42, whereas the CZH2 region (residues 1693 -2069) was shown to be sufficient to cause nucleotide exchange on Cdc42 (72). It was also found that slightly longer constructs, such as 1512-2069 seemed to bind more efficiently to Cdc42. Later, a region similar to CZH2 and Docker was identified by Cote and Vuori and was termed DHR-2, encompassing residues 1111-1636 in DOCK1 and 1505-2069 in DOCK9 (73). The DHR-2 from DOCK1 and DOCK9 were sufficient for binding nucleotide-depleted Rac1 and Cdc42 respectively and were able to cause nucleotide exchange in *in vitro* nucleotide exchange assays.

The mechanism of nucleotide exchange catalyzed by DOCK proteins has been elucidated by the recent structure of Cdc42 bound to the DHR-2 of DOCK9 (79). DOCK9 was observed to fold into three distinct lobes; A, B, and C. Lobe A mediated dimerization, but did not directly contact Cdc42, while lobes B and C stabilized a nucleotide-free conformation of Cdc42. The insertion of F28 in switch 1 of Cdc42 into a hydrophobic groove in lobe B of DOCK9 stabilizes conformational changes in switch 1 that disrupts interactions with the guanine base of the bound nucleotide. A major contribution to nucleotide exchange comes from the insertion of V1951 of DOCK9 into a position that would directly occlude binding of the Mg^{2+} that coordinates nucleotide binding. Movement of C18 also contributes to nucleotide exchange by breaking an interaction with the α-phosphate of the bound nucleotide (79). Other DOCK family proteins are likely to catalyze nucleotide exchange in a similar manner as key residues such as V1951 are conserved throughout the DOCK family. However, the regulation of DOCK family protein and the contributions of accessory proteins such as ELMO to nucleotide exchange remain unanswered questions.

The structural determinants for specificity between Rho GTPases and DOCK proteins are only beginning to be examined. One study aiming to identify the GTPase residues required for interaction with DOCK2 and DOCK9 found that a chimeric GTPase containing residues 1-58 of Cdc42 fused to 59-192 of Rac2 was robustly activated by DOCK9, but not DOCK2/ELMO. In contrast, the reciprocal chimera (1-58 of Rac2 fused to 59-192 of Cdc42) was unable to be activated by DOCK9, but was activated by DOCK2/ELMO implicating that key determinants for specificity were located within residues 1-58 of the GTPase (144). Further mutational analysis indicated that GTPase residues 27, 30, 51, 52, and 56 were essential in mediating the specificity of the GTPase/DOCK interaction. The structure of DOCK9 bound to Cdc42 showed direct interactions between DOCK9 and residues 27, 30, and 56 of Cdc42. Based on contacts with Cdc42 specificity determinants, it would be predicted that DOCK9 residues 1812, 1814, 1832, 1941, and 1944 would be important in GTPase selectivity, but this has not been experimentally verified.

Part 3: SmgGDS is a unique and atypical GEF

Discovery of SmgGDS as a GEF with unusually broad specificity

SmgGDS is the only identified member of a class of atypical guanine nucleotide exchange factors composed entirely of armadillo repeats (145). There are two known SmgGDS splice variants, SmgGDS-608 (henceforth referred to as SmgGDS) and SmgGDS-558. SmgGDS has one additional ARM repeat inserted in the middle of the protein compared to SmgGDS-558. SmgGDS has a unique tertiary structure from all known exchange factors suggesting the mechanism by which it catalyzes nucleotide

exchange may be distinct. SmgGDS has been reported to catalyze exchange on a diverse array of small GTPases including the Ras family members Rap1A (146), Rap1B (147), and K-Ras (146), as well as the Rho family members Rac1(148), Rac2 (149), Cdc42 (150), RhoA (151) , and RhoB (151). Consistent with the unique structural characteristics of SmgGDS, it is the only GEF reported to activate GTPases from different families making SmgGDS an attractive target. While SmgGDS was the first mammalian GEF identified for the Rho GTPase family, there remains very little clarity to the mechanism, specificity, or biological role of SmgGDS (151).

The SmgGDS protein was first discovered by the Takai group in 1990 during a search for proteins with the capacity to promote the dissociation of GDP from small GTPases. It was originally partially purified from bovine brain extracts in a fraction able to stimulate release of GDP from RhoA and RhoB, but not H-Ras, Rap1A, or Rab3 (151). However, the sample was contaminated by G proteins and could not be fully analyzed. Subsequent studies with further purified protein samples led to the conclusion that SmgGDS could activate Rap1A and Rap1B, but not H-Ras, Rab3, or RhoB (147). These studies establish conflicting reports of GTPase specificity, and mark a trend of inconsistent results reported on SmgGDS.

SmgGDS was first observed to stimulate nucleotide exchange on K-Ras during a study in which Rap1A, Rap1B, RhoB and RhoA were also activated (146). Using GTPases purified from mammalian cells, SmgGDS was shown to activate Rac1 (148). These results were significant as SmgGDS was now proposed to activate multiple branches of the Rho GTPase family (RhoA, RhoB, and Rac1) as well as multiple branches of the Ras GTPase family (Rap1A, Rap1B, and K-Ras). However, SmgGDS

was reported to be inactive on a number of small GTPases from the Rab GTPase family (148).

SmgGDS was cloned in 1991 and the properties of bacterially expressed protein were shown to be nearly identical to those of SmgGDS purified from brain extracts with regards to protein size, limited proteolysis fragments, and exchange activity upon Rap1B (152). An enzymatic study comparing bacterially expressed SmgGDS reported a K_{cat} of 0.37 nmol/min/nmol and K_m of 220 nM for catalyzing the dissociation of [^{3}H]GDP from prenylated K-Ras with no activity being observed on nonprenylated protein (153). Work by Yaku *et al.* first identified the ability of SmgGDS to exchange on Cdc42 and illustrated that SmgGDS could most efficiently exchange on RhoA, followed by Cdc42, then K-Ras, Rap1B, and least efficiently on Rac1 (150). A similar study found SmgGDS caused a greater increase in nucleotide exchange on RhoA than Rac1 (154). These data provide evidence that RhoA is the preferred substrate for SmgGDS.

An elegant study by Hutchinson and Eccleston characterized the ability of SmgGDS to catalyze nucleotide exchange on RhoA using fluorescently labeled MANT nucleotides (155). Their data suggested that SmgGDS operated by an associative displacement mechanism that required a conformational change in the SmgGDS/RhoA/nucleotide intermediate prior to nucleotide release. They saw a 5,000-fold enhancement of nucleotide release from RhoA when adding SmgGDS in single turnover conditions, but observed that the steady-state rate of nucleotide exchange would be markedly lower due to very slow dissociation of SmgGDS from the ternary SmgGDS/RhoA/nucleotide complex (155). No published work has characterized the

ability of SmgGDS to exchange on other GTPases using this methodology therefore the activity of SmgGDS on RhoA cannot be quantitatively compared to other GTPases.

The first and only study to examine the GTPase specificity of SmgGDS GEF activity outside of the Takai group was published from the Bokoch laboratory in 1994. In their hands, SmgGDS was found to be highly active on both RhoA and Rac2 and less active on Rap1A, Cdc42, and Rac1 (149). SmgGDS activity was observed on both prenylated and nonprenylated GTPases in contrast to earlier claims that prenylation was required for SmgGDS catalyzed exchange. While other studies have examined the ability of SmgGDS to bind various GTPases, as well as biological functions of SmgGDS, no additional studies have addressed the specificity of SmgGDS GEF activity.

In addition to examining nucleotide exchange activity, a number of early studies investigated the ability of SmgGDS to bind to GTPases using sucrose gradient ultracentrifugation. The first such study found that SmgGDS was able to form a 1:1 molar complex with Rap1B in both the GDP-bound and the GTPγS-bound state (156). A subsequent study found that SmgGDS could form complexes with prenylated K-Ras, Rap1B, and RhoA although the nucleotide state of the complexes was not determined (146). In the same study, no complexes were detected in the presence of non-prenylated proteins. Work by Nakanishi *et al.* demonstrated that SmgGDS would form a stable ternary complex with both [^{35}S]GTPγS-RhoA and [^{3}H]GDP-RhoA (157). The observation that SmgGDS would form a ternary complex with nucleotide-bound GTPases is surprising given that GEFs work through stabilization of the nucleotide-free GTPase conformation. An independent group also found that SmgGDS could form a complex

with GTPγS-bound, GDP-bound, and nucleotide-depleted Rac1 using an *in vitro* GST pull-down assay with purified proteins (149).

Despite the fact that SmgGDS has never been reported to exchange on H-Ras or N-Ras, one intriguing report indicates that SmgGDS can bind both nucleotide-free H-Ras and N-Ras (158). The binding was found to require the C-terminal 100 residues of SmgGDS. Interestingly, the C-terminus was dispensable for the interaction of SmgGDS with K-Ras, RhoA, Rac1, and Rap1A, suggesting separate binding sites for H-Ras and N-Ras compared to the other GTPases that are activated by SmgGDS. Indeed, their data suggested that addition of increasing amounts of dominant negative N-Ras to equimolar concentrations of RhoA and SmgGDS allowed for binding of N-Ras without disrupting RhoA binding. Since no other reports have identified H-Ras or N-Ras as able to bind to or modulate SmgGDS activity, the role of nonsubstrate GTPases in SmgGDS function remains unknown.

In addition to its ability to stimulate nucleotide exchange on GTPases, early studies on SmgGDS indicate that SmgGDS is able to extract small GTPases from membranes. The first study to examine this possibility established that purified SmgGDS could inhibit both the binding of Rap1B to synaptic plasma membranes and stimulate the release of prebound Rap1B in a dose-dependent fashion (156). These results suggested that SmgGDS could bind to prenylated proteins and remove them from the plasma membrane as observed for the GDI proteins.

A follow-up study found that SmgGDS could extract K-Ras from membranes as well as inhibit the binding of K-Ras to membranes (159). However, in addition to examining the activity of purified proteins this study also looked at the relative

abundance of K-Ras in membrane and soluble fractions of COS-7 cells transfected with SmgGDS. Consistent with the results using purified proteins, co-transfection of SmgGDS increased the amount of soluble K-Ras from 25% to 50%. In addition, the ability of SmgGDS to activate K-Ras *in vivo* was examined by labeling co-transfected cells with radioactive phosphate prior to lysis. After cells were lysed, the radioactive K-Ras was immunoprecipitated and the GTP and GDP-bound forms were separated by thin-layer chromatography. The authors were able to see a small 6% increase in the amount of GTP-bound K-Ras when SmgGDS was cotransfected with K-Ras. This is likely an underestimate of the true activation as some GTP was likely hydrolyzed during the assay. This was the first study to show *in vivo* activation of a GTPase by SmgGDS.

SmgGDS homologues have been studied from two species. The *Xenopus laevis* homologue of SmgGDS was discovered in a yeast two-hybrid assay through its interaction with the RalB homologue (160). Interaction was also detected with both GDP and GTP bound forms of RalA and required prenylation of the GTPase. This was the first report of the presence of an inserted ARM repeat in SmgGDS sequences. The *Dictyostelium discoideum* homologue of SmgGDS, Darlin, was found to bind both constitutively active and dominant negative mutations of a Rac homologue (161). Darlin is the most divergent SmgGDS homologue with a large C-terminal extension that is not conserved in SmgGDS homologues from other species.

The GTPase polybasic region and isoprenoid modification and SmgGDS function

In addition to the nebulous GTPase specificity of SmgGDS, there are numerous conflicting reports regarding the necessity of a polybasic region or prenyl group on the C-

terminus of a GTPase for proper exchange by SmgGDS. The first indication that the polybasic region or prenyl group of GTPases might be necessary for SmgGDS catalyzed nucleotide exchange came from testing an N-terminal fragment of Rap1B produced by limited proteolysis that retained nucleotide binding and GTPase activity, but lacked the ability to bind to membranes. SmgGDS was unable to catalyze nucleotide exchange on the N-terminal fragment, but was active on full length Rap1B (162). These data suggested that the polybasic region, the prenyl group, or both might be necessary for the ability of SmgGDS to exchange on small GTPases. A subsequent study showed that SmgGDS could exchange on RhoA purified from insect cell membranes, but not produced in bacteria, indicating that the prenyl moiety was necessary for SmgGDS catalyzed exchange (163).

The work of Mizuno *et al.* compared the ability of SmgGDS to exchange on the prenylated and nonprenylated forms of K-Ras, RhoA, RhoB, and Rap1B purified from insect cells. SmgGDS only catalyzed exchange on prenylated proteins purified from the membrane fraction of insect cells (146). This work strengthened the claim that lipid modification of the GTPase was necessary for SmgGDS catalyzed exchange. However, RhoB which lacks a polybasic region in its C-terminus remained activated by SmgGDS, arguing against the necessity of polybasic tail for SmgGDS catalyzed nucleotide exchange.

A library of peptides derived from the C-terminus of Rap1B were tested for their ability to bind and inhibit SmgGDS (164). Among these, only a geranylgeranylated polybasic peptide could directly bind to SmgGDS as measured by sucrose density ultracentrifugation, but not the non-lipid modified analogous peptide. It was also shown

that the geranylgeranylated polybasic peptide could competitively inhibit the ability of SmgGDS to exchange on Rap1B with an estimated K_i of 400 nM. To test lipid specificity, a number of different lipid moieties were attached to the peptide, and although geranylgeranyl was the most effective in inhibiting SmgGDS, farnesyl and palmitoyl groups also caused inhibition. Interestingly, reversing charges on the four basic residues of the geranylgeranylated peptide resulted in a geranylgeranylated polyacidic peptide that also inhibited SmgGDS catalyzed exchange on Rap1B. However, the affinity of this polyacidic peptide for SmgGDS was decreased approximately 10-fold. These data strongly suggest that the polybasic region could enhance interaction with SmgGDS, but was not necessary for SmgGDS catalyzed exchange.

In contrast to the initial observations of the Takai group, a number of subsequent studies convincingly showed that SmgGDS could activate nonprenylated proteins. SmgGDS was first shown to stimulate nucleotide exchange on nonprenylated Rac1, Cdc42, and RhoA (149). A separate group was able to characterize SmgGDS exchange on a nonprenylated RhoA produced in bacteria (165). The two splice variants of SmgGDS are reported to differ in their preference for binding to prenylated and nonprenylated GTPases, with the shorter SmgGDS-558 isoform only binding prenylated GTPases *in vivo* and the longer isoform preferring to interact with nonprenylated GTPases (T. Berg and C. Williams, personal communication). However, all of the conflicting studies regarding the necessity of GTPase prenylation for SmgGDS function used the shorter SmgGDS-558 isoform. Thus, the discrepancy in the results cannot be explained.

The importance of the GTPase polybasic region in SmgGDS activity stems from the fact that, with the exception of conflicting reports on RhoB, all proposed SmgGDS substrates have a strong polybasic region in their C-terminus. Cross-linking experiments suggested the polybasic region of Rap1B could directly interact with SmgGDS (166). Mutating the polybasic region of small GTPases causes a decrease in the ability of dominant negative Rac1and RhoA to associate with SmgGDS in cells (167). The acidic membrane phospholipids phosphatidylinositol, phosphatidylinosoitol-4,5-bisphosphate, phosphatidic acid, and phosphatidylserine all inhibited both the ability of SmgGDS to exchange on Rap1B and the ability of SmgGDS to extract Rap1B from membranes (168). The acidic phospholipids were suggested to compete with SmgGDS for interactions in the polybasic region of Rap1B. However, no study has directly tested the importance of an intact polybasic region on small GTPases using nucleotide exchange assays.

Phosphorylation of Rap1B by protein kinsase A on S179 in the polybasic region enhances the ability of SmgGDS to stimulate nucletotide exchange (169). This result is somewhat counterintuitive as the addition of a phosphate group to the serine would reduce the basic nature of the C-terminus. However another study by the same group showed no difference in the ability of phosphorylated and non-phosphorylated Rap1B derived C-terminal peptides to inhibit the ability of SmgGDS to exchange on Rap1B (164). Since no other studies have followed up on these initial conflicting observations, the question still remains, if phosphorylation of small GTPases in the polybasic region affects their interactions with or exchange by SmgGDS.

SmgGDS in biology and in cancer

One of the earliest studies concerning the biological function of SmgGDS looked at effects on DNA synthesis from microinjecting SmgGDS protein into Swiss 3T3 cells (170). While, injection of SmgGDS alone had no effect, co-injection of SmgGDS and K-Ras significantly increased the level of DNA synthesis to a level equivalent to that obtained with the injection of the activated K-Ras V12 mutant. It was also observed that injection of Rap1B-GDP caused no increase in DNA synthesis, while both Rap1B-GTPγS and Rap1B-GDP coinjected with SmgGDS significantly increased DNA synthesis. These results suggest that SmgGDS activates Rap1B and K-Ras in cells to promote DNA synthesis.

SmgGDS has been reported to have interactions with only two proteins that are not GTPases. SmgGDS was found to interact with SMAP in a yeast-two hybrid assay using a human brain cDNA library (171). Although direct interaction was shown using purified protein, SMAP had no effect on SmgGDS GEF activity. SMAP is an ARM-repeat protein that is highly homologous to SpKAP115, an accessory protein for sea urchin kinesin II molecular motor implicated in vesicular transport along microtubules. The mouse homologue of SMAP, KAP3, forms a complex with KIF3A and KIF3B molecular motors which are implicated in protein transport out of the Golgi apparatus (172) as well as for roles in mitosis (173) and endosome movement (174). Other ARM repeat containing proteins such as APC and β-catenin are also thought to be transported by the KAP3/KIF3A/KIF3B complex (175). In analogous fashion, SMAP may be involved in SmgGDS transport, but has not been implicated in affecting its GEF activity.

SmgGDS has also been found to interact with the Dbl family GEF β-Pix in co-immunoprecipitation studies (176). Knockdown of SmgGDS abolished the ability of β-

Pix to activate Rac, but mutations in the DH domain of β-Pix did not affect the ability of the SmgGDS/ β-Pix complex to activate Rac. These results suggest that SmgGDS is the catalytic component of the SmgGDS/ β-Pix complex, although it does not rule out a direct role for β-Pix. The binding of SmgGDS to β-Pix was mapped to the GIT-binding domain of β-Pix. Activation of Rac by the SmgGDS/β-Pix complex led to neurite outgrowth in PC12 cells (176). However, it has not been examined if β-Pix directly influences the GEF activity of SmgGDS or if β-Pix influences other SmgGDS functions such as subcellular localization.

SmgGDS has been proposed to play a role in the nucleocytoplasmic shuttling of Rho family GTPases. Indeed, RhoA is normally found in both the cytoplasm and the nucleus. SmgGDS expression modifies the localization of RhoA so that it is entirely cytoplasmic (177). Consistent with the ability of SmgGDS to increase the amount of cytoplasmic RhoA, a nuclear export sequence has been found in N-terminal residues 4-13 of SmgGDS (167). SmgGDS was also found to have increased concentration in the nucleus when co-expressed with Rac1 proteins, most notably with a constitutively active Rac1 mutant. The interpretation of these data are unclear because the same study showed that SmgGDS was only able to co-immunoprecipitate with dominant negative Rac1 proteins (167). However, a model was proposed in which SmgGDS interactions with Rac1 led to a GTP-dependent translocation of SmgGDS to the nucleus whereas interactions with RhoA inhibited this translocation by retaining SmgGDS in the cytoplasm (167,178).

Although the GTPase specificity of SmgGDS remains ill-defined, SmgGDS has been reported to be most active on RhoA in a number of studies (149-150,154). The

strongest cellular evidence linking SmgGDS with RhoA signaling comes from an examination of the role of SmgGDS in vascular smooth muscle cells (179). RhoA is known to control contraction in vascular smooth muscle cells. Knock-down of SmgGDS led to a loss of contraction equivalent to the one seen from knock-down of RhoA expression. Similarly, knock-down of both SmgGDS and RhoA caused disruption of myosin organization. More importantly, knockdown of SmgGDS inhibited the ability of serum to stimulate an increase in the levels of active RhoA suggesting that SmgGDS is necessary for the full activation of RhoA (179).

SmgGDS may also regulate cellular growth and survival. SmgGDS knockout mice show enhanced apoptosis of cardiomyocytes, thymocytes, and neuronal cells and 70% of the mice die of heart failure within 5 days of birth (180). Expression of SmgGDS in knockout thymocytes reduces sensitivity to apoptotic signals by a factor of four. The functional pathways linking SmgGDS to growth and survival remain to be elucidated.

The first insights into the role of SmgGDS in cancer came from a study by Fujioka *et al.* in 1992, in which SmgGDS was over-expressed in NIH-3T3 cells (181). Transfection of SmgGDS or K-Ras alone had no significant effect, but co-transfection of both SmgGDS and K-Ras led to significant increases in transformation as measured by foci formation and anchorage independent growth measured in soft agar assays. The effects were not as robust as transfection of an activated K-Ras V12 mutant, but they provided evidence that SmgGDS could activate K-Ras *in vivo* to lead to transformation.

SmgGDS protein levels are known to be upregulated in both non-small lung cancer carcinomas (182) as well as prostate carcinomas (183). In both cancer lines,

knockdown of SmgGDS inhibits proliferation of cancer cell lines in addition to causing reduced migration of cancer cells. SmgGDS is also found to be fused to the NUP98 nucleoporin protein in acute T-cell lymphocytic leukemia (184). This is intriguing considering that nucleoporins are involved in mediating the nucleocytoplasmic shuttling of proteins and SmgGDS was previously reported to control nuclear localization of Rho family GTPases. A more thorough understanding of SmgGDS may elucidate how its aberrant function can influence disease states such as cancer.

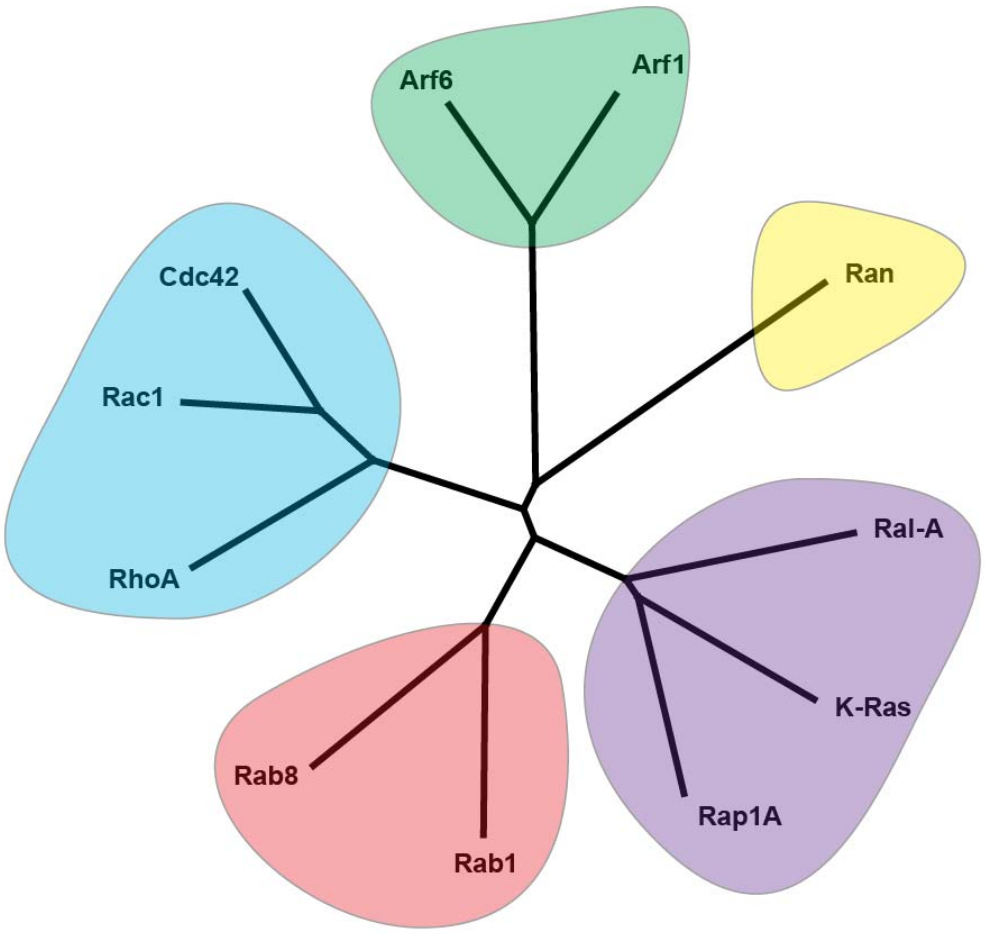

Figure 1: The Ras superfamily of small GTPases.
A cladogram of select Ras superfamily members illustrating the existence of five individual families: Arf (green), Ran (yellow), Ras (purple), Rab (red), and Rho (blue).

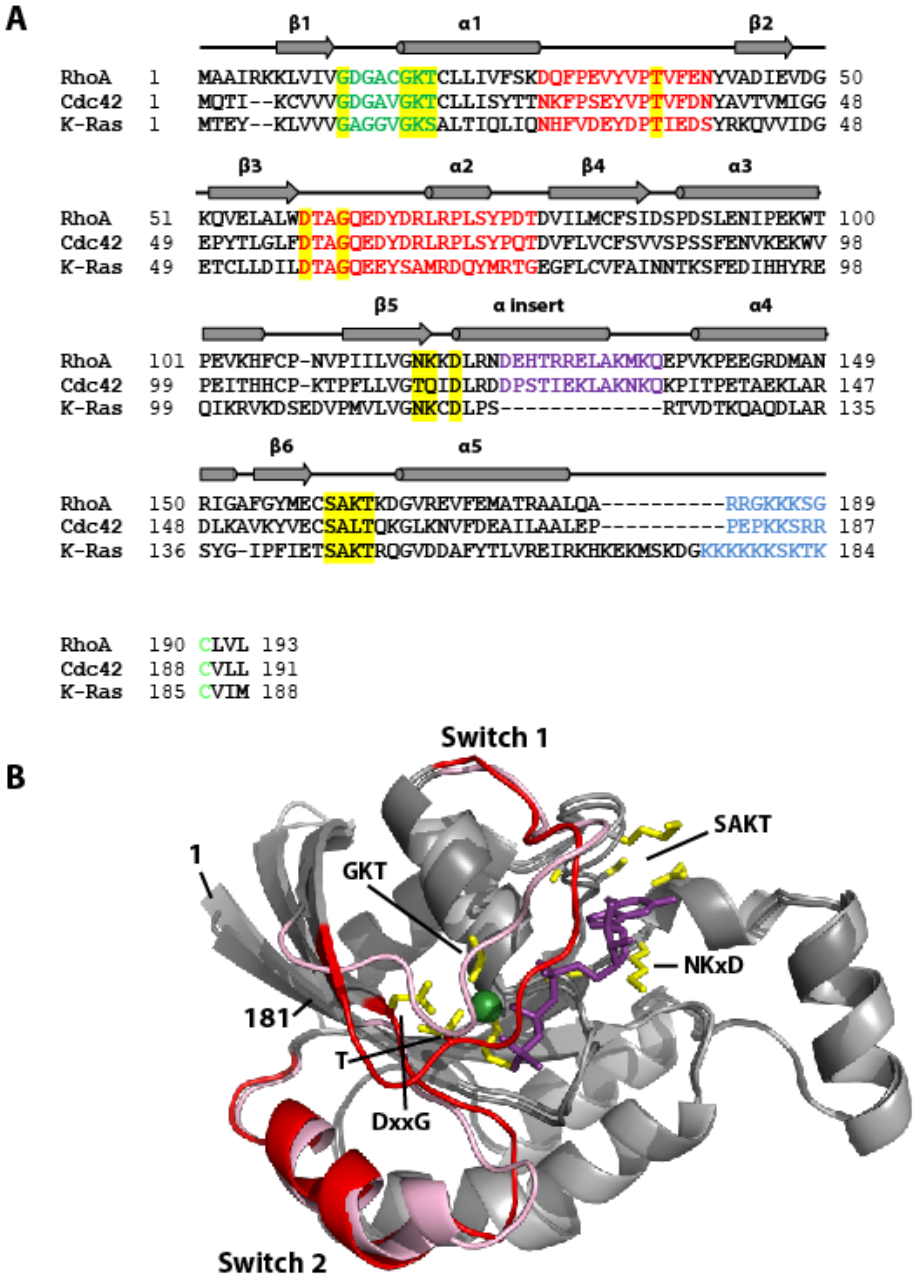

Figure 2: The ability of a GTPase to bind nucleotides and alter switch region conformation derives from its structure.

A, Alignment of RhoA, Cdc42, and K-Ras with P-loop (dark green) switch regions (red), Rho insert region (purple), residues that contact guanine nucleotide (highlighted in yellow), polybasic motif (blue), and prenylation site (light green). *B*, Superimposition of GTP- and GDP-bound RhoA with Mg^{2+} (green sphere), GTP (purple sticks), nucleotide binding residues from GTP-bound RhoA (yellow sticks), and switch region conformation from GTP-bound RhoA (red) and GDP-bound RhoA (pink). Not shown: GDP, Mg^{2+}, and nucleotide binding residues from GDP-bound RhoA.

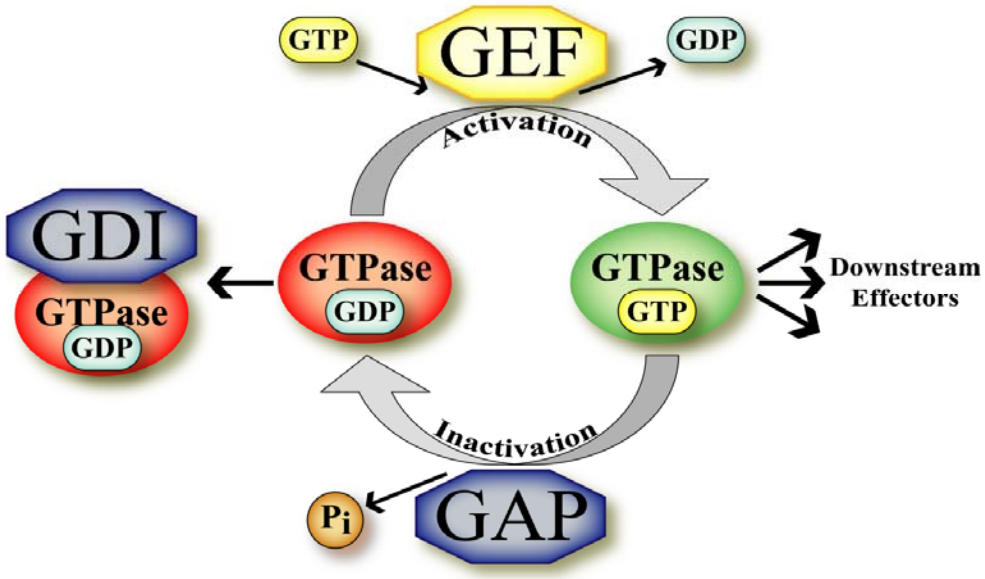

Figure 3: GTPases alter their nucleotide state through a dynamic cycle with the help of regulatory proteins.

A GTPase is only able to signal to downstream effectors when in the "on" state (green). The intrinsic hydrolysis of phosphate is greatly increased by GAP proteins resulting in the inactive GDP-bound conformation (red). This conformation can be stabilized by the GDI proteins which extract GTPases from the cell membrane and sequester them in an inactive state. GTPases are activated through the actions of GEFs which stabilize a nucleotide-free conformation of the GTPase that is driven to the active conformation by the high intracellular concentration of GTP.

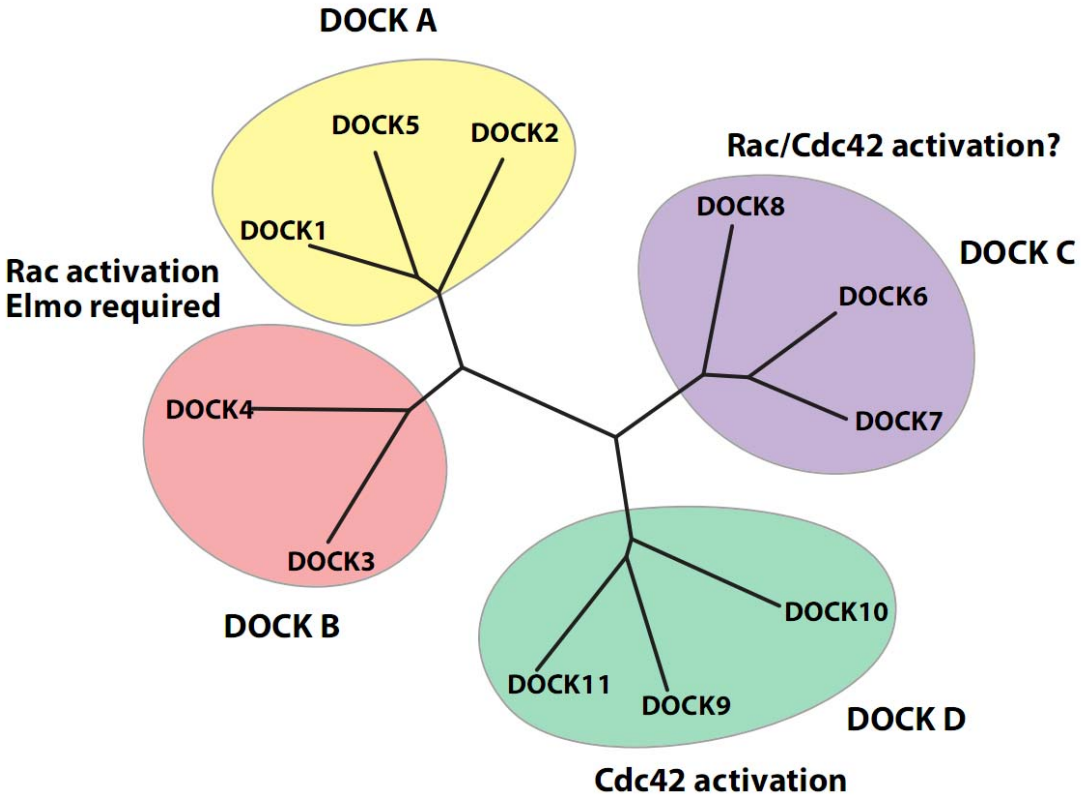

Figure 4: The DOCK family of Rho GEFs.
The DOCK family of Rho GEFs is subdivided into four subfamilies: DOCK A consisting of DOCK1, DOCK2, and DOCK5 (yellow), DOCK B consisting of DOCK3 and DOCK4 (red), DOCK C consisting of DOCK6, DOCK7, and DOCK8 (purple), and DOCK D consisting of DOCK9, DOCK10, and DOCK11 (green). DOCKs 1-5 activate Rac in an ELMO-dependent manner. The GTPase specificity of DOCKs 6-8 are not fully elucidated, but members have been suggested to activate both Rac and Cdc42. DOCKs 9-11 specifically activate Cdc42.

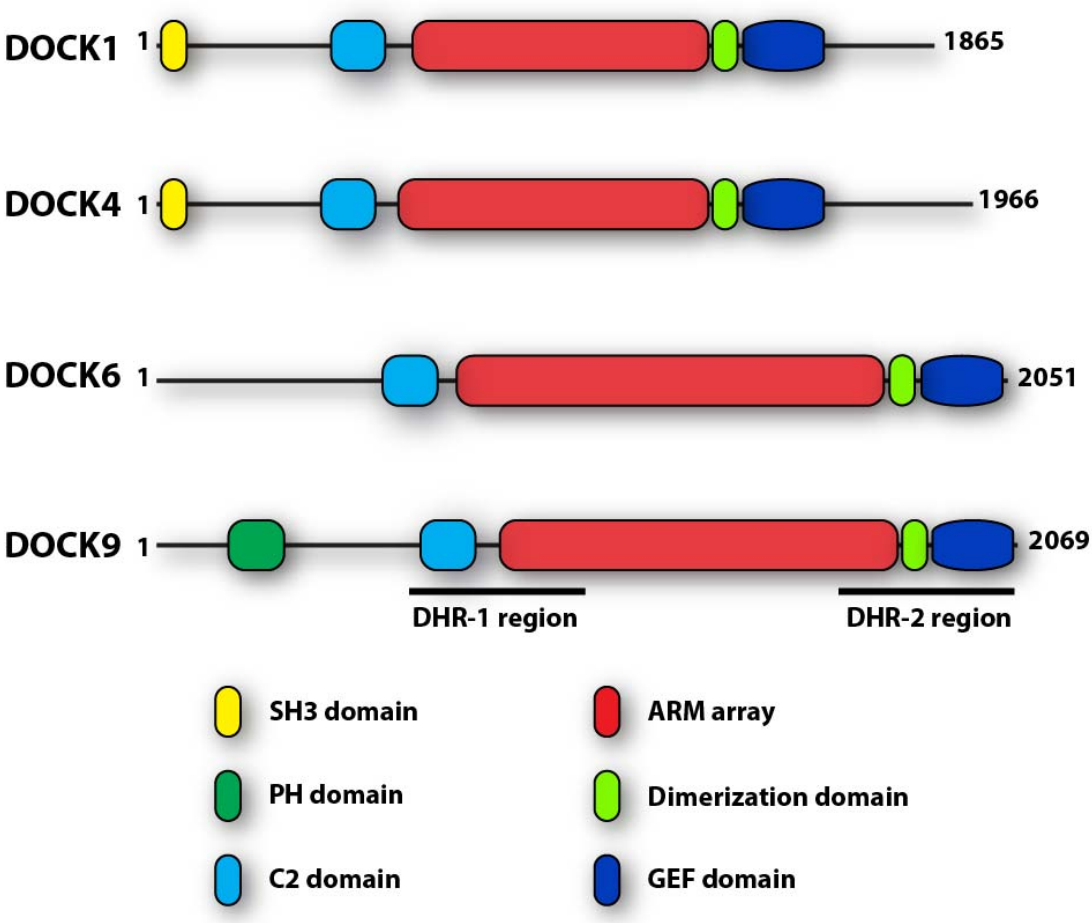

Figure 5: DOCK family members share a conserved core domain structure.
The DOCK family of GEFs share a number of conserved domains including a central C2 domain (cyan) followed by ARM repeats (red), a dimerization interface (light green), and a GEF domain (dark blue). DOCKs 1-5 have an SH3 domain (yellow) in their N-terminus while DOCKs 9-11 contain a PH domain (dark green). DOCK family members were identified based on sequence homology in two regions, DHR-1 and DHR-2, whose locations are shown underneath the domain structure of DOCK9.

CHAPTER 2: PURIFICATION AND ANALYSIS OF DOCK9 CONSTRUCTS ABLE TO CATALYZE NUCLEOTIDE EXCHANGE ON CDC42

Background

The DOCK (dedicator of cytokinesis) proteins are a family of Rho GEFs structurally and mechanistically distinct from the more numerous Dbl family Rho GEFs. There are 11 DOCK members in humans (**Fig. 4**). DOCKs 1-5 activate Rac, but require binding to an accessory protein, ELMO, for their GEF activity. DOCKs 9-11 activate Cdc42 in an ELMO-independent fashion, while DOCKs 6-8 may activate both Rac and Cdc42. DOCK proteins are evolutionary conserved in many species including *D. melanogaster* (113), *C. elegans* (114-115), and *A. thaliana* (117).

The DOCK family was originally identified to contain two regions of sequence conservation, DOCK homology region 1 and 2 (DHR-1 and DHR-2) (72-73). DHR-1 is implicated in membrane association (74-75), while DHR-2 is essential for GEF activity (76-77). Recent structures have revealed that the DHR-1 contains a C2 lipid binding domain which can specifically interact with phosphatidylinositol-(3,4,5)-triphosphate (78) while the DHR-2 domain contains both a dimerization domain and the catalytic GEF domain (79). The C2 and dimerization domains are linked by an ARM array such that all DOCK family members contain a conserved central domain architecture consisting of a C2 domain followed by an ARM array, a dimerization domain, and a GEF

domain (**Fig. 5**). In addition, DOCKs 1-5 have an SH3 domain in their N-terminus and a polyproline rich C-terminus, while the DOCKs 9-11 contain a PH domain in the N-terminus. No additional domains have been identified for DOCKs 6-8. The PH (121) and C2 (74-75,78) domains are implicated in membrane targeting, while the SH3 domain of DOCKs 1-5 is involved in autoinhibition of GEF activity (94) and binding to ELMO (119-120). Although they do not contain an SH3 domain, DOCK6 (128) and DOCK9 (121) are also believed to be autoinhibited in a basal state.

The founding family member, DOCK1, was originally identified as a 180 kDa protein that bound to the CRK adaptor protein (63) and led to activation of Rac1 (64). DOCK1 regulates cell motility and phagocytosis and is ubiquitously expressed (67). Other DOCK proteins show more limited tissue specific expression. DOCKs with neuronal functions include: DOCK3, implicated in axonal transport (103); DOCK6, a regulator of neurite outgrowth (128); and DOCK7, necessary for axon determination (128). DOCKs shown to have a role in the immune system include: DOCK2, required for proper lymphocyte chemotaxis (97); DOCK8, essential for B-cells to undergo affinity maturation (137); and DOCK10, implicated in interleukin-4 dependent activation of Rho GTPases in lymphocytes (138). DOCK proteins have also been shown to regulate cell-cell adhesion (104-105), gene expression (105), protein degradation (109), cell growth and differentiation (110), and myoblast fusion (112).

DOCK proteins have been implicated to play roles in a number of disease states. DOCK2 knockout mice show markedly reduced allograft rejection with mismatched tissues, implicating DOCK2 as necessary for normal immune system function (98). DOCK2 has also been linked to HIV through the ability of the HIV-virulence factor Nef

to activate Rac through the DOCK2/ELMO complex (100). DOCK3 is likely involved in the progression of Alzheimer's disease. DOCK3 interacts with the presenilin protein in a yeast two-hybrid assay (70), is found to aggregate in regions of the brain affected by Alzheimer's disease (102), and regulates the amount of β-amyloid precursor protein (102). DOCK4 is mutated by a number of different missense mutations in both prostate and ovarian cancer cell lines (106). DOCK8 has been shown deleted or to have its expression down-regulated in almost 90% of lung cancers (133) and 45% of hepatocellular carcinomas (134). DOCK8 has also been found to be deleted in multiple patients with mental retardation (135) and with a novel heritable immunodeficiency disease (136). DOCK10 has been shown to regulate the motility of invading melanoma cells (142). DOCK family GEFs are implicated in an array of disease states and are critical for their progression. Therefore, a thorough understanding of their normal biological activities is imperative.

In this study, we used DOCK9 as a model DOCK family member. DOCK9 (also known as zizimin1) is a GEF for Cdc42 *in vitro* and induces formation of filopodia (72). DOCK9 is ubiquitously expressed, including neuronal tissues where it regulates dendrite growth (139). DOCK9 was the first DOCK protein shown to dimerize; however, the biological significance of dimerization remains unknown as C-terminal constructs lacking the dimerization domain retain the full ability to catalyze nucleotide exchange on Cdc42 (121). DOCK9 contains a PH domain and the C2 domain that are critical for membrane association, but how these domains orient the catalytic domain at the membrane surface has not been determined (139). DOCK9 has been shown to be

autoinhibited with three separate N-terminal regions able to bind the DHR-2, yet the structural mechanism of autoinhibition and its relief is an open question (121).

The mechanism of nucleotide exchange catalyzed by DOCK proteins has been elucidated by the structure of Cdc42 bound to the DHR-2 of DOCK9 (79). DOCK9 was observed to fold into three distinct lobes; A, B, and C. Lobe A mediated dimerization, but did not directly contact Cdc42, while lobes B and C stabilized a nucleotide-free conformation of Cdc42. The insertion of F28 in switch 1 of Cdc42 into a hydrophobic groove in lobe B of DOCK9 stabilizes conformational changes in switch 1 of Cdc42 that disrupts interactions with the guanine base of the bound nucleotide. A major contribution to nucleotide exchange comes from the insertion of V1951 of DOCK9 into a position that directly occludes binding of the Mg^{2+} that coordinates nucleotide binding. Movement of C18 of Cdc42 also contributes to nucleotide exchange by breaking an interaction with the α-phosphate of the bound nucleotide (79). Other DOCK family proteins are likely to catalyze nucleotide exchange in a similar manner as key residues such as V1951 are conserved throughout the entire DOCK family.

The major aim of this study was to determine the mechanism of nucleotide exchange catalyzed by DOCK9 on Cdc42, unknown at the outset of these experiments. The recent structure of the DHR-2 of DOCK9 bound to Cdc42 has answered many of the questions this research was attempting to address. However, multiple questions remain to be addressed. Only the catalytic domain and dimerization domains of DOCK9 have been structurally visualized. A full-length DOCK9/Cdc42 structure is necessary to understand the relative orientation and contribution of all domains, including membrane binding domains that are likely important for correctly orientating DOCK9 at the surface

of the membrane. DOCK9 has been proposed to be autoinhibited, but little is known about the contacts that mediate this autoinhibition or how it is relieved to activate DOCK9 signaling. A structure of full-length autoinhibited DOCK9 would address the autoinhibited conformation. Additionally, there is evidence that full-length DOCK9 is markedly more active than the isolated GEF domain in catalyzing nucleotide exchange suggesting other regions of the protein may also contribute to nucleotide exchange. While the present work is unable to address all of these extended questions, it describes progress made towards purification of constructs necessary for future functional and mechanistic studies of DOCK9.

Experimental procedures

Domain architecture prediction

The domains of DOCK9 were predicted using the SMART domain prediction server (185), 3D-PSSM (186), and the Coils2 server (187). Multiple sequence alignments were created and analyzed with ClustalX (188). Secondary structure prediction was conducted using PSIPRED (189), PROF (190), sspro (191), and samT-99 (192).

Molecular Constructs

DOCK9 (GenBank accession: AB028981) and DOCK6 (GenBank accession: AB037816) were obtained from the Kazusa DNA Research Institute (Kisarazu, Japan). PCR amplification was used to clone DOCK constructs into a modified pET-21a vector (Novagen) using a ligation-independent cloning strategy (LIC) (193). The pLIC-His

vector expresses an N-terminal His_6 tag, a tobacco etch virus (TEV) cleavage site, and the inserted protein construct. Cdc42 was cloned and expressed as previously described and contained an N-terminal His_6 tag followed by a tobacco etch virus (TEV) cleavage site (60). The Cdc42 Δpbr construct, deleting the polybasic region, was generated using deletion mutagenesis as described in the QuickChange manual (Stratagene). Full-length DOCK9 (DOCK9 FL) was PCR-amplified and cloned into the pCR8/GW/Topo entry vector (Invitrogen) and recombined into the pDEST10 destination vector (Invitrogen) used for bacmid formation. The pDEST10 vector encodes a N-terminal His_6 tag, followed by a TEV cleavage site prior to the protein construct. The deletion construct of DOCK9 (DOCK9 LD), deleting residues 1184-1286, was generated using pDEST10 DOCK9 FL as a template and PCR-based deletion mutagenesis as described in the QuickChange manual.

Protein Expression and Purification

Bacterial constructs were transformed into BL21(DE3) *Escherichia coli* cells, grown at 37° C in LB media supplemented with 0.1 mg/mL ampicillin to an A_{600} of 0.6 to 0.8, and then induced for 15 to 18 hours with 1 mM IPTG before cells were harvested. Baculoviruses generated from DOCK9 FL and DOCK9 LD constructs using the Bac-to-Bac system (Invitrogen) were used to infect High-Five insect cells. Cells were harvested 48 hours post-infection. As necessary, cell pellets were frozen at -20° C prior to protein purification.

Cells were resuspended in 20 mM Tris pH 8.0, 300 mM NaCl, 10% glycerol and 5 mM imidazole (N1 buffer) and supplemented with EDTA-free protease inhibitors

(Roche). Cells were lysed with an Emulisflex (Avestin) and centrifuged at 180,000 x g for 50 minutes at 4°C. Supernatant was loaded onto a His-Trap chromatography column (GE Healthcare), washed with 20 CV of N1 and 15 CV of N1 containing 50 mM imidazole prior to elution with N1 containing 400 mM imidazole. Eluted protein was analyzed with SDS-PAGE followed by Coomassie staining and fractions containing the desired protein were pooled. The His_6 tag was cleaved overnight with TEV protease in N1 buffer containing 2 mM DTT.

DOCK9 CT and Cdc42 Δpbr were dialyzed into 20 mM Tris pH 8.0, 20 mM NaCl, 5% glycerol and 1 mM DTT (Q1 buffer) and loaded onto a Source Q anion exchange column (GE Healthcare). Protein was eluted with a linear gradient of Q1 containing 50 to 400 mM NaCl. A two-fold molar excess of Cdc42 was added to DOCK9 CT in the presence of 50 mM EDTA and incubated on ice for 20 minutes. The proteins were then isolated over a Superdex-75 size-exclusion chromatography column (GE Healthcare) in a buffer containing 20 mM Tris pH 8.0, 200 mM NaCl, 5% glycerol, 1 mM DTT, and 20 mM EDTA. The Cdc42 Δpbr/DOCK9 CT complex was dialyzed into a final buffer containing 20 mM Tris pH 8.0, 100 mM NaCl, 5% glycerol, and 1 mM DTT, prior to concentration and storage at -80° C.

Guanine Nucleotide Exchange Assays

Ability of DOCK9 CT to catalyze guanine nucleotide exchange was determined by unloading of MANT-GDP as previously described (194). Exchange assays were performed with a LS-55 fluorescence spectrometer (PerkinElmer) with λ_{ex} = 360 nm and λ_{em} = 430 nm and slits of 5nm. The exchange assay buffer was 50 mM NaCl, 20 mM

Tris-HCl pH 8.0, 5 mM $MgCl_2$, 5% glycerol, 1 mM DTT, and 100 μM GDP. Loading of MANT-GDP was conducted under similar conditions using 400 nM MANT-GDP in place of GDP in the exchange buffer.

Results

Prediction of DOCK9 domain architecture

A combination of secondary and tertiary structure prediction techniques coupled with analyses of multiple sequence alignments were utilized to determine domain boundaries in DOCK9 (**Fig. 6**). The only readily identified domain in DOCK9 predicted by the SMART domain identification server was a PH domain encompassing residues 175-283. The NMR solution structure of this domain has been deposited in the PDB (1WG7) confirming that this region folds as a canonical PH domain.

Although the most highly conserved regions among all DOCK family members are the DHR-1 and DHR-2 regions, visual analysis of a DOCK family multiple sequence alignment highlighted a region of relatively high sequence conservation located prior to the DHR-1 region in DOCKs 9-11. Indeed, the average ClustalX consensus score for this region was 83.3 compared to an average of 61.7 for the rest of the region between the PH domain and DHR-1. Thus, due to higher sequence conservation, this region was designated "conserved region", CR. Fold prediction servers such as 3D-PSSM were unable to identify any known structure for this region, suggesting it could adopt a novel conformation.

The 3D-PSSM server predicted that the DHR-1 contained a C2 domain from 615-747, although the confidence level of the prediction did not reach the 95% level.

However, other DOCK family members such as DOCK1 and DOCK6 were also predicted to contain a C2 domain in this region increasing the likelihood that there is a C2 domain conserved among all DOCK proteins. In fact, the presence of a C2 domain in DOCK1 was confirmed by a recent crystal structure (78).

Secondary structure predictions from PSIPRED, PROF, sspro, and samT-99 servers were generated and aligned to derive a consensus secondary structure prediction for the entire length of DOCK9. A long region of α-helical content was predicted for residues 823-1775 of DOCK9. 3D-PSSM predicted the region to contain ARM repeats. Similar predictions were observed for DOCK1 suggesting ARM repeats were conserved in all DOCK family members. A large region predicted to lack secondary structure, residues 1180-1296, was present in the middle of the putative DOCK9 ARM array. No such region was seen in the DOCK1 ARM array which suggested that 1180-1296 of DOCK9 could encompass an unstructured loop inserted into the middle of the DOCK9 ARM repeats. The presence of ARM repeats has not been structurally confirmed in any DOCK protein, although the presence of the repeats has been suggested by multiple groups.

The region of highest sequence conservation among all DOCK proteins exists in the DHR-2 which is known to be sufficient for catalytic activity. Thus, we predicted that the GEF domain would follow the ARM region. The GEF domain was thought to end at residue 1948 due to prediction of coiled-coil regions at 1949-1981 and 2039-2067 by the COILS2 server. Structure prediction servers such as 3D-PSSM did not predict the presence of any known domains in the GEF region of DOCK9 suggesting it contains a novel fold. The structure of this region did reveal a novel fold, but also illustrated two

incorrect predictions we made regarding DOCK9 domain boundaries. First, the ARM repeats do not extend as far as we predicted. The C-terminal region of our predicted ARM array actually folds as five α-helices that form a dimerization interface. Second, no coiled-coil region was observed in the C-terminus of DOCK9. Instead, this region was part of the GEF domain which contacted Cdc42. Although predictions of DOCK9 domain boundaries were not identical to what we now know the actual structure to contain, they were utilized for the design of DOCK9 expression constructs.

Cloning and expression of DOCK9 yields soluble C-terminal constructs

Having identified domain boundaries in DOCK9, a large array of bacterial expression constructs were generated and tested for their ability to yield soluble recombinant protein (**Fig. 6C**). Clones were generated from a full-length DOCK9 template obtained from the Kasuza DNA research institute. A high-throughput ligation independent cloning (LIC) methodology was utilized which allowed precise domain boundaries to be selected without the need for restriction enzymes or ligation reactions. Constructs started at the N-terminus, the beginning of the PH domain, an area of higher sequence conservation (CR), the C2 domain, the ARM array, after a loop in the ARM array, at the start of the DHR-2, or at the end of the ARM array. Constructs ended at either the C-terminus, the start of a predicted coiled-coil motif, or the end of the ARM array.

With the exception of full-length DOCK9 that could not be successfully PCR amplified, all desired constructs were inserted into the pLIC-His vector that expresses a His_6 affinity tag, a TEV cleavage site, and the inserted construct. Protein expression was

tested in small-scale cultures. In general, constructs of intermediate and smaller lengths expressed well, while larger constructs did not over-express. All proteins observed to over-express were then tested for solubility. Only two C-terminal fragments, DOCK9 1775-2069 (henceforth, DOCK9 CT) and DOCK9 1775-1949 were soluble and able to be purified. Significantly more DOCK9 CT was soluble, and therefore this construct was used in further experiments. DOCK9 CT was predicted to start at the end of the ARM array and proceed to the C-terminus. We now know that DOCK9 CT starts after the dimerization domain, contains the entire GEF domain that directly contacts Cdc42, and ends at the C-terminus.

DOCK9 CT activates Cdc42

Having successfully purified a C-terminal fragment of DOCK9, DOCK9 CT, we tested its ability to bind to Cdc42. A two-fold molar excess of Cdc42 was incubated with DOCK9 CT in the presence of EDTA and the complex was isolated over size-exclusion chromatography. The DOCK9 CT/Cdc42 complex eluted sooner than either DOCK9 CT or Cdc42 alone implying the formation of a higher molecular weight complex (**Fig. 7**). SDS-PAGE analysis of the eluted fractions confirmed the formation of a 1:1 complex of DOCK9 CT and Cdc42. The complex formation between DOCK9 CT and Cdc42 was dependent upon the presence of EDTA, suggesting that Cdc42 is in a nucleotide-free state. The ability to stabilize nucleotide-depleted GTPase is a key characteristic of an exchange factor indicating that DOCK9 might be able to catalyze nucleotide exchange on Cdc42.

The ability of DOCK9 CT to catalyze the unloading of MANT-GDP from various GTPases was tested. DOCK9 CT was observed to activate specifically Cdc42, but not Rac1 nor RhoA (**Fig. 8**). The specificity of DOCK9 CT for Cdc42 recapitulates the previously reported specificity of full-length DOCK9, arguing that DOCK9 CT contains all regions necessary for determining substrate specificity. As expected, DOCK9 CT induced unloading of Cdc42 in a concentration-dependent manner such that the exchange rate was proportional to the amount of DOCK9 CT present. In summary, DOCK9 CT contains all structural elements necessary to specifically catalyze nucleotide exchange on Cdc42.

Purification and crystal trials with DOCK9 CT and a Cdc42/DOCK9 CT complex

Having in hand a DOCK9 construct capable of binding and exchanging on Cdc42, protein purification was optimized and crystallization trials were initiated. The purity original Cdc42/DOCK9 CT complex was not judged sufficient for crystallization (**Fig. 9A**). The additional bands present in the Cdc42/DOCK9 CT complex were believed to be either contaminating proteins or proteolytic fragments of DOCK9 CT. Cdc42 formed a doublet band, which is often observed in full-length GTPases from cleavage of the C-terminal polybasic region. To improve complex purity a more rigorous protein purification scheme was developed (**Fig. 9B**). Cdc42 lacking the polybasic region, Cdc42 Δpbr, was used instead of Cdc42. Both Cdc42 Δpbr and DOCK9 CT were purified from bacterial cells using Ni^{2+} affinity chromatography, followed by TEV-cleavage of the His_6 affinity tag, and anion-exchange chromatography with a Source Q column. Cdc42 Δpbr and DOCK9 CT were then incubated together in the presence of

EDTA and isolated over a Superdex-75 size-exclusion chromatography column. The resulting protein complex was judged to be 99% pure as visualized by SDS-PAGE (**Fig. 9C**). The Cdc42 Δpbr/DOCK9 CT complex was then concentrated and frozen in aliquots for use in crystal trials. DOCK9 CT was also purified in isolation, concentrated, and frozen in aliquots for use in crystal trials.

Numerous crystal trials were undertaken with both DOCK9 CT and the Cdc42 Δpbr/DOCK9 CT complex. As an initial step, commercial screens were tested including both sparse matrix screens and grid screens. Custom-made screens were also utilized to extend the chemical space being tested. No crystals were observed, but conditions giving phase separation or other positive indications were further investigated with custom-made screens. However, these also failed to yield crystals. Screens were conducted at a range of protein concentrations (8 to 32 mg/mL) and temperatures (4°C and 18°C). A high-throughput screen of 1536 different conditions conducted by the Hauptman-Woodward Institute using the Cdc42 Δpbr/DOCK9 CT complex also failed to generate any crystals. In summary, approximately 1800 conditions were tested in-house and another 1536 conditions at the Hauptman-Woodward Institute without generating a positive condition for crystal growth. Having probed many diverse conditions with DOCK9, we used a different isozyme, DOCK6, to increase our chances of successful crystallization.

DOCK6 CT does not behave as a monomer in solution

The homologous clone to DOCK9 CT was created in DOCK6. DOCK6 CT starts at the end of the dimerization domain and proceeds to the C-terminus. Surprisingly,

DOCK6 CT elutes in both monomeric and dimeric forms over size-exclusion chromatography (**Fig. 10**). When the monomer peak is collected and isolated over size-exclusion chromatography both monomer and dimer peaks are again observed, indicating that DOCK6 CT exists in equilibrium of monomer and dimer forms. Our results imply that DOCK6 contains additional contacts outside of the characterized dimerization domain that also contribute to homodimer formation. DOCK6 CT was not observed to stimulate nucleotide exchange on RhoA, Cdc42, or Rac1 leaving its GTPase specificity unknown. Crystal trials with DOCK6 CT did not yield any conditions with crystal formation.

Full-length DOCK9 expressed in insect cells cannot be purified from proteolytic fragments

Since large constructs of DOCK9 were not soluble in bacteria, we used an insect cell expression system to test expression of larger DOCK9 fragments. Full length DOCK9 (DOCK9 FL) was successfully expressed in High Five cells. However, purified DOCK9 FL had many contaminating bands identified as C-terminal proteolytic fragments by western blotting with a DOCK9 specific antibody. The contaminating fragments could not be separated from DOCK9 FL by ion-exchange or size-exclusion chromatography, despite a large size difference between the fragments (**Fig. 11**). N-terminal constructs of DOCK9 are known to bind the C-terminus of DOCK9 so it seems likely that the proteolytic fragments bind to the full-length protein resulting in its inability to be successfully purified. These contaminating proteolytic fragments are likely to be inhibitory to crystallization of full-length DOCK9.

Full-length DOCK9 is significantly more potent than DOCK9 CT in catalyzing nucleotide exchange

Despite the fact that DOCK9 FL was not entirely pure, its GEF activity was tested on Cdc42. DOCK9 FL robustly catalzyed the loading of MANT-GDP onto Cdc42 (**Fig. 12**). At a concentration 20-fold less than DOCK9 CT, DOCK9 FL was significantly more active towards Cdc42. This strongly implies that regions outside of the putative GEF domain may enhance the ability of DOCK9 to exchange on Cdc42. While we cannot exclude the possibility that one of the contaminating proteolytic fragments is responsible for the enhanced nucleotide exchange activity, it would not alter our conclusion that other domains likely contribute to the nucleotide exchange mechanism. This finding is supported by a similar result obtained with DOCK11 where the full-length protein was significantly more active towards Cdc42 than the isolated DHR-2 domain (143). Thus, the possibility that domains outside of the identified GEF region influence exchange may be a conserved feature of DOCK proteins and indicate a pressing need for more complete structural characterization.

Deletion of a loop containing proteolysis sites does not abolish proteolysis of DOCK9

Secondary structure analyses of DOCK9 predicted a large ~100 residue insertion in the middle of the ARM array (**Fig. 13**). Trypsin digestion of the loop region of DOCK9 has been previously reported (195). Furthermore, the sizes of the observed proteolytic fragments of DOCK9 FL correspond to a cut site within the loop region. N-terminal Edman sequencing of proteolytic fragments revealed two sites of proteolysis

within the identified loop. A DOCK9 construct lacking this loop, DOCK9 LD, was produced in insect cells with the expectation that the proteolysis would be greatly reduced. To our surprise, significant proteolysis still occurred despite deletion of the known cleavage sites. Like the original proteolysis fragments, these breakdown products could not be separated by traditional chromatography techniques. The absence of the major proteolytic sites likely renders other regions of the protein more susceptible to proteolysis. For example, there is a smaller putative loop before the large unstructured region. This loop may be a minor site of proteolysis in DOCK9 FL, yet become a major proteolytic site in DOCK9 LD. Therefore, it may be wiser to screen for protease-resistant family members than attempt to delete all proteolytic sites in DOCK9 and potentially destabalize its structure.

Discussion

DOCK9 activates Cdc42 in a number of *in vitro* and cell-based activity assays (72,76,139,144). The aim of this work was to uncover the structural details of DOCK9 activation of Cdc42. At the outset of this work only a PH domain was readily identified by standard domain prediction tools. Using a combination of secondary and tertiary structure prediction tools coupled with analyses of multiple sequence alignments the boundaries of DOCK9 domains were predicted. The bioinformatics analyses correctly identified a C2 domain in the DHR-1 region as well as correctly predicted the start of the GEF domain that binds to Cdc42. ARM arrays were predicted to extend until the start of the GEF domain, but it is now known that an α-helical dimerization domain exists prior

to the GEF domain. The prediction of extended ARM repeats was based on results from 3D-PSSM that aligned this region to a known ARM-repeat protein. Although it has not been directly observed, DOCK9 is still believed to have an extended region of ARM repeats between the C2 and dimerization domains. Predictions of coiled-coil regions in the C-terminus of DOCK9 were also not supported by the actual crystal structure which revealed the GEF domain to extend to near the C-terminus of the protein. Although not accurate in every detail, the predicted domain architecture was successfully utilized to design a minimal DOCK9 construct capable of catalyzing nucleotide exchange.

Utilizing the identification of putative domain boundaries, multiple constructs of DOCK9 were tested for expression and solubility. Large constructs of DOCK9 either did not express or were not soluble in bacterial cells. However, a C-terminal fragment of DOCK9, DOCK9 CT, was successfully purified from bacteria and shown both to bind nucleotide-depleted Cdc42 and to catalyze nucleotide exchange in a fluorescence-based nucleotide exchange assay. Interestingly, comparison of DOCK9 CT with the crystal structure of DOCK9 bound to Cdc42 shows that DOCK9 CT contains all the structural elements needed to contact Cdc42, but none of the elements involved in dimerization DOCK9. It has been suggested that Cdc42 shows positive cooperativity in binding to DOCK9 dimers so that binding of Cdc42 to one half of the dimer may increase the affinity of the other half for an additional Cdc42 molecule (121). Comparison of DOCK9 CT with larger DOCK9 fragments may be of use in dissecting the role of dimerization and potential cooperativity of Cdc42 binding in the nucleotide exchange reaction.

Protein purification of a Cdc42 Δpbr/DOCK9 CT complex was optimized and extensive crystallization trials were attempted. Unfortunately, conditions conducive to

the formation of crystals were not obtained. Comparison of our construct with the published DOCK9 construct used for crystallization suggests two major factors that may have contributed to the inability of the Cdc42 Δpbr/DOCK9 CT to crystallize. First, the published crystal structure contained an additional dimerization domain that was not present in our construct. This domain mediated extensive crystal contacts that likely contributed to the formation of crystals. Second, in the published crystal structure a 14-residue insertion at the beginning of the GEF domain was identified and deleted. This insertion was present in our construct and may have contributed to increased conformational flexibility that inhibited crystallization. With accurate knowledge of the domain structure of DOCK9 in hand, attempts to crystallize other DOCK family members bound to their cognate GTPases can be optimized. Such structures will be important in determining how DOCK proteins achieve their specificity in GTPase activation.

Full length DOCK9 was able to be expressed and partially purified from insect cells, but could not be separated from proteolytic fragments. The inability of proteolytic fragments to be separated from full length DOCK9 likely stems from the fact that multiple portions of the DOCK9 N-terminus bind to the DOCK9 C-terminus to cause autoinhibition (121). Thus, both N and C-terminal proteolytic fragments can be expected to interact with full-length DOCK9. Proteolytic sites were identified in a loop in the middle of the ARM array. However, despite deletion of this loop, DOCK9 LD was still liable to proteolysis. Thus, additional proteolytic sites were made accessible upon deletion of the main cleavage site. Multiple proteases are expressed in insect cells and their inhibition is both difficult to achieve and cost-prohibitive (196-197). While other

DOCK proteins such as DOCK1 have also been found to be proteolyzed in insect cells, DOCK2 was not proteolyzed. Thus, it may be more efficient to screen for DOCK proteins resistant to proteolysis than to attempt to identify and delete proteolytic sites from protease sensitive DOCK proteins. Alternatively, other eukaryotic expression systems such as *Pichia pastoris* that may not cause proteolysis could be utilized to express full length DOCK9.

The activity of DOCK9 FL was found to be markedly more robust than DOCK9 CT implicating domains outside of the core GEF domain as contributing to nucleotide exchange. This highlights a pressing need for more complete structural information in regards to DOCK9 activation of Cdc42. A structure of full length DOCK9 bound to Cdc42 would address the contribution of additional domains as well as allow for predictions of how the PH and C2 membrane binding domains orient DOCK9 at the plasma membrane. In addition, there is evidence that DOCK9 is autoinhibited in the basal state and structural information illustrating such a conformation would help us to understand the molecular details of DOCK9 autoinhibition and propose a model for its regulation. If full length DOCK9 cannot be purified from alternative expression systems, the close homologues DOCK10 or DOCK11 could be used in structural studies. Although the basic mechanism of DOCK catalyzed nucleotide exchange is now known, outstanding questions in the field require further structural models to be solved.

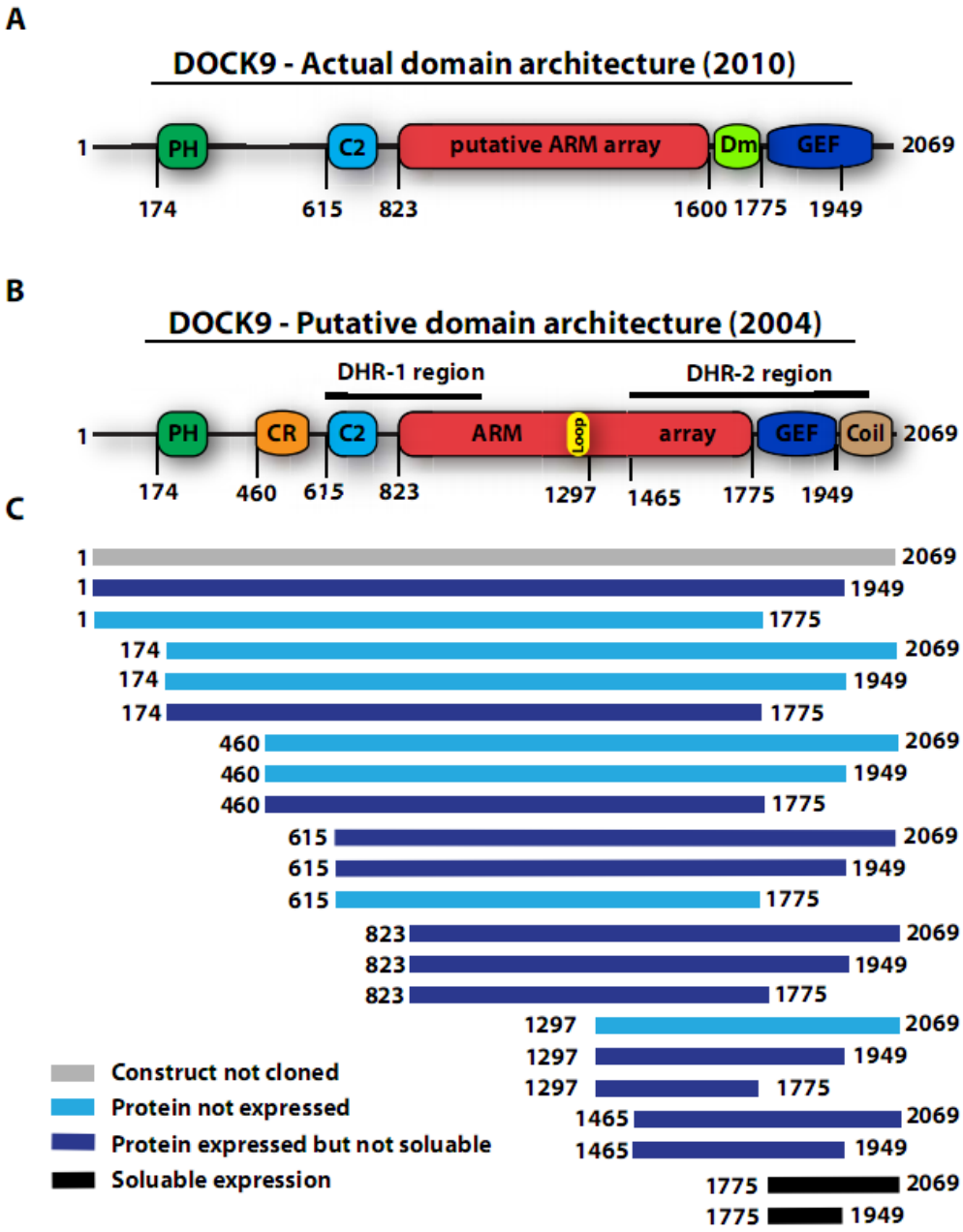

Figure 6: Cloning and expression of DOCK9 yields soluble C-terminal constructs. The actual (*A*) and putative (*B*) domain architecture of DOCK9. *C*, DOCK9 constructs were designed and cloned into the pLIC-His vector and tested for expression and solubility. Full-length DOCK9 was not successfully cloned (grey) while other constructs were not expressed (cyan) or not soluble (dark blue). Two C-terminal fragments yielded soluble recombinant protein (black).

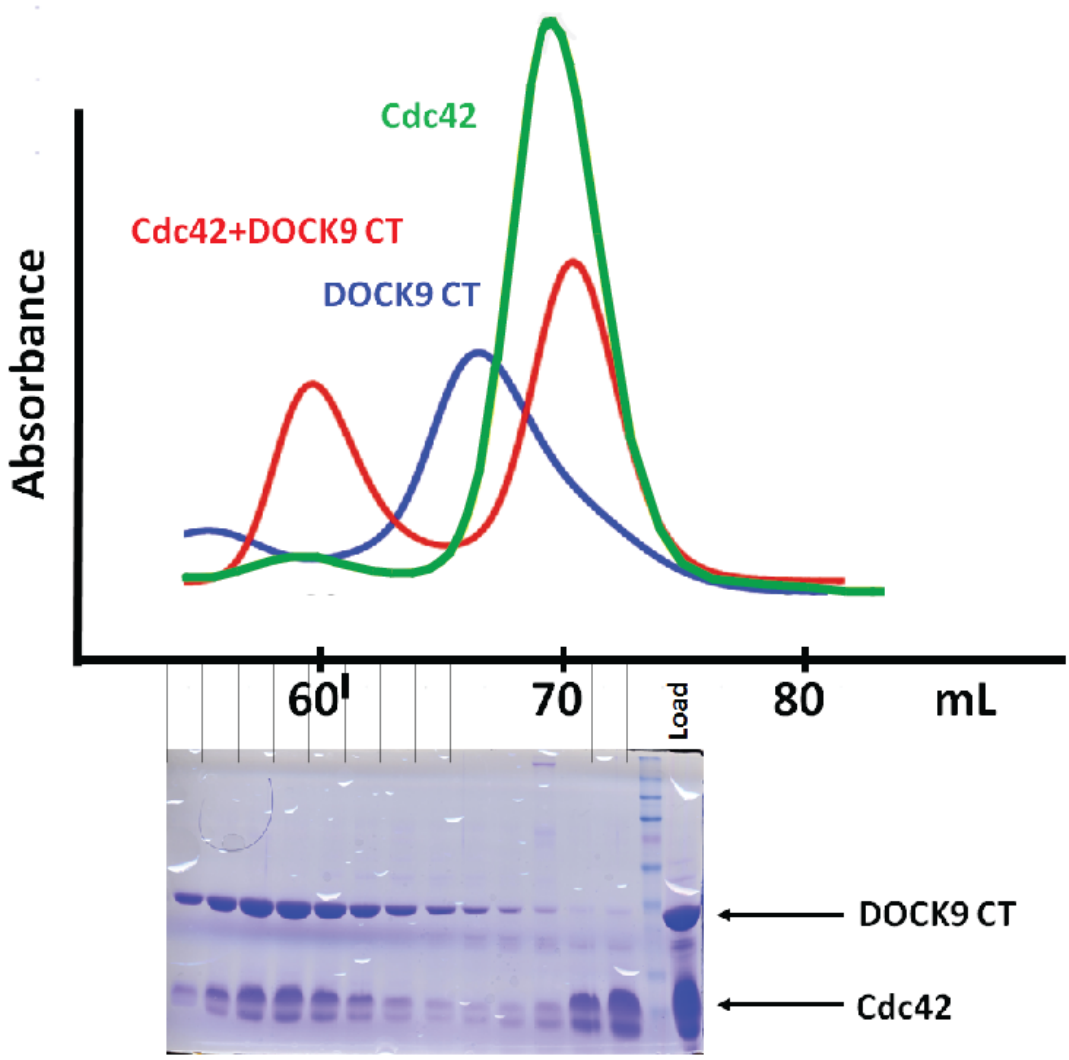

Figure 7: DOCK9 CT forms a complex with nucleotide-depleted Cdc42.
DOCK9 CT and a two-fold molar excess of Cdc42 were incubated in an EDTA-containing buffer for 30 min and isolated over size exclusion chromatography. The Cdc42/DOCK9 CT complex (red) was shifted relative to the elution profile of DOCK9 CT (blue) or Cdc42 (green) alone. SDS-PAGE shows the formation of a 1:1 complex of DOCK9 CT and Cdc42.

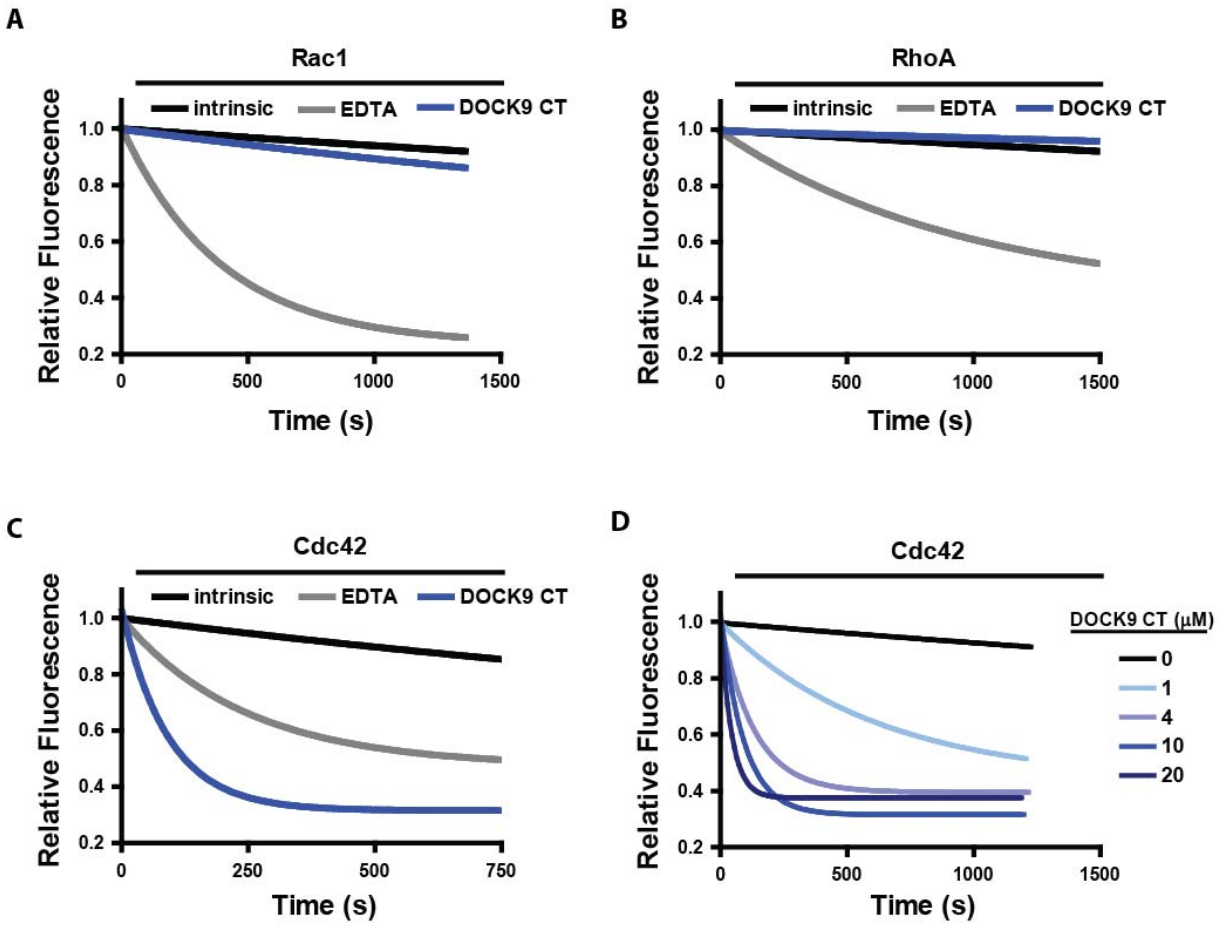

Figure 8: DOCK9 CT specifically activates Cdc42 in a concentration dependent manner.

Rac1 (*A*), RhoA (*B*), and Cdc42 (*C*) GTPases (500 nM) preloaded with MANT-GDP were incubated in exchange buffer and nucleotide exchange was stimulated by the addition of EDTA (10 μM) or DOCK9 CT (10 μM). *D*, Cdc42 (500 nM) preloaded with MANT-GDP was incubated in exchange buffer and nucleotide exchange was stimulated by the indicated amounts of DOCK9 CT.

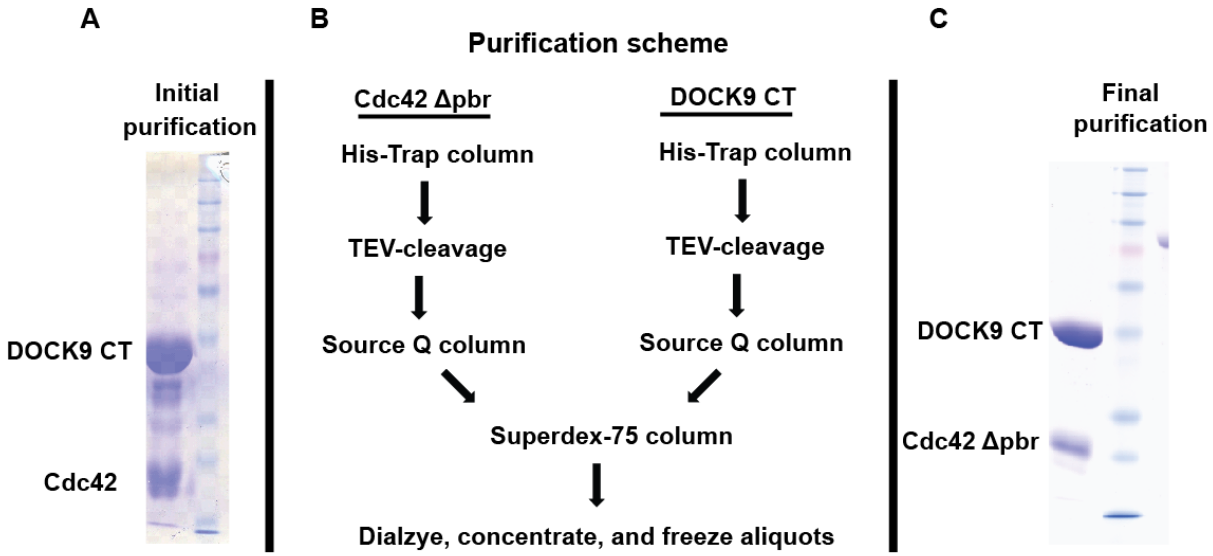

Figure 9: Purification scheme for crystal trial quality Cdc42 Δpbr/DOCK9 CT complex.

A, Initial purification of Cdc42/DOCK9 CT yielded a final complex with numerous contaminants. *B*, Addition of ion-exchange chromatography and the use of Cdc42 lacking the C-terminal polybasic region, Cdc42 Δpbr, yielded the final purification scheme for a Cdc42 Δpbr/DOCK9 CT complex. *C*, Final purification gave a complex void of any observable contaminants.

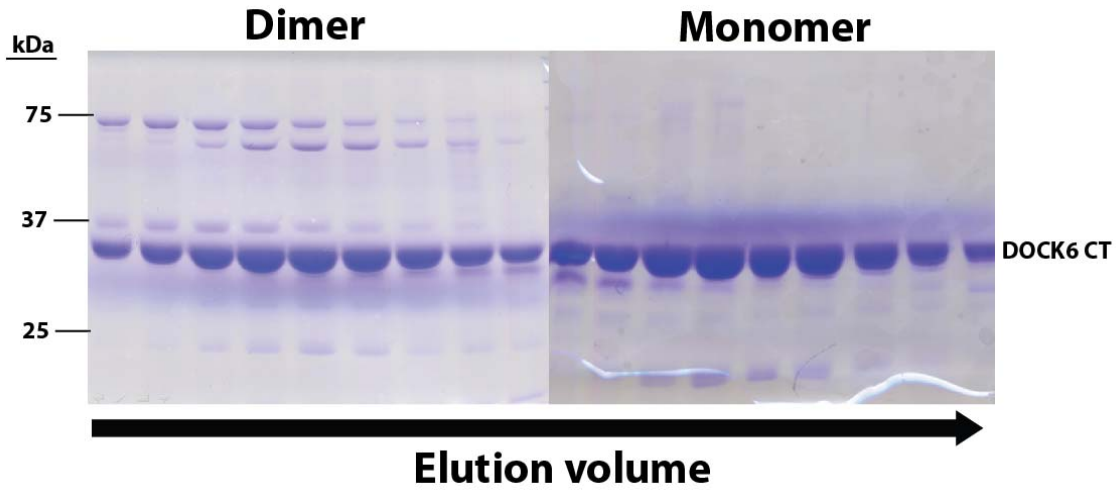

Figure 10: DOCK6 CT elutes as both monomer and dimer from size exclusion chromatography.

DOCK6 CT was eluted from a Sephacryl-200 size-exclusion chromatography column and fractions were analyzed via SDS-PAGE. Two major peaks were observed which corresponded to predicted molecular weights of a DOCK6 CT monomer (37 kDa) and dimer (75 kDa).

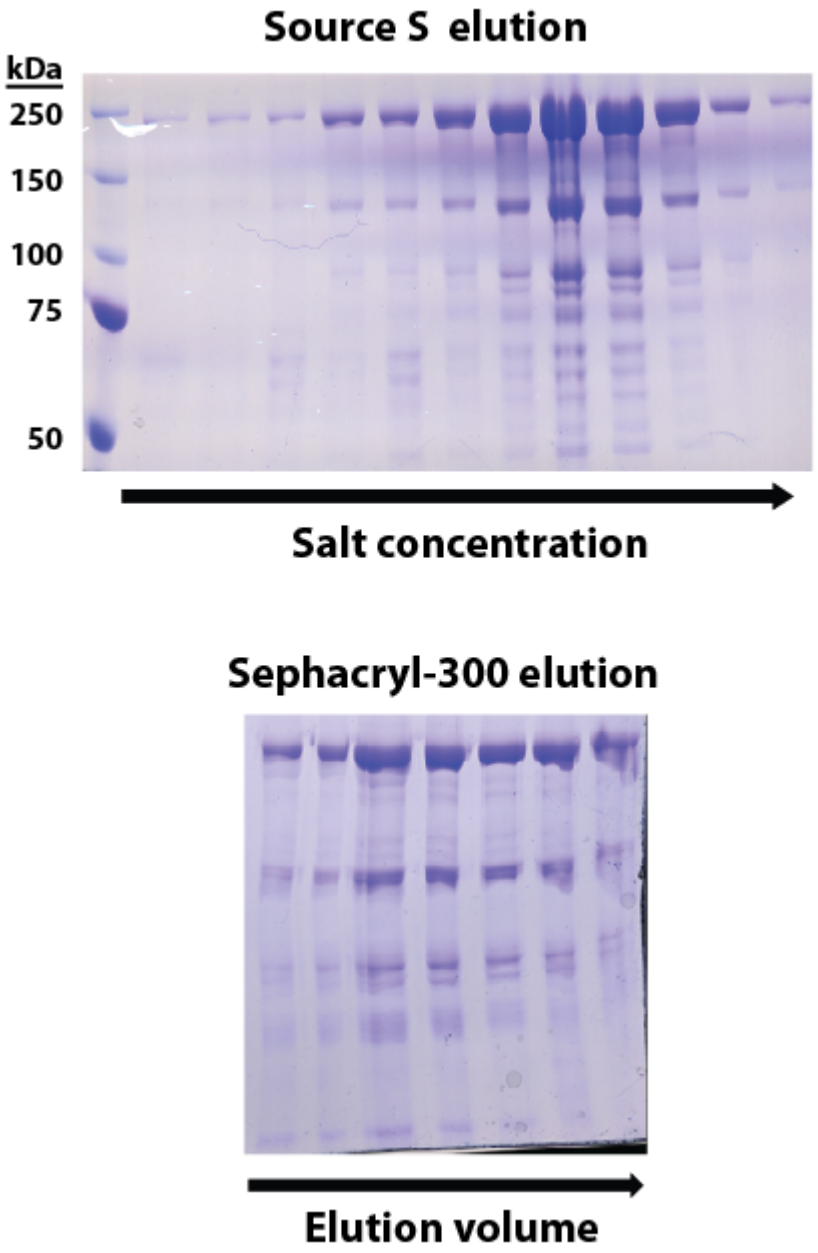

Figure 11: Full-length DOCK9 cannot be separated from proteolytic fragments by ion exchange or size exclusion chromatography.

Full-length DOCK9 purified from insect cells was found to contain excessive proteolytic fragments. Attempts to separate full-length DOCK9 from breakdown products with Source S ion exchange chromatography (top) or Sephacryl-300 size exclusion chromatography (bottom) did not yield appreciable purification.

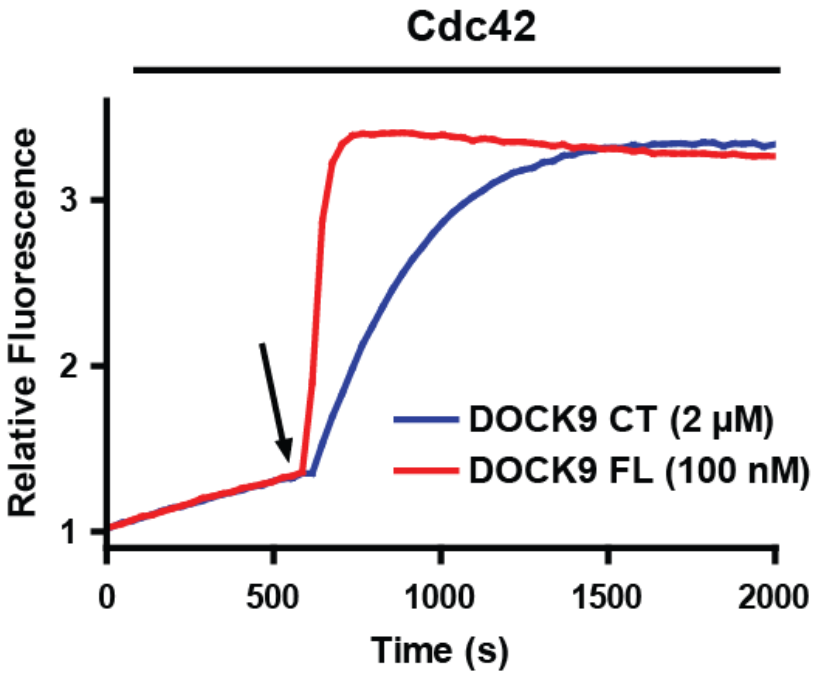

Figure 12: Full-length DOCK9 is significantly more potent than DOCK9 CT in catalyzing nucleotide exchange.

Cdc42 (1 µM) was incubated in exchange buffer containing MANT-GDP and at the indicated time (arrow) nucleotide exchange was stimulated with the addition of DOCK9 CT (2 µM) or DOCK9 FL (100 nM).

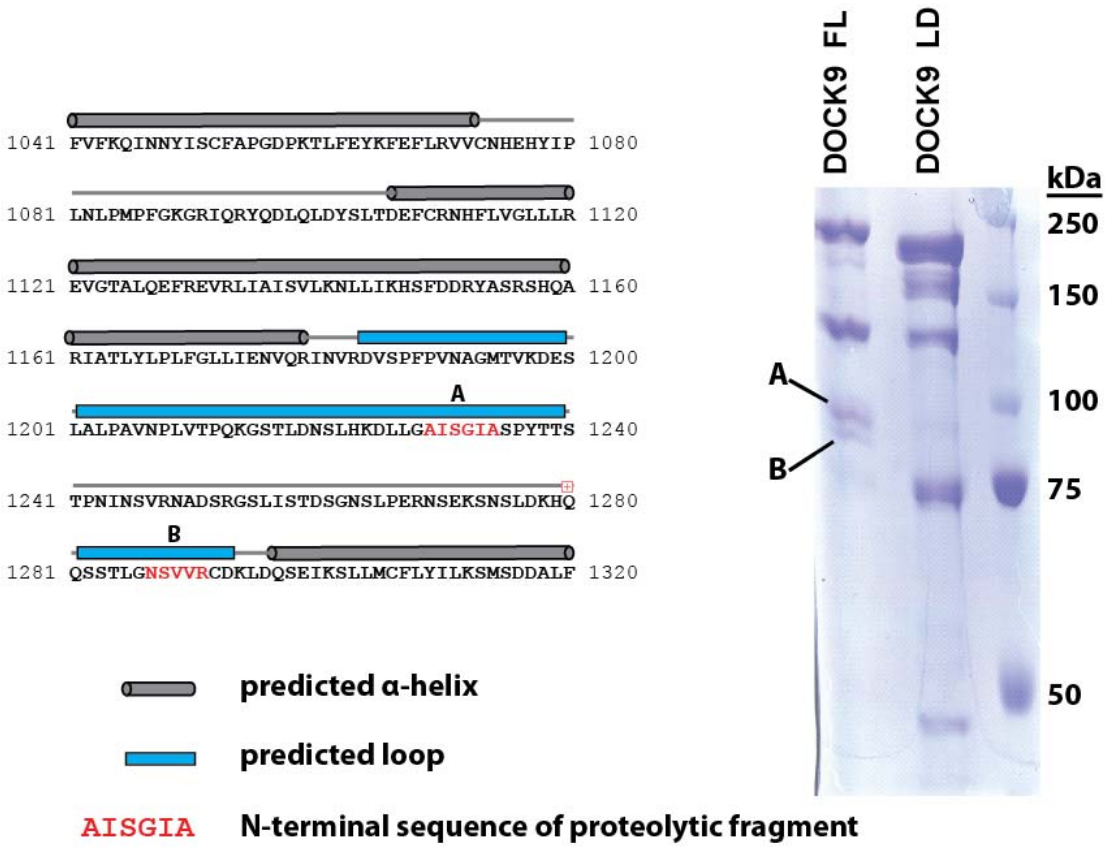

Figure 13: Deletion of loop containing identified proteolysis sites does not inhibit proteolysis.

The predicted secondary structure of DOCK9 indicated the presence of a large unstructured loop in the middle of the ARM array (left, residues 1180-1296). Proteolytic fragments (A and B) were identified by Edman sequencing (red letters) confirming that full-length DOCK9 (DOCK9 FL) was being proteolyzed in the predicted loop. However, DOCK9 lacking the loop region (DOCK9 LD) was still proteolyzed (right).

CHAPTER 3: SMGGDS IS A GUANINE NUCLEOTIDE EXCHANGE FACTOR THAT SPECIFICALLY ACTIVATES RHOA AND RHOC

Background

Rho family GTPases play essential signaling roles in a wide variety of cellular processes ranging from migration (*20*), to cell cycle regulation (*21*), neurite outgrowth (20,24), exocytosis (*23*), and mitosis (198). As nucleotide-dependent switches, Rho GTPases exist in two states: an inactive GDP-bound state, and a GTP-bound state in which they signal to downstream partners. Rho family guanine nucleotide exchange factors (RhoGEFs) stimulate the activation of Rho GTPases by promoting the dissociation of GDP from the inactive GTPase and the subsequent binding of GTP. High-resolution structures have allowed elucidation of the mechanism of nucleotide exchange for both Dbl family (58-59) and DOCK family (79) RhoGEFs. SmgGDS is a GEF comprised entirely of armadillo (ARM) repeats that does not belong to either the Dbl or DOCK family (145). Despite its early identification as a GEF (147,151), details on the mechanism by which SmgGDS stimulates nucleotide exchange remain sparse.

SmgGDS was originally characterized as activating multiple GTPases including Ras family members Rap1A (146), Rap1B (147), and K-ras (146), as well as the Rho family members Rac1(148), Rac2 (149), Cdc42 (150), RhoA (151), and RhoB (151). However, early experiments using SmgGDS were performed with crudely purified

protein samples (147,151) and yielded inconsistent reports of GTPase specificity. In addition, some groups claimed SmgGDS only activated prenylated small GTPases (146,162) while others reported activity independent of prenylation state (149,155). SmgGDS has only been well characterized for its ability to stimulate nucleotide exchange upon RhoA (155), but nothing is known about the regions of SmgGDS necessary to catalyze exchange.

The biological function of SmgGDS has also been questioned. SmgGDS has not yet been shown to activate any small GTPases *in vivo.* An alternative function of controlling the nucleocytoplasmic shuttling of small GTPases has been proposed (167). Despite the unknowns surrounding the function of SmgGDS, it has been implicated in a number of disease states. For example, SmgGDS has been found to be upregulated in both prostate cancer (183) and non-small lung carcinoma (182) where it promotes both proliferation and migration of cancer cells.

This study revisits the substrate specificity of SmgGDS using highly purified proteins, and for the first time implicates specific residues of SmgGDS as being necessary for interaction and exchange upon small GTPases. We show that SmgGDS is solely able to activate RhoA and RhoC using *in vitro* nucleotide exchange assays. The interaction depends on the presence of an intact polybasic region on the C-terminus of the RhoA, but does not utilize the same mechanism of nucleotide exchange as seen for traditional Dbl family exchange factors. We also demonstrate that ability to exchange upon RhoA requires both an electronegative surface patch and a highly conserved binding pocket on the surface of SmgGDS. Finally, we show that transfection of

SmgGDS leads to a specific increase in the level of active RhoA and RhoC in cells, but not other GTPases including RhoB.

Experimental procedures

Molecular Constructs

SmgGDS-558 (GenBank accession: NM_174666, bovine) was kindly provided by L. Quilliam. SmgGDS (GenBank accession: NM_001100426) was obtained from the IMAGE Consortium. PCR amplification was used to subclone each isoform into a modified pET-21a vector (Novagen) using a ligation-independent cloning strategy (LIC) (193). The pLIC-His vector expresses an N-terminal His_6 tag, a TEV cleavage site, and the inserted protein construct. Rho and Ras family GTPases were cloned into bacterial expression vectors as previously described (59-60,199). The mammalian expression constructs of human SmgGDS-558 (GenBank accession: NP_001093899) and SmgGDS (GenBank accession: NP_001093897) in pcDNA3.1+ (Invitrogen) were kind gifts from C. Williams. Point mutations in constructs were generated using the QuickChange site-directed mutagenesis kit (Stratagene) and were verified by DNA sequencing of the entire open reading frame.

Protein Expression and Purification

SmgGDS and GTPase constructs were transformed into BL21(DE3) *Escherichia coli*, grown at 37° C in LB media supplemented with 0.1 mg/mL ampicillin to an A_{600} of 0.6 to 0.8, and then induced for 15 to 18 hours with 200 μM IPTG. Cells were harvested, lysed, and His_6-tagged proteins were purified via Ni^{2+} affinity chromatography. The His_6

tag was cleaved from SmgGDS and RhoA with TEV protease. Proteins were further purified with size-exclusion chromatography. Proteins were then concentrated and final concentration was determined using A_{280} and the extinction coefficient as calculated by ProtParam (ExPASy server (200)) prior to storage at -80° C.

Gel Filtration Binding Assay

SmgGDS was incubated on ice for 30 minutes with a two-fold molar excess of RhoA in buffer containing 150 mM NaCl, 20 mM Tris-HCl pH 8.0, 5% glycerol, 1 mM DTT, and either 25 mM EDTA (nucleotide-free buffer), 2 mM $MgCl_2$ and 30 μM GDP (GDP buffer), or 2 mM $MgCl_2$ and 30 μM GTPγS (GTPγS buffer). Protein was then separated using a Superdex-75 gel filtration column (GE Healthcare) and fractions were analyzed via SDS-PAGE.

Guanine Nucleotide Exchange Assays

Ability of SmgGDS to catalyze guanine nucleotide exchange was determined with a MANT-GDP loading assay as previously described (194). Exchange assays were performed with a LS-55 fluorescence spectrometer (PerkinElmer) with λ_{ex} = 360 nm and λ_{em} = 430 nm and slits of 5nm. The exchange assay buffer was 50 mM NaCl, 20 mM Tris-HCl pH 8.0, 5 mM $MgCl_2$, 5% glycerol, 1 mM DTT, and 400 nM MANT-GDP. Dissociation of nucleotide from RhoA preloaded with MANT-GDP was measured in a buffer containing 50 mM NaCl, 20 mM Tris-HCl pH 8.0, 100 μM $MgCl_2$, and 100 μM free GDP.

Homology Modeling

Since there is not a template structure in the PDB with enough ARM repeats to model all of SmgGDS, the SmgGDS sequence was broken into four sections of overlapping ARM repeats (1-184, 118-362, 254-528, and 472-608) and a homology model of each section was created using β-catenin (pdb: 2Z6G) as a template (201). Alignment of SmgGDS repeats to β-catenin was computed with the HHpred server (http://toolkit.tuebingen.mpg.de/hhpred) and homology models were created using the InsightII molecular modeling package (Accelrys, San Diego, CA). Overlapping repeats were superimposed to yield a single model of SmgGDS that was analyzed with Profiles3D, yielding a score of 0.87, indicating the model characteristics are similar to known protein structures (202). The electrostatic surface potential of the homology model was calculated using PBEQ Solver (http://www.charmm-gui.org, (203)) and displayed using PyMOL (http://www.pymol.org/). A multiple sequence alignment of SmgGDS sequences from 23 different species was aligned using ClustalX (188). The clustal consensus scores from the MSA were mapped onto the surface of SmgGDS and colored according to degree of sequence conservation.

Rho and Rac Family Activity Assays

Rho family activity assays were performed as previously described (34). SmgGDS constructs were transfected into HEK293 cells using Lipofectamine 2000 according to manufacturer's instructions (Invitrogen). Cells were grown for 24 hours and lysed in 50 mM Tris-HCl pH 7.6, 500 mM NaCl, 1% Triton X-100, 0.1% SDS, 0.5% deoxycholate, 10 mM $MgCl_2$, 200 μM orthovanadate and protease inhibitors to assess

Rho activity and 50 mM Tris-HCl pH 7.6, 150 mM NaCl, 1% Triton X-100, 10 mM $MgCl_2$, 200 μM orthovanadate and protease inhibitors to assess Rac activity. Lysates were clarified by centrifugation, equalized for total volume and protein concentration, and rotated for 30 minutes with 30 μg of purified GST-RBD (Rho binding domain of either Rhotekin for Rho or PAK for Rac bound to glutathione-sepharose beads). The bead pellets were washed in 50 mM Tris-HCl pH 7.6, 150 mM NaCl, 1% Triton X-100, 10 mM $MgCl_2$, 200 μM orthovanadate, with protease inhibitors, and subsequently processed for SDS-PAGE.

Western Blotting

Cell lysates were subjected to SDS-PAGE and transferred to polyvinylidene fluoride (PVDF) membranes (Millipore). For western blotting, membranes were incubated with primary antibody overnight at 4°C and secondary antibodies for 1 hour at room temperature. Blots were developed with SuperSignal West Pico Chemiluminescent Substrate (Pierce) and visualized using Kodak BioMax film (Kodak).

Circular Dichroism

Purified SmgGDS proteins at a concentration of 0.15 mg/mL in 20 mM sodium phosphate, pH 7.5 were analyzed by circular dichroism at 25° C using a Pistar-180 spectrometer with a 0.5 mM path length. Data were plotted as mean residue ellipticity as a function of wavelength and were analyzed for α-helical content using CDPro software (http://lamar.colostate.edu/~sreeram/CDPro).

Results

SmgGDS specifically activates RhoA and RhoC

Due to the fact that many earlier characterizations of SmgGDS were not performed with highly purified recombinant proteins and used sub-optimal activity assays, we wished to revisit the substrate specificity of SmgGDS using highly purified proteins and a robust fluorescence-based guanine nucleotide exchange assay. Consequently, we purified to homogeneity SmgGDS and a panel of small GTPases previously reported to be SmgGDS substrates. We then tested the ability of SmgGDS to promote nucleotide exchange on the small GTPases using a time resolved fluorescence assay where the fluorescence of MANT-GDP increases as it is loaded onto the GTPase. In this assay, SmgGDS increases the nucleotide exchange rate for RhoA and RhoC and does not activate other Rho or Ras family GTPases (**Fig. 14**). The ability of each GTPase to be stimulated was confirmed by addition of EDTA, which chelates the Mg^{2+} ion that stabilizes nucleotide binding. Our results suggest that previous characterizations of SmgGDS as a promiscuous exchange factor were in error and that it is highly specific for RhoA and RhoC.

Activation of RhoA by SmgGDS is catalytic and independent of SmgGDS isoform

Two major splice variants of SmgGDS are expressed in human tissues, a 608 amino acid protein (SmgGDS) containing 14 ARM repeats and a 558 amino acid protein (SmgGDS-558) containing 13 ARM repeats (**Fig. 15A**). The fourth ARM repeat, ARM D, is not present in the shorter isoform. We first compared the ability of each isoform to activate RhoA, the most well characterized SmgGDS substrate (**Fig. 15B**). There is no

significant difference in the activation of RhoA by SmgGDS-558 and SmgGDS, indicating that the isoforms have equivalent nucleotide exchange abilities *in vitro*. We also detected no difference in GTPase specificity between SmgGDS and SmgGDS-558 (data not shown). Thus, the remainder of the *in vitro* exchange assays in this paper utilize the longer SmgGDS isoform.

Consistent with the ability of GEFs to stimulate both the loading and unloading of guanine nucleotides at catalytic concentrations, SmgGDS catalyzes the unloading of MANT-GDP from RhoA as indicated by a decrease in fluorescence, as well as the loading of MANT-GDP in a concentration dependent fashion (**Fig. 15C-D**). The difference in the time scales for the reactions may reflect both the slower intrinsic exchange of MANT-labeled nucleotides from RhoA and the potential of the fluorophore to interfere with exchange factor binding (204). Due to these potential caveats, loading of labeled nucleotide is the preferred method to examine SmgGDS function.

SmgGDS preferentially forms a high-affinity complex with RhoA in the nucleotide-free state

Another key characteristic of GEFs is that they specifically stabilize the nucleotide-free state of cognate GTPases allowing weak interactions with nucleotide-bound forms, but very tight binding to nucleotide-free versions. Incubating a small GTPase in buffer containing excess Mg^{2+} and nucleotide stabilizes the nucleotide-bound form. Inclusion of excess EDTA in the buffer will lead to chealation of the Mg^{2+} ion that stabilizes nucleotide binding, thus favoring formation of the nucleotide-free state. When SmgGDS and a two-fold molar excess of RhoA are isolated over size exclusion

chromatography in different buffer conditions, SmgGDS only binds RhoA in the nucleotide-depleted state as expected for a guanine nucleotide exchange factor (**Fig. 16**).

SmgGDS has a unique exchange mechanism compared to Dbl family RhoGEFs and SopE

The exchange mechanism has been elucidated for canonical Dbl family GEFs and the bacterial exchange factor SopE. Despite having distinct structures, both SopE and Dbl family GEFs stabilize nearly identical conformations of the GTPase switch regions (**Fig. 17A**). Both GEFs stabilize a conformation where a salt bridge is formed between K18 of the phosphate binding loop and E64 of switch 2. The side chain of A61 in switch 2 is stabilized in a position that directly occludes Mg^{2+} binding and is thus a major contributor for nucleotide exchange. Mutations in either A61 or E64 of RhoA ablate the ability of Dbs (a Dbl family GEF) to catalyze exchange. These mutations have no effect on SmgGDS catalyzed exchange (**Fig. 17B**). The fact that RhoA residues A61 and E64 are not essential for SmgGDS catalyzed nucleotide exchange supports the conclusion that SmgGDS bound to nucleotide-free RhoA has a distinct conformation from previously seen Rho GEFs that require A61 to stimulate nucleotide exchange.

SmgGDS requires a C-terminal polybasic region to interact with RhoA

SmgGDS did not activate RhoB, yet activated RhoA and RhoC. Sequence alignment of RhoA, B, and C reveals a high degree of similarity with the exception of the C-terminal polybasic regions (**Fig. 18A**). RhoB lacks a strong polybasic sequence, implicating this motif as critical for SmgGDS binding. Size-exclusion chromatography

was performed with SmgGDS and a RhoA truncation lacking the polybasic region, RhoA Δpbr, and a stable complex could not be observed in the nucleotide-depleted state (**Fig. 18B**). In addition, SmgGDS was unable to catalyze exchange on either RhoA Δpbr or RhoA containing mutations in the polybasic region, while Dbs retained activity on all proteins (**Fig. 18C**). These data are consistent with previous reports that SmgGDS requires the polybasic region of RhoA for binding. To test the specificity of the RhoA/SmgGDS interaction, chimeric GTPases with the polybasic region of Rac1 replacing that region in RhoA, RhoA (Rac1pbr), and the polybasic region of RhoA replacing that in Rac1, Rac1 (RhoApbr), were generated. SmgGDS was able to catalyze nucleotide exchange on RhoA (Rac1pbr), but not on Rac1 (RhoApbr) (**Fig. 18D**). These data illustrate that the presence of a strong polybasic region is necessary for the interaction of SmgGDS and RhoA, but the specific sequence of the polybasic is not essential. Importantly, it also implies that there are differences outside of the polybasic region that account for the ability of SmgGDS to exchange upon RhoA and not Rac1.

Homology model of SmgGDS reveals conserved electrostatic surface patch required for RhoA interaction

The necessity of a polybasic region on RhoA suggests that it interacts with a complementary acidic region on the surface of SmgGDS. Lacking a high-resolution structure of SmgGDS, we created a homology model using the armadillo repeats of β-catenin as a template structure. The electrostatic potential surface was calculated and a region of strong electronegative charge was observed near the center of the protein (**Fig. 19A**). To identify which acidic residues were likely to be critical for interaction with

RhoA, a multiple sequence alignment was created with SmgGDS sequences from 23 different species. The alignment was analyzed to yield residues that were completely conserved as acidic in all species or the majority of species (**Fig. 19B**). Charge reversal mutations were made in clusters of acidic residues and the mutant SmgGDS proteins were tested for their ability to activate RhoA (**Fig. 19C**). As predicted, mutating any of the conserved acidic clusters that comprise the electronegative patch on SmgGDS ablated the ability of SmgGDS to activate RhoA. However, mutating a large modestly conserved acidic cluster on the opposite surface of SmgGDS had no effect on the ability of SmgGDS to activate RhoA. Circular dichroism was used to confirm that mutant SmgGDS proteins had the same secondary structure characteristics as the wild-type protein and were not misfolded (**Fig. 21**).

Mapping sequence conservation on homology model identifies a highly conserved binding surface essential for nucleotide exchange

Having observed that the conserved acidic residues were essential for the ability of SmgGDS to activate RhoA, we investigated whether there were additional conserved regions on the surface of SmgGDS that could be implicated in the interaction with RhoA. Using the multiple sequence alignment of SmgGDS we mapped the sequence conservation of SmgGDS onto the surface of the homology model (**Fig. 19D**). We observed two regions of high sequence conservation. The smaller region of sequence conservation was between the N-terminus and the electronegative patch while the larger region was in a super-helical groove formed by the curvature of the ARM repeats and was C-terminal to the electronegative patch. Mutations of conserved residues were

generated and mutant proteins were tested for their ability to activate RhoA (**Fig. 19E**). Mutating the N-terminal conserved region (R112A+N116A+Y119A) did not affect the ability of SmgGDS to exchange upon RhoA, while mutations in the C-terminal binding groove abolished (N388A; H380A+S384A; N343A+R346A) or impaired (M307A; N395A+K396) activity on RhoA. None of the SmgGDS mutations had altered α-helical content as measured by circular dichroism, indicating that they retained their proper fold (**Fig. 21**). These results suggest that the conserved binding groove of SmgGDS might bind the body of the GTPase with the polybasic tail extending over the electronegative surface patch.

SmgGDS specifically activates RhoA and RhoC in vivo

Having characterized specificity for RhoA and RhoC *in vitro*, we wished to test the *in vivo* specificity of SmgGDS. SmgGDS isoforms were transiently transfected into HEK293 cells and lysates were probed for active Rho proteins using an effector pull-down assay (**Fig. 20A**). Interestingly, only the SmgGDS-558 isoform was able to activate RhoA *in vivo,* in contrast to the *in vitro* observation of equivalent activities between isoforms. Activation was abolished with mutations in the conserved binding groove of SmgGDS. Similar assays conducted with Rac1 GTPase showed no activation of Rac1 by either SmgGDS isoform (**Fig. 20B**). The ability of SmgGDS-558 to specifically activate RhoA and RhoC, but not RhoB was tested by blotting with antibodies specific to each Rho isoform (**Fig. 20C**.) SmgGDS-558 was observed only to activate RhoA and RhoC as seen *in vitro*. These results demonstrate that the *in vivo*

specificity of SmgGDS-558 recapitulates our *in vitro* results, but indicates a need for future studies concerning the role of each SmgGDS isoform *in vivo*.

SmgGDS crystallization trials

The most direct method to determine the mechanism of SmgGDS catalyzed nucleotide exchange is to solve the crystal structure of SmgGDS bound to RhoA. The structure of SmgGDS alone is also unknown and its solution would provide valuable information on the conformation of SmgGDS residues that are necessary for nucleotide exchange. A large number of SmgGDS constructs were generated from both the SmgGDS and SmgGDS-558 isoforms (**Fig. 22**). Crystallization trials were conducted with isolated SmgGDS constructs and SmgGDS constructs bound to RhoA. Although numerous conditions were searched, only one condition was found that resulted in successful crystal diffraction.

A truncation construct of SmgGDS-558 from residues 1-472 (SmgGDS-472) crystallized from a 4°C sitting drop over a reservoir containing 21% PEG 3350, 300 mM NaCl, and 100 mM Bis-tris propane pH 7.5 (**Fig. 23A**). Crystals were observed to be small rods, rarely longer than 50 μm, that appeared overnight with a mixture of precipitate. Excessive nucleation was observed that could have limited the growth of larger crystals. Macroseeding existing crystals into a condition containing a lower concentration of PEG 3350 was not sufficient to initiate spontaneous nucleation and resulted in the formation of significantly longer and thicker crystals, reaching lengths of 600 μm (**Fig. 23B-C**). Improvement of crystals was attempted with the addition of numerous additives as well as alterations in precipitant, protein, buffer, pH, temperature,

and other conditions. However, no changes were observed that significantly improved the appearance of crystals.

Crystals diffracted, typically in the 7-9 Å resolution range, with one crystal reaching 4.6 Å diffraction at a synchrotron. Although diffraction was not of sufficient resolution to solve the SmgGDS structure, the data was indexed to space group $p4_122$ with unit cell dimensions: $a = b = 82.02$ Å, $c = 340.07$ Å, and $\alpha = \beta = \gamma = 90°$. The long axis along c likely corresponds with the extended structure of ARM repeats. The diffraction of approximately 100 crystals were screened using a variety of crystallization and cryo-protection conditions, but increased resolution was not be obtained. The lack of high resolution diffraction led us to expand our search for an improved crystallization condition.

Knowing that SmgGDS-472 was susceptible to crystal formation, derivatives were made to increase the likelihood of well-diffracting crystals. Highly entropic surface residues, such as lysine, can inhibit crystal formation. Reductive methylation alters surface exposed lysines to reduce their mobility and has been shown to assist crystallization in some instances (205). Alternatively, clusters of entropic surface residues can be predicted by computational techniques, such as the SERp server (206), and directly mutated to alanines. Both reductive methylation and site-directed mutations (predicted by the SERp server) were conducted on SmgGDS-472. The modified proteins were then broadly screened for new crystallization conditions. Unfortunately, conditions resulting in reproducible crystal formation were not identified.

Analysis of RhoA mutant effects on Dbs and SmgGDS catalyzed nucleotide exchange

Mutational analyses of SmgGDS were successful in the identification of a RhoA binding surface. Therefore, using the same approach we sought to find mutations in RhoA that would affect the ability of SmgGDS to cause exchange, and thus identify the surface of RhoA that interacts with SmgGDS. A sequence alignment of small GTPases was used to select residues for mutation. Because SmgGDS activates RhoA, but neither Rac nor Cdc42, residues unique to RhoA were initially targeted. Mutations were first isolated to regions known to be important for binding to GEFs, such as the switch regions, but later extended to include other areas of the GTPase. Mutant proteins were examined for their ability to be exchanged upon by Dbs and SmgGDS (Table 1). Multiple residues known to be critical for Dbs activation of RhoA, such as A61, E64, and D67, were necessary for exchange by Dbs, but not SmgGDS. Surprisingly, only mutations in the polybasic region of RhoA were seen to inhibit the exchange ability of SmgGDS. None of the residues in the area traditionally associated with GEF interactions, the switch regions, were essential, nor were any of the other residues tested. Although we cannot exclude the involvement of the switch region in the absence of a crystal structure, it appears that an alternative interface of RhoA is utilized to contact SmgGDS. Further mutational analyses or the solution of a RhoA/SmgGDS structure will be required to understand how RhoA interacts with SmgGDS.

Discussion

While previous reports indicate SmgGDS activates a wide variety of small GTPases, including members of the Ras and Rho families, this broad specificity would be unique compared to all other known GEFs. In addition, the existing literature on the

ability of SmgGDS to activate many small GTPases contains conflicting results. Reports from the same group claim SmgGDS stimulates exchange on Rap1b (152) or RhoB (146) in one paper yet demonstrate no exchange in other papers (147,151). Early work with SmgGDS was performed with SmgGDS purified from brain extracts and contamination by other GEFs could account for the conflicting results. In addition, it has been shown that SmgGDS can associate with other exchange factors such as β-Pix lending credence to this possibility (176). Using highly purified recombinant proteins, we have demonstrated that SmgGDS is only capable of stimulating nucleotide exchange on RhoA and RhoC *in vitro*. This selective substrate specificity aligns with the paradigm of GEFs activating specific GTPases to cause downstream signaling events as opposed to broadly activating many diverse GTPases.

SmgGDS has also been proposed to have an alternative function in regulating the nucleocytoplasmic transport of small GTPases (167). Our present studies do not examine this role of SmgGDS and it is possible that the protein may have multiple biological roles. However, we show convincing evidence that SmgGDS shares the characteristics expected of a true guanine nucleotide exchange factor. First, it is able to catalyze both the loading and the unloading of MANT-GDP on RhoA in a catalytic and concentration-dependent fashion. In addition, it specifically forms a high affinity complex only with nucleotide-depleted RhoA and not with GDP- or GTPγS-bound RhoA, displaying a key characteristic of GEFs to stabilize the nucleotide-free state.

The mechanism of nucleotide exchange of Rho family GTPases was originally thought to occur via a conserved nucleotide-free state, and to include divergent GEFs such as SopE which despite having a unique tertiary structure from Dbl family RhoGEFs

stabilizes a similar nucleotide-free conformation (207). When Rho GTPases bind to Dbl family GEFs or SopE, residues A61 and E64 stabilize a conformation of switch 2 that is incompatible with efficient nucleotide binding and mutation of either residue abolishes the ability of these GEFs to catalyze nucleotide exchange. Our data show that these residues are not essential for the ability of SmgGDS to exchange upon RhoA, indicating SmgGDS has a unique exchange mechanism that stabilizes a distinct nucleotide-free state of the GTPase. DOCK9 has also been shown to stabilize a nucleotide-free conformation of Cdc42 which does not rely on A61 or E64 (79). DOCK9 catalyzes nucleotide exchange by the direct insertion of a valine side chain into the nucleotide binding pocket. Although no homology exists between the DOCK9 GEF region and SmgGDS, we cannot rule out that SmgGDS may work in a similar fashion. However, it is likely that SmgGDS stabilizes another distinct nucleotide-free conformation and that there are in fact multiple nucleotide-free states for Rho GTPases that can be stabilized by GEFs.

The definitive method to address the mechanism of nucleotide exchange catalyzed by SmgGDS is solution of its crystal structure bound to nucleotide-depleted GTPase. Despite extensive efforts with multiple SmgGDS constructs, crystals of RhoA bound to SmgGDS have not been obtained. Crystals of an isolated SmgGDS truncation were obtained, but did not diffract to sufficient resolution for structural visualization. Screening additional crystallization conditions may yield better crystals, but alternative approaches may also be productive. Very little is known about biological partners of SmgGDS, but if additional binding partners were identified they could be incorporated into crystallization trials where they might assist crystal formation.

We have for the first time identified regions of SmgGDS that are essential for its ability to catalyze exchange upon RhoA. The necessity of a polybasic region on RhoA for exchange by SmgGDS implied that there would be acidic residues on the surface of SmgGDS to mediate this interaction. The electrostatic surface potential generated from the SmgGDS homology model illustrated an electronegatively charged region essential for the ability of SmgGDS to interact with RhoA. Correspondingly, mutations in the electronegatively charged region ablated the ability of SmgGDS to exchange upon RhoA highlighting its importance. A highly conserved binding surface was also identified in the groove of the ARM-repeat superhelical structure. Mutation of residues in the conserved binding surface also rendered SmgGDS unable to catalyze nucleotide exchange on RhoA. Given the relative positions of the two regions, we postulate that the body of RhoA lies in the conserved hydrophobic groove and that the polybasic tail of RhoA extends over the surface of the electronegative region. Without any structural information on the orientation of RhoA on the surface of SmgGDS it was not possible to model complex formation, but given the distance between the N-terminal edges of the conserved binding groove and the electronegative region (~25 Å) and the length of the polybasic region (~32 Å if fully extended) such a conformation is physically reasonable.

We attempted to find sites on RhoA that interact with SmgGDS, but were unable to identify areas outside of the polybasic region necessary for nucleotide exchange catalyzed by SmgGDS. Our initial analysis focused on residues in switch 1 and switch 2 as these have been implicated as the major sites of interaction with all known Rho GEFs (47,79). However, nearly the entire binding surface of a GTPase has been shown to be utilized in binding to other proteins (208). Thus, it is entirely possible that the switch

regions are not major determinants of the SmgGDS binding interface. Additional mutagenesis over the entire surface of RhoA would likely lead to identification of residues involved in binding SmgGDS. Other approaches such as hydrogen/deuterium exchange on isolated versus SmgGDS-bound RhoA followed by mass spectrometry could implicate potential sites of interaction. Knowing the face of RhoA that interacts with SmgGDS may provide information to explain the RhoA/C specificity of SmgGDS and may also give constraints to guide docking of RhoA onto the SmgGDS homology model.

The observation that only SmgGDS-558 was able to activate RhoA *in vivo* was surprising given that no difference in GEF activity could be discerned between the isoforms *in vitro*. It has recently been discovered that SmgGDS-558 may only interact with prenylated GTPases *in vivo* while SmgGDS may prefer nonprenylated GTPases *in vivo* (T. Berg and C. Williams, personal communication). The reasons for these preferences are not yet fully understood, but may involve the binding of accessory proteins to SmgGDS. Interestingly, the highly conserved residues of R112, N116, and Y119 are directly N-terminal to the extra ARM repeat of SmgGDS. An intriguing speculation is that they may serve as a binding site for a protein that regulates the prenylation-dependent binding of GTPases to the different SmgGDS isoforms.

A unique property of SmgGDS is that it requires the RhoA polybasic region to interact with and promote nucleotide exchange upon RhoA. To date, no other nucleotide exchange factor clearly requires the polybasic region of the GTPase for exchange. The traditional paradigm suggests that the polybasic region functions for proper membrane localization of the GTPase and that is not generally used for binding to protein partners. The polybasic region of the GTPase ends at the CAAX box, with the cysteine residue

being the site of prenylation for the GTPase. Early reports on the function of SmgGDS suggested that SmgGDS required prenylation of the GTPase to stimulate exchange and was capable of causing extraction of GTPase from cell membranes (146,163). Given the close proximity of the prenylation site and the polybasic region, it would be interesting to investigate if prenylation of GTPase alters affinity for either SmgGDS isoform.

This study is the first identification of SmgGDS residues essential for nucleotide exchange. SmgGDS is observed to specifically activate only RhoA and RhoC *in vitro* and this specificity is supported *in vivo*. The mechanism of nucleotide exchange is shown to be distinct from Dbl family GEFs and to require the presence of a C-terminal polybasic region in the GTPase. Future structural characterization of a SmgGDS/RhoA complex will be necessary for a full understanding of the SmgGDS nucleotide exchange mechanism.

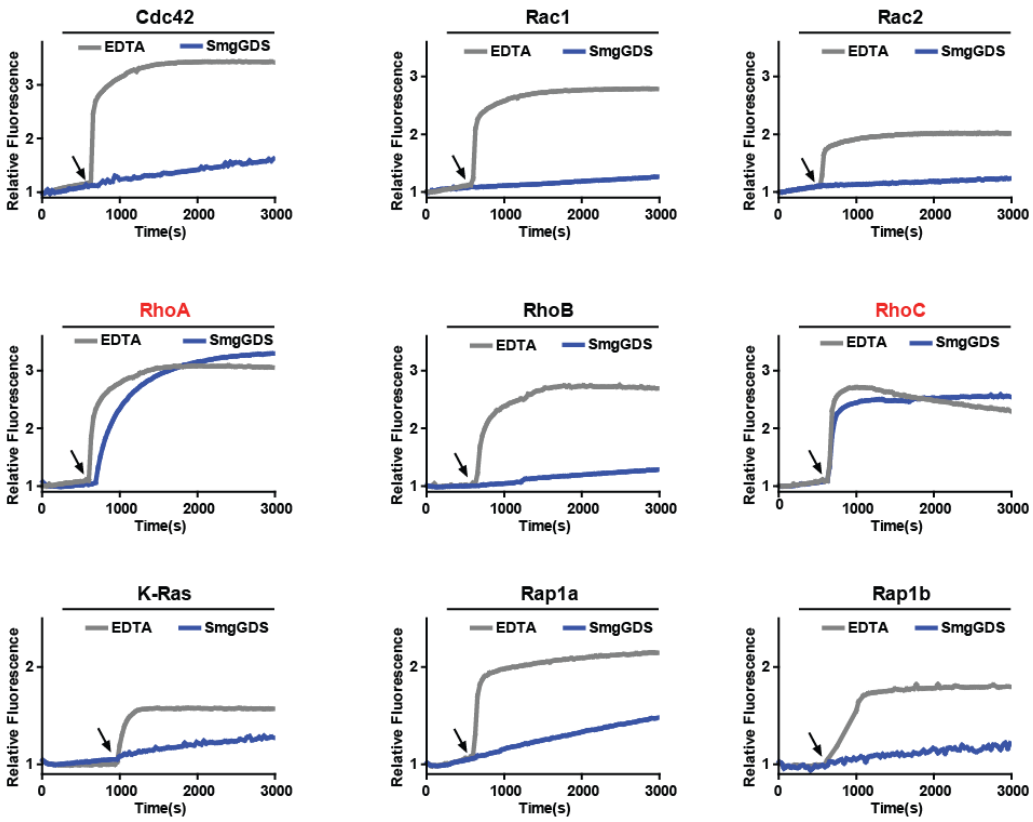

Figure 14: Purified SmgGDS specifically activates RhoA and RhoC *in vitro*.
The intrinsic exchange of the indicated GTPase (1 μM) was measured for 600 s in exchange buffer. At the indicated time (arrow), SmgGDS (50 μM) or EDTA (25 μM) was added to stimulate nucleotide exchange.

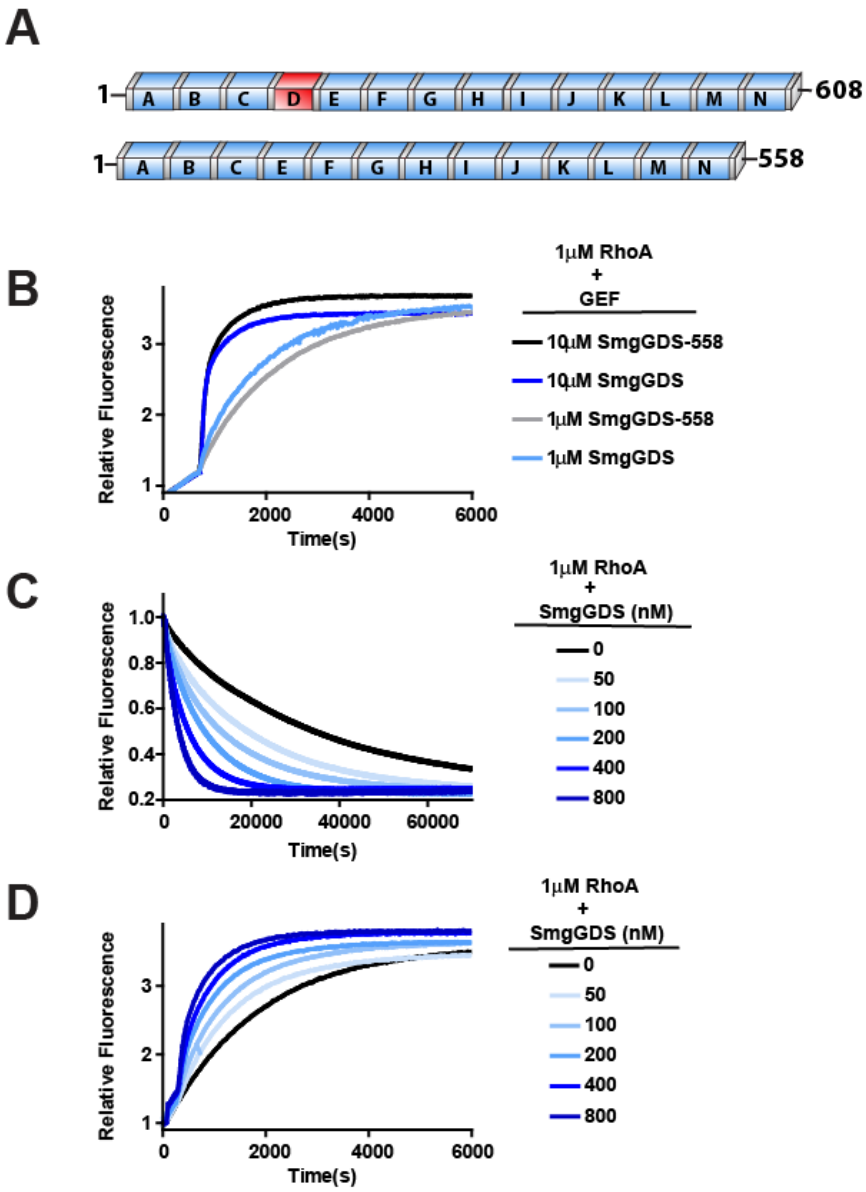

Figure 15: Activation of RhoA by SmgGDS is catalytic and independent of SmgGDS isoform.

A, Domain architecture of SmgGDS isoforms with conserved ARM repeats labeled in blue boxes and inserted repeat in red. *B*, The intrinsic exchange of RhoA (1 μM) was measured for 600 s in exchange buffer before stimulation of nucleotide exchange with the indicated concentration of SmgGDS or SmgGDS-558. *C*, MANT-GDP loaded RhoA (1 μM) was incubated in unloading exchange buffer for 600 s before addition of SmgGDS at the indicated concentrations to stimulate nucleotide exchange. *D*, RhoA (1 μM) was incubated in unloading exchange buffer with 400 nM MANT-GDP in place of the free GDP for 600 s before addition of SmgGDS at the indicated concentrations to stimulate nucleotide exchange.

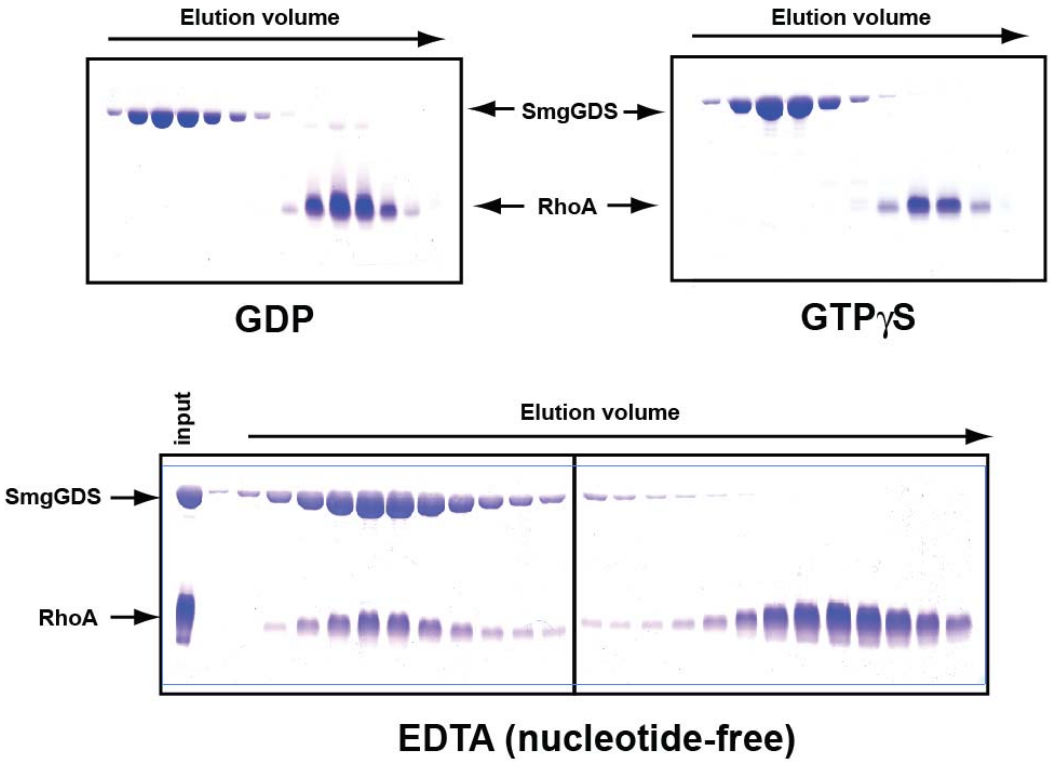

Figure 16: SmgGDS preferentially forms a high-affinity complex with RhoA in the nucleotide-free state.

SmgGDS was incubated on ice for 30 min with a two-fold molar excess of RhoA loaded with GDP (top left), GTPγS (top right), or in the presence of EDTA (bottom), separated over size-exclusion chromatography, and fractions were analyzed via SDS-PAGE.

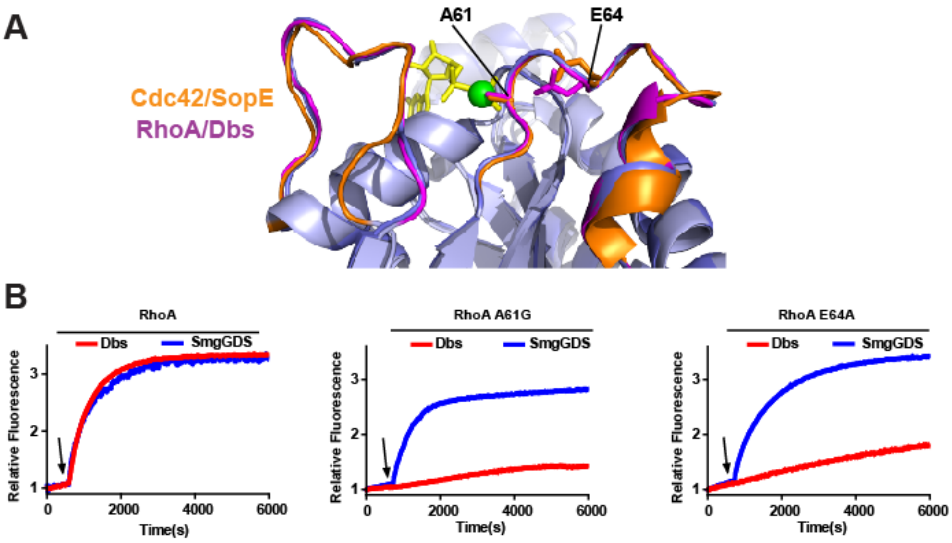

Figure 17: SmgGDS utilizes a distinct exchange mechanism to activate RhoA compared to traditional RhoGEFs.

A, Cartoon representation of nucleotide-free RhoA (slate with switch regions in magenta) from Dbs/RhoA complex (PDB 1LB1) superimposed upon nucleotide-free Cdc42 (slate with switch regions in orange) from the SopE/Cdc42 complex (PDB 1GZS) with Mg^{2+} ion (green ball) and GDP (yellow sticks) from the RhoA-GDP structure (PDB 1FTN) and residues numbered according to RhoA sequence. *B,* The intrinsic exchange of 1μM RhoA (left), RhoA A61G (center), and RhoA E64A (right) was measured for 600 s in exchange buffer. At the indicated time (arrow), SmgGDS (50 μM) or Dbs (200 nM) was added to stimulate nucleotide exchange.

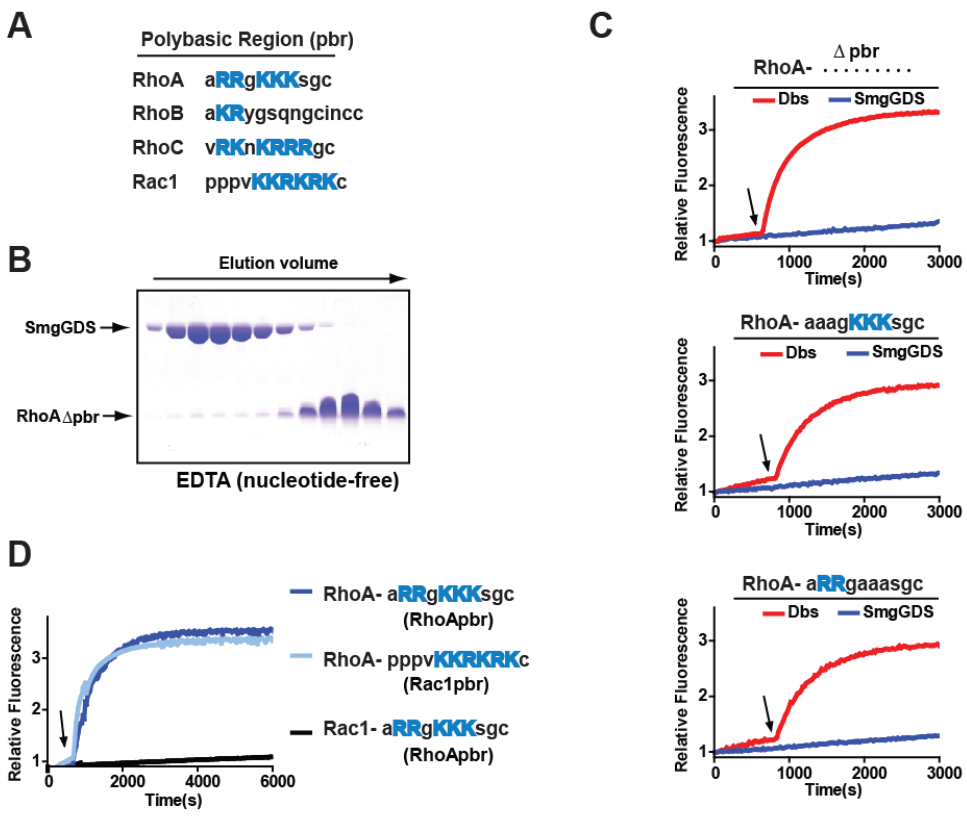

Figure 18: C-terminal polybasic region of RhoA is required for activation by SmgGDS.

A, Sequence comparison of Rho family member polybasic regions with basic residues highlighted in blue. *B*, SmgGDS was incubated on ice for 30 min with a two-fold molar excess of RhoA lacking the polybasic region (RhoA Δpbr) in the presence of EDTA, separated over size-exclusion chromatography, and fractions were analyzed via SDS-PAGE. *C*, The intrinsic exchange of 1μM RhoA Δpbr (top left), RhoA R182A+R183A (top right), and RhoA K185A+K186A+K187A (bottom left) was measured for 600 s in exchange buffer. At the indicated time (arrow), SmgGDS (50 μM) or Dbs (200 nM) was added to stimulate nucleotide exchange. *D*, The intrinsic exchange of the indicated GTPase construct (1 μM) was measured for 600 s in exchange buffer. At the indicated time (arrow), SmgGDS (50 μM) was added to stimulate nucleotide exchange.

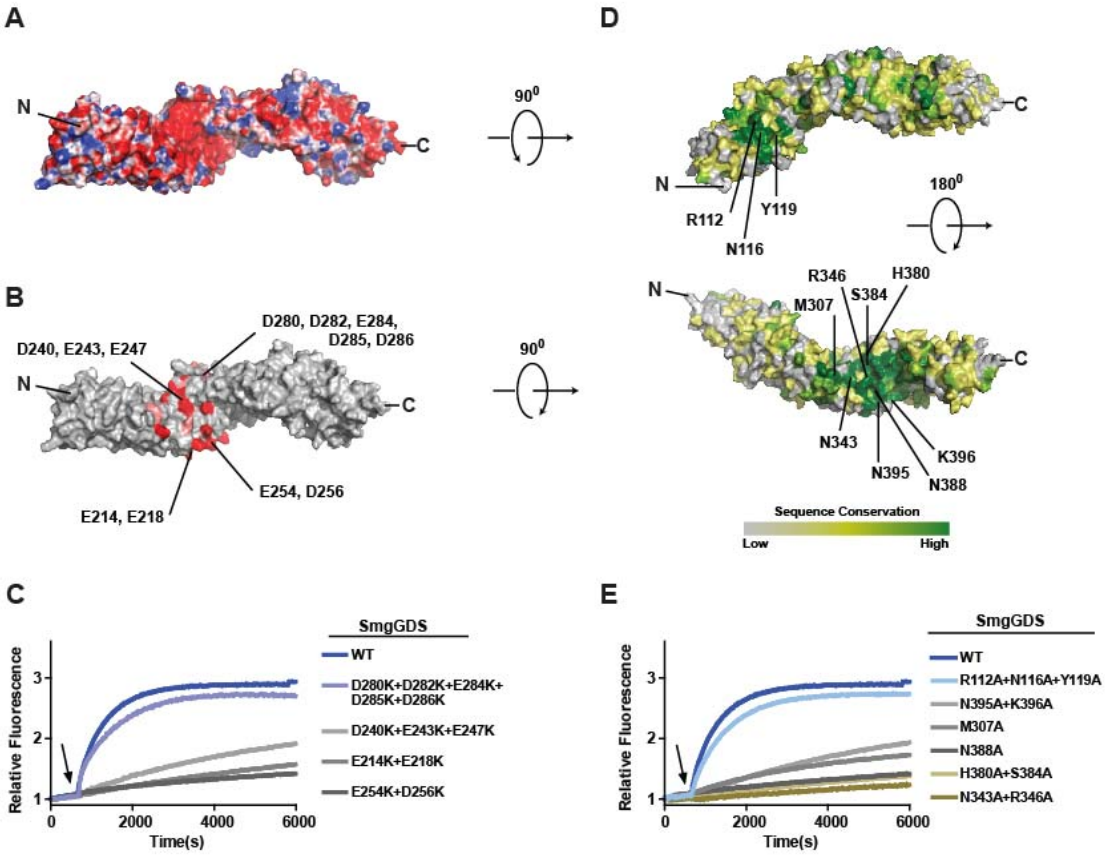

Figure 19: An electronegative patch and highly conserved binding groove on SmgGDS facilitate interaction with RhoA.

A, Electrostatic surface potential of SmgGDS homology model with acidic regions in red and basic regions in blue. *B,* Homology model of SmgGDS with completely (red) and highly (pink) conserved acidic residues. *C,* The intrinsic exchange of RhoA (1 μM) was measured for 600 s in exchange buffer. At the indicated time (arrow), wild-type or mutant SmgGDS (50 μM) was added to stimulate nucleotide exchange. *D,* Homology model of SmgGDS colored according to ClustalX consensus score from a multiple sequence alignment of full-length SmgGDS isoforms. *E,* The intrinsic exchange of RhoA (1 μM) was measured for 600 s in exchange buffer. At the indicated time (arrow), wild-type or mutant SmgGDS (50 μM) was added to stimulate nucleotide exchange.

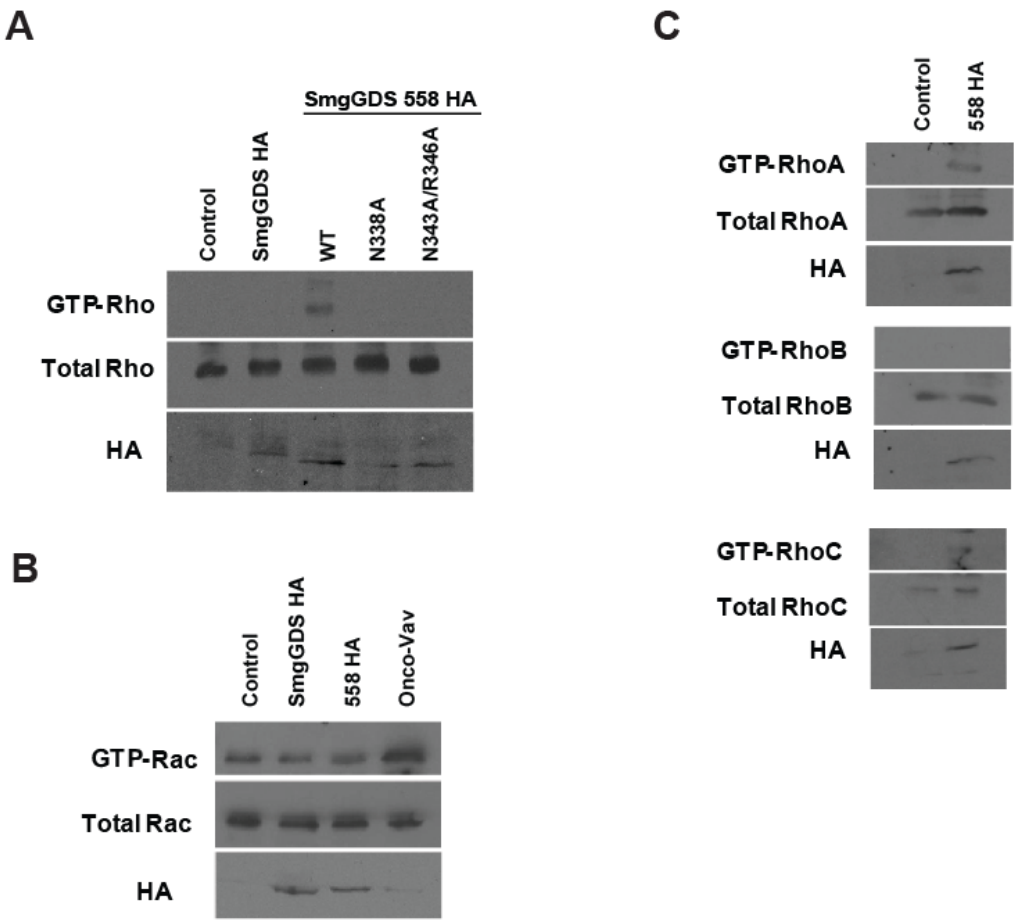

Figure 20: SmgGDS specifically activates RhoA *in vivo*.

HEK293 cells were transfected with the indicated GEF expression vectors then lysed. *A*, Active RhoA was precipitated from lysate with GST-RBD. Pull-down (GTP Rho) and lysates (total Rho) were imunoblotted with a pan-Rho antibody. Expression of GEF was confirmed with immunoblot for HA antibody. *B*, Active Rac was precipitated from lysate with GST-RBD. Pull-down (GTP Rac) and lysates (total Rac) were imunoblotted with a pan-Rac antibody. *C*, Active Rho proteins were precipitated from lysate with GST-RBD. Pull-down (GTP Rho) and lysates (total Rho) were imunoblotted with antibodies specific for RhoA, B, or C.

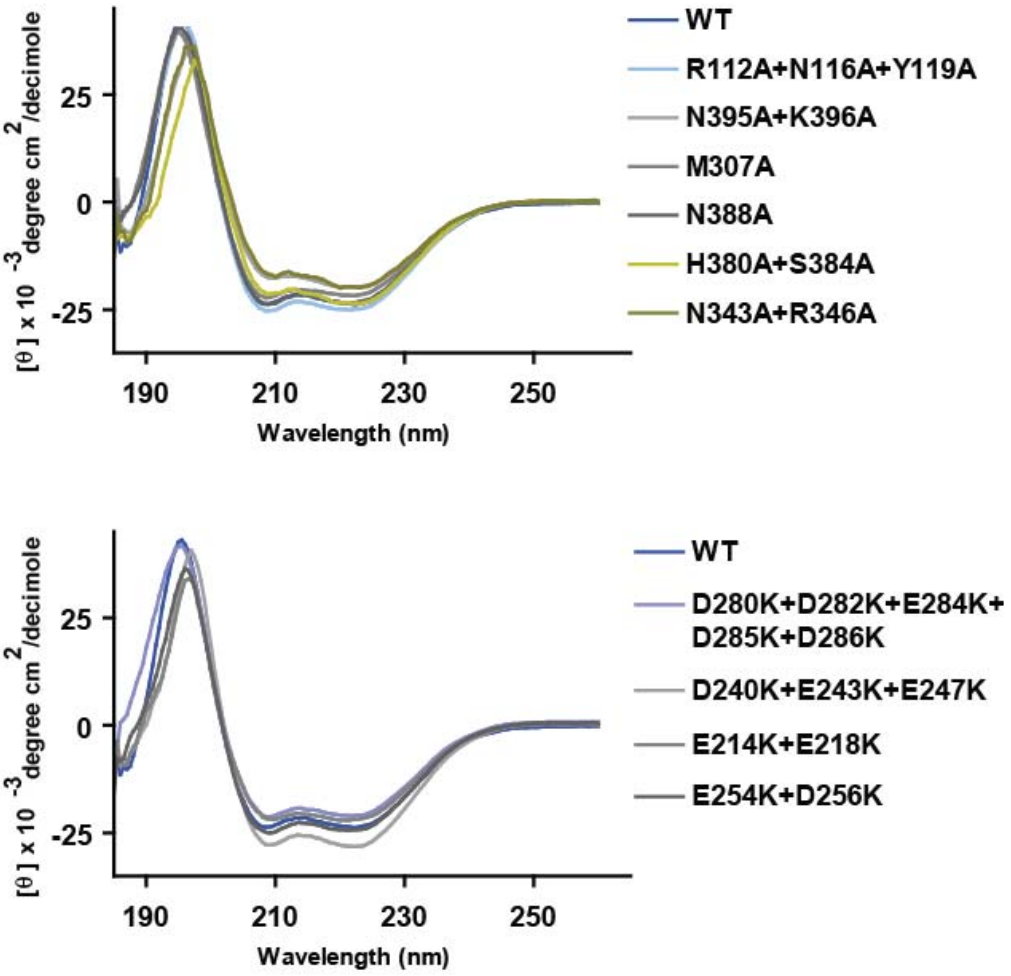

Figure 21: Circular dichroism confirms mutant SmgGDS proteins possess the same fold as wild-type SmgGDS.

The mean residue ellipticity as a function of wavelength for purified SmgGDS proteins at a concentration of 0.15 mg/mL in 20 mM $NaPO_4$ analyzed using a Pistar-180 spectrometer with a 0.5 mM pathlength at 25°C. Mutations in either the conserved GTPase binding groove (top) or the electropositive patch (bottom) do not affect the α-helical fold of SmgGDS.

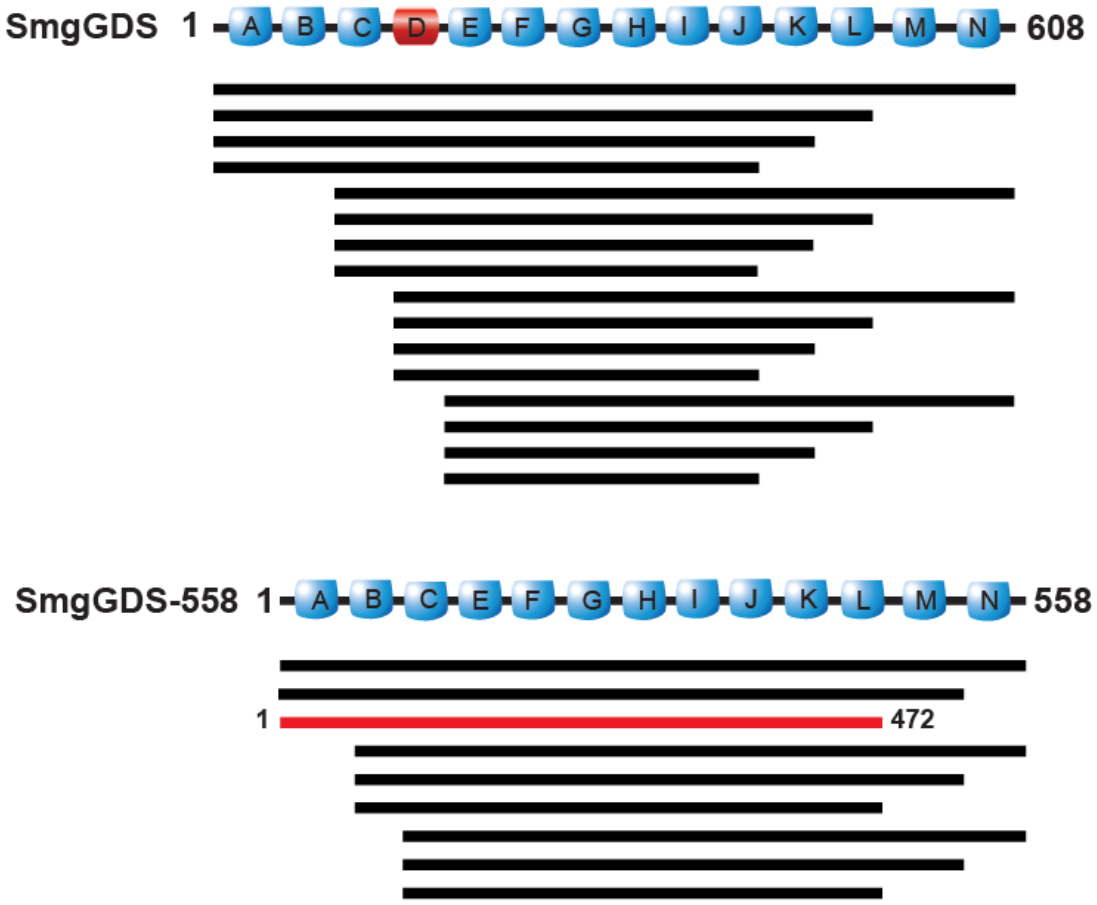

Figure 22: SmgGDS constructs used for crystallization trials.
The constructs of SmgGDS used for crystallization trials (black lines) with construct resulting in crystal formation (red).

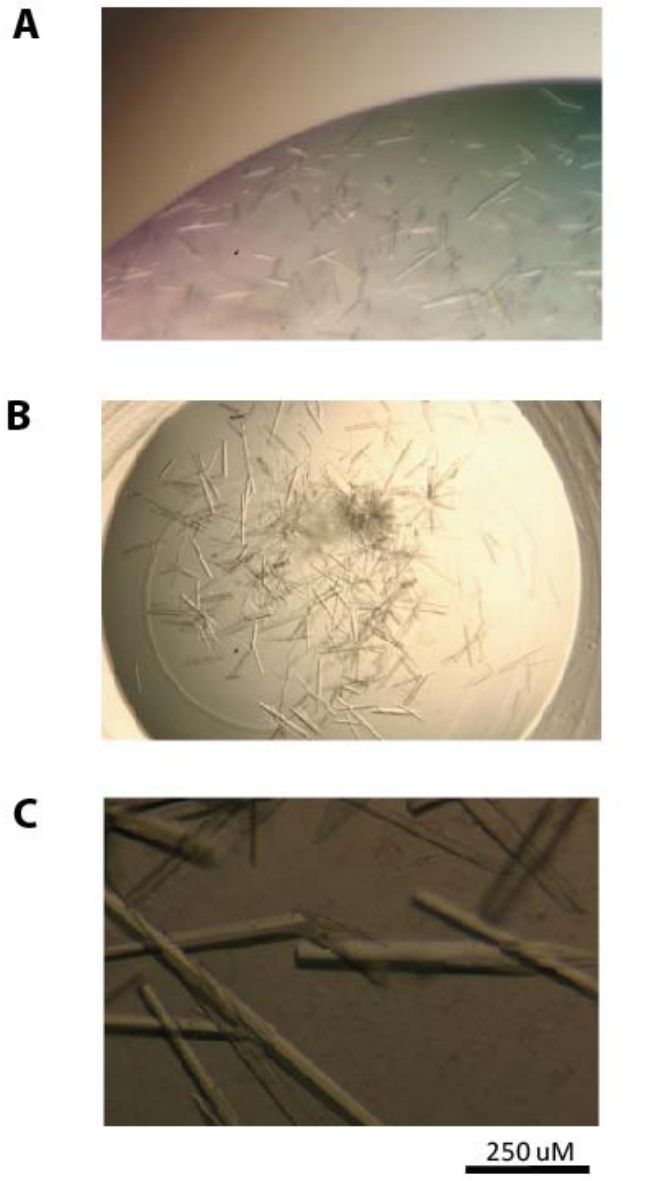

Figure 23: SmgGDS-472 crystal size is improved through macroseeding.
A, Initial crystals of SmgGDS-472 grew from sitting drops over a reservoir containing 21% PEG 3350, 300 mM NaCl, and 100 mM Bis-tris propane pH 7.4 and reached a maximum length of 50 µM. *B*, Macroseeding of crystals into 18% PEG 3350, 300 mM NaCl, and 100 mM Bis-tris propane pH 7.5 *C*, Close-up view of crystals from B.

Residue	Loss of Dbs catalyzed exchange	Loss of SmgGDS catalyzed exchange	No loss of exchange
R5A			X
C16S			X
C16V			X
C16A			X
S26A, K27A, D28A			X
S26			X
E32G			X
V33S			X
E32G, V33S			X
E32G, V33E			X
E32A, V33A			X
E40A			X
E40N			X
D45A			X
D45N			X
D45N, E47N			X
D45A, E47A, D49A			X
D45N, E47N, D49N			X
E47N			X
E47N, D49N			X
A61G	X		
E64A	X		
D67A	X		
D76A			X
D76S			X
D76Q			X
D76A, D78A			X
D76Q, I80F			X
I80F			X
K104A, P108A, N109A			X
F154A			X
R168A, E169A, E172A			X
R176A			X
R182A, R183A		X	
K185A, K186A, K187A		X	

Table 1: Multiple mutations in RhoA have no effect on SmgGDS catalyzed exchange.

The abilities of SmgGDS and Dbs to catalyze nucleotide exchange on mutant RhoA proteins were tested using a MANT-GDP loading assay. Mutants were classified as having loss of Dbs catalyzed exchange, loss of SmgGDS catalyzed exchange, or no loss of exchange by either Dbs or SmgGDS.

CHAPTER 4: CONCLUSIONS AND FUTURE DIRECTIONS

Conclusions

Rho family GTPases play vital roles in cellular processes. Their activation is tightly controlled by GEFs, which stabilize nucleotide-free GTPases. While the mechanism of traditional Dbl family GEFs is known, many questions surround the ability of DOCK family proteins and SmgGDS to catalyze exchange. Our results suggest that DOCK9 may utilize regions outside of the identified GEF domain to increase its ability to cause exchange on Cdc42. SmgGDS is shown to catalyze nucleotide exchange specifically on RhoA and RhoC using a mechanism distinct from Dbl family GEFs that depends on the presence of the Rho polybasic region. Importantly, both DOCK9 and SmgGDS appear to catalyze exchange through unique mechanisms that suggest the existence of multiple nucleotide-free GTPase conformations.

The domains of DOCK9 were identified through secondary and tertiary structure prediction techniques coupled with analysis of multiple sequence alignments. Numerous truncation constructs were generated and tested for solubility in bacteria. Although it was not successfully crystallized, a C-terminal truncation, DOCK9 CT, was shown both to bind nucleotide-depleted Cdc42 and to catalyze exchange on Cdc42. Full-length DOCK9 was expressed in insect cells, but could not be separated from proteolytic break-down products. However, DOCK9 FL was significantly more active than DOCK9 CT in

catalyzing exchange on Cdc42 implicating domains other than the GEF domain as contributing to nucleotide exchange. Regions outside of the GEF domain have also been implicated as contributing to exchange for DOCK11, where binding of active Cdc42 to its N-terminus led to increased GEF activity of the full-length protein compared to the isolated DHR-2 (143).

SmgGDS has been reported to activate multiple GTPases of both the Rho and Ras families. However, early reports utilized impure protein and contained conflicting results. We show SmgGDS specifically activates RhoA and RhoC both *in vitro* and *in vivo*. SmgGDS activation of RhoA is shown to be catalytic and concentration-dependent. The ability of SmgGDS to interact with RhoA depends on the presence of an intact polybasic region on the GTPase, a unique requirement not known to be essential for any other GTPase/GEF interaction. The sites of SmgGDS interaction with RhoA are mapped to a highly conserved binding groove and an acidic surface patch, both of which are necessary for the capacity of SmgGDS to cause exchange. This is the first study to examine regions of SmgGDS necessary for its nucleotide exchange ability and the resulting mutants will be useful in probing *in vivo* SmgGDS function.

While traditional Rho GEFs utilize a conserved mechanism for exchange, it has also been shown the structurally distinct bacterial GEF, SopE, stabilizes a nearly identical nucleotide-free GTPase conformation as Dbl family GEFs. Both Dbs and SopE stabilize conformational rearrangements that lead to occlusion of nucleotide binding by a conserved residue located in switch 2, A61. Therefore, it was hypothesized that a single nucleotide-free state for Rho GTPases might exist. However, the structure of DOCK9 bound to Cdc42 illustrated a distinct nucleotide-free GTPase conformation that did not

utilize A61. Instead, a valine from DOCK9 was directly inserted into the nucleotide binding pocket. We also show that SmgGDS does not require A61 to catalyze nucleotide exchange on RhoA. While the exact exchange mechanism is unknown, SmgGDS has no homology to DOCK proteins suggesting it may stabilize another distinct nucleotide-free GTPase conformation. Thus, it appears that in contrast to previous adopted hypotheses, Rho GEFs can stabilize multiple nucleotide-free GTPase conformations.

Future directions

Despite the structure of a Cdc42/DOCK9 complex, multiple questions regarding the mechanism of exchange for the DOCK family GEFs remain. What are the contributions of domains outside of the core GEF domain to exchange? How do PH and C2 domains orient DOCK9 at the membrane? What are the autoinhibited states of DOCK9 and DOCK1 and what is the contribution of ELMO to the exchange mechanism of DOCKs 1-5? These questions are unlikely to be fully answered in the absence of structures of full-length DOCK proteins both bound to cognate GTPases and in isolation. Thus, a critical future direction is the purification of full-length DOCK proteins that are not degraded by proteases. This may be obtained by screening multiple DOCK isozymes in insect cell and yeast expression systems. For example, a complex of full-length DOCK2 bound to ELMO has been successfully purified from insect cells with negligible proteolytic contamination (Cynthia Holley, unpublished observations).

The exact mechanism of SmgGDS exchange on RhoA is unlikely to be determined in the absence of a crystal structure. The use of newly acquired robotics for

setting up and imaging crystallization trials will allow more conditions to be screened effectively. The primary effort was put towards the SmgGDS-472 construct because it was the only construct observed to form crystals. However, there are many other SmgGDS constructs that could be screened more thoroughly. In addition, now that we know the importance of the polybasic region, polybasic peptides can be used as co-crystallization targets with SmgGDS to aid in understanding the molecular details of the RhoA/SmgGDS interaction.

The regions of RhoA that interact with SmgGDS have not been determined, although many RhoA mutants have been tested. A brute-force approach could be utilized to test additional RhoA mutants. Alternatively, cross-linking of RhoA and SmgGDS followed by protease digests and mass spectrometry could reveal distance constraints for SmgGDS and RhoA residues. Hydrogen-deuterium exchange rates of RhoA residues involved in binding SmgGDS would also be expected to be altered upon complex formation that could be detected through mass spectrometry. SmgGDS has been identified as interacting with other GTPases including Rac and Rap, in cells using pull-down experiments. It may prove insightful to probe if they directly interact with SmgGDS; and if so, is a different interface utilized that accounts for their inability to be exchanged upon?

SmgGDS is a relatively inefficient exchange factor *in vitro* compared to Dbl family GEFs. However, it is possible that unknown accessory proteins *in vivo* increase the capacity of SmgGDS to catalyze exchange. The necessity of a binding partner for maximal activity has been observed for other GEFs such as DOCKs 1-5, which are activated by ELMO. The ability of purified SmgGDS to catalyze exchange could be

compared to SmgGDS immunoprecipitated from cells to test this hypothesis. Additionally, to identify SmgGDS binding partners pull-downs of SmgGDS from cell lysates could be analyzed by mass spectrometry. Identification of additional binding partners will also allow them to be incorporated into crystallization trials, which may increase the possibility of obtaining well-diffracting crystals.

Finally, the identification of SmgGDS mutants lacking the ability to bind RhoA will be useful in probing the *in vivo* function of SmgGDS. Both isoforms of SmgGDS can be knocked-down by siRNA and SmgGDS knock-down has been shown to effect cellular migration and growth. We now have the ability to reintroduce both wild-type SmgGDS and RhoA binding defective mutants into knocked-down cells. This will allow us to address if the biological effects of SmgGDS depend upon RhoA activation or if there are alternative cellular functions of SmgGDS that do not depend upon RhoA binding. These studies have been initiated in collaboration with Carol Williams.

REFERENCES

1. Vetter, I. R., and Wittinghofer, A. The guanine nucleotide-binding switch in three dimensions. (2001) *Science* **294**, 1299-1304

2. Noble, C. G., and Song, H. Structural studies of elongation and release factors. (2008) *Cell Mol Life Sci* **65**, 1335-1346

3. Johnston, C. A., and Siderovski, D. P. Receptor-mediated activation of heterotrimeric G-proteins: current structural insights. (2007) *Mol Pharmacol* **72**, 219-230

4. Wennerberg, K., Rossman, K. L., and Der, C. J. The Ras superfamily at a glance. (2005) *J Cell Sci* **118**, 843-846

5. Cook, A., Bono, F., Jinek, M., and Conti, E. Structural biology of nucleocytoplasmic transport. (2007) *Annu Rev Biochem* **76**, 647-671

6. Weis, K. Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle. (2003) *Cell* **112**, 441-451

7. Stenmark, H. Rab GTPases as coordinators of vesicle traffic. (2009) *Nat Rev Mol Cell Biol* **10**, 513-525

8. Gillingham, A. K., and Munro, S. The small G proteins of the Arf family and their regulators. (2007) *Annu Rev Cell Dev Biol* **23**, 579-611

9. Myers, K. R., and Casanova, J. E. Regulation of actin cytoskeleton dynamics by Arf-family GTPases. (2008) *Trends Cell Biol* **18**, 184-192

10. Karnoub, A. E., and Weinberg, R. A. Ras oncogenes: split personalities. (2008) *Nat Rev Mol Cell Biol* **9**, 517-531

11. Shields, J. M., Pruitt, K., McFall, A., Shaub, A., and Der, C. J. Understanding Ras: 'it ain't over 'til it's over'. (2000) *Trends Cell Biol* **10**, 147-154

12. DeFeo, D., Gonda, M. A., Young, H. A., Chang, E. H., Lowy, D. R., Scolnick, E. M., and Ellis, R. W. Analysis of two divergent rat genomic clones homologous to the transforming gene of Harvey murine sarcoma virus. (1981) *Proc Natl Acad Sci U S A* **78**, 3328-3332

13. Roberts, P. J., and Der, C. J. Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. (2007) *Oncogene* **26**, 3291-3310

14. Caron, E. Cellular functions of the Rap1 GTP-binding protein: a pattern emerges. (2003) *J Cell Sci* **116**, 435-440

15. Stork, P. J. Does Rap1 deserve a bad Rap? (2003) *Trends Biochem Sci* **28**, 267-275

16. Nomura, K., Kanemura, H., Satoh, T., and Kataoka, T. Identification of a novel domain of Ras and Rap1 that directs their differential subcellular localizations. (2004) *J Biol Chem* **279**, 22664-22673

17. Pannekoek, W. J., Kooistra, M. R., Zwartkruis, F. J., and Bos, J. L. Cell-cell junction formation: the role of Rap1 and Rap1 guanine nucleotide exchange factors. (2009) *Biochim Biophys Acta* **1788**, 790-796

18. Lafuente, E. M., Iwamoto, Y., Carman, C. V., van Puijenbroek, A. A., Constantine, E., Li, L., and Boussiotis, V. A. Active Rap1, a small GTPase that induces malignant transformation of hematopoietic progenitors, localizes in the nucleus and regulates protein expression. (2007) *Leuk Lymphoma* **48**, 987-1002

19. Wennerberg, K., and Der, C. J. Rho-family GTPases: it's not only Rac and Rho (and I like it). (2004) *J Cell Sci* **117**, 1301-1312

20. Etienne-Manneville, S., and Hall, A. Rho GTPases in cell biology. (2002) *Nature* **420**, 629-635

21. Coleman, M. L., Marshall, C. J., and Olson, M. F. RAS and RHO GTPases in G1-phase cell-cycle regulation. (2004) *Nat Rev Mol Cell Biol* **5**, 355-366

22. Jaffe, A. B., and Hall, A. Rho GTPases: biochemistry and biology. (2005) *Annu Rev Cell Dev Biol* **21**, 247-269

23. Wu, H., Rossi, G., and Brennwald, P. The ghost in the machine: small GTPases as spatial regulators of exocytosis. (2008) *Trends Cell Biol* **18**, 397-404

24. Nikolic, M. The role of Rho GTPases and associated kinases in regulating neurite outgrowth. (2002) *Int J Biochem Cell Biol* **34**, 731-745

25. Olenik, C., Aktories, K., and Meyer, D. K. Differential expression of the small GTP-binding proteins RhoA, RhoB, Cdc42u and Cdc42b in developing rat neocortex. (1999) *Brain Res Mol Brain Res* **70**, 9-17

26. Erickson, J. W., and Cerione, R. A. Multiple roles for Cdc42 in cell regulation. (2001) *Curr Opin Cell Biol* **13**, 153-157

27. Heasman, S. J., and Ridley, A. J. Mammalian Rho GTPases: new insights into their functions from in vivo studies. (2008) *Nat Rev Mol Cell Biol* **9**, 690-701

28. Arias-Romero, L. E., and Chernoff, J. A tale of two Paks. (2008) *Biol Cell* **100**, 97-108

29. Wang, D. A., and Sebti, S. M. Palmitoylated cysteine 192 is required for RhoB tumor-suppressive and apoptotic activities. (2005) *J Biol Chem* **280**, 19243-19249

30. Wheeler, A. P., and Ridley, A. J. Why three Rho proteins? RhoA, RhoB, RhoC, and cell motility. (2004) *Exp Cell Res* **301**, 43-49

31. Narumiya, S., Tanji, M., and Ishizaki, T. Rho signaling, ROCK and mDia1, in transformation, metastasis and invasion. (2009) *Cancer Metastasis Rev* **28**, 65-76

32. Kim, D. M., Chung, K. S., Choi, S. J., Jung, Y. J., Park, S. K., Han, G. H., Ha, J. S., Song, K. B., Choi, N. S., Kim, H. M., Won, M., and Seo, Y. S. RhoB induces apoptosis via direct interaction with TNFAIP1 in HeLa cells. (2009) *Int J Cancer* **125**, 2520-2527

33. Lebowitz, P. F., and Prendergast, G. C. Functional interaction between RhoB and the transcription factor DB1. (1998) *Cell Adhes Commun* **6**, 277-287

34. Arthur, W. T., Ellerbroek, S. M., Der, C. J., Burridge, K., and Wennerberg, K. XPLN, a guanine nucleotide exchange factor for RhoA and RhoB, but not RhoC. (2002) *J Biol Chem* **277**, 42964-42972

35. Kitzing, T. M., Wang, Y., Pertz, O., Copeland, J. W., and Grosse, R. Formin-like 2 drives amoeboid invasive cell motility downstream of RhoC. (2010) *Oncogene*

36. Karlsson, R., Pedersen, E. D., Wang, Z., and Brakebusch, C. Rho GTPase function in tumorigenesis. (2009) *Biochim Biophys Acta* **1796**, 91-98

37. Horiuchi, A., Imai, T., Wang, C., Ohira, S., Feng, Y., Nikaido, T., and Konishi, I. Up-regulation of small GTPases, RhoA and RhoC, is associated with tumor progression in ovarian carcinoma. (2003) *Lab Invest* **83**, 861-870

38. Pille, J. Y., Denoyelle, C., Varet, J., Bertrand, J. R., Soria, J., Opolon, P., Lu, H., Pritchard, L. L., Vannier, J. P., Malvy, C., Soria, C., and Li, H. Anti-RhoA and anti-RhoC siRNAs inhibit the proliferation and invasiveness of MDA-MB-231 breast cancer cells in vitro and in vivo. (2005) *Mol Ther* **11**, 267-274

39. Faried, A., Faried, L. S., Usman, N., Kato, H., and Kuwano, H. Clinical and prognostic significance of RhoA and RhoC gene expression in esophageal squamous cell carcinoma. (2007) *Ann Surg Oncol* **14**, 3593-3601

40. Wang, H. B., Liu, X. P., Liang, J., Yang, K., Sui, A. H., and Liu, Y. J. Expression of RhoA and RhoC in colorectal carcinoma and its relations with clinicopathological parameters. (2009) *Clin Chem Lab Med* **47**, 811-817

41. Takami, Y., Higashi, M., Kumagai, S., Kuo, P. C., Kawana, H., Koda, K., Miyazaki, M., and Harigaya, K. The activity of RhoA is correlated with lymph node metastasis in human colorectal cancer. (2008) *Dig Dis Sci* **53**, 467-473

42. Simpson, K. J., Dugan, A. S., and Mercurio, A. M. Functional analysis of the contribution of RhoA and RhoC GTPases to invasive breast carcinoma. (2004) *Cancer Res* **64**, 8694-8701

43. Bellovin, D. I., Simpson, K. J., Danilov, T., Maynard, E., Rimm, D. L., Oettgen, P., and Mercurio, A. M. Reciprocal regulation of RhoA and RhoC characterizes the EMT and identifies RhoC as a prognostic marker of colon carcinoma. (2006) *Oncogene* **25**, 6959-6967

44. Ikoma, T., Takahashi, T., Nagano, S., Li, Y. M., Ohno, Y., Ando, K., Fujiwara, T., Fujiwara, H., and Kosai, K. A definitive role of RhoC in metastasis of orthotopic lung cancer in mice. (2004) *Clin Cancer Res* **10**, 1192-1200

45. Faried, A., Faried, L. S., Kimura, H., Nakajima, M., Sohda, M., Miyazaki, T., Kato, H., Usman, N., and Kuwano, H. RhoA and RhoC proteins promote both cell proliferation and cell invasion of human oesophageal squamous cell carcinoma cell lines in vitro and in vivo. (2006) *Eur J Cancer* **42**, 1455-1465

46. Zhang, S., Tang, Q., Xu, F., Xue, Y., Zhen, Z., Deng, Y., Liu, M., Chen, J., Liu, S., Qiu, M., Liao, Z., Li, Z., Luo, D., Shi, F., Zheng, Y., and Bi, F. RhoA regulates G1-S progression of gastric cancer cells by modulation of multiple INK4 family tumor suppressors. (2009) *Mol Cancer Res* **7**, 570-580

47. Dvorsky, R., and Ahmadian, M. R. Always look on the bright site of Rho: structural implications for a conserved intermolecular interface. (2004) *EMBO Rep* **5**, 1130-1136

48. Bos, J. L., Rehmann, H., and Wittinghofer, A. GEFs and GAPs: critical elements in the control of small G proteins. (2007) *Cell* **129**, 865-877

49. DerMardirossian, C., and Bokoch, G. M. GDIs: central regulatory molecules in Rho GTPase activation. (2005) *Trends Cell Biol* **15**, 356-363

50. Goody, R. S., and Hofmann-Goody, W. Exchange factors, effectors, GAPs and motor proteins: common thermodynamic and kinetic principles for different functions. (2002) *Eur Biophys J* **31**, 268-274

51. Boriack-Sjodin, P. A., Margarit, S. M., Bar-Sagi, D., and Kuriyan, J. The structural basis of the activation of Ras by Sos. (1998) *Nature* **394**, 337-343

52. Freedman, T. S., Sondermann, H., Friedland, G. D., Kortemme, T., Bar-Sagi, D., Marqusee, S., and Kuriyan, J. A Ras-induced conformational switch in the Ras activator Son of sevenless. (2006) *Proc Natl Acad Sci U S A* **103**, 16692-16697

53. Goldberg, J. Structural basis for activation of ARF GTPase: mechanisms of guanine nucleotide exchange and GTP-myristoyl switching. (1998) *Cell* **95**, 237-248

54. Renault, L., Kuhlmann, J., Henkel, A., and Wittinghofer, A. Structural basis for guanine nucleotide exchange on Ran by the regulator of chromosome condensation (RCC1). (2001) *Cell* **105**, 245-255

55. Itzen, A., Pylypenko, O., Goody, R. S., Alexandrov, K., and Rak, A. Nucleotide exchange via local protein unfolding--structure of Rab8 in complex with MSS4. (2006) *Embo J* **25**, 1445-1455

56. Delprato, A., and Lambright, D. G. Structural basis for Rab GTPase activation by VPS9 domain exchange factors. (2007) *Nat Struct Mol Biol* **14**, 406-412

57. Dong, G., Medkova, M., Novick, P., and Reinisch, K. M. A catalytic coiled coil: structural insights into the activation of the Rab GTPase Sec4p by Sec2p. (2007) *Mol Cell* **25**, 455-462

58. Rossman, K. L., Der, C. J., and Sondek, J. GEF means go: turning on RHO GTPases with guanine nucleotide-exchange factors. (2005) *Nat Rev Mol Cell Biol* **6**, 167-180

59. Worthylake, D. K., Rossman, K. L., and Sondek, J. Crystal structure of Rac1 in complex with the guanine nucleotide exchange region of Tiam1. (2000) *Nature* **408**, 682-688

60. Rossman, K. L., Worthylake, D. K., Snyder, J. T., Siderovski, D. P., Campbell, S. L., and Sondek, J. A crystallographic view of interactions between Dbs and Cdc42: PH domain-assisted guanine nucleotide exchange. (2002) *Embo J* **21**, 1315-1326

61. Derewenda, U., Oleksy, A., Stevenson, A. S., Korczynska, J., Dauter, Z., Somlyo, A. P., Otlewski, J., Somlyo, A. V., and Derewenda, Z. S. The crystal structure of RhoA in complex with the DH/PH fragment of PDZRhoGEF, an activator of the Ca(2+) sensitization pathway in smooth muscle. (2004) *Structure* **12**, 1955-1965

62. Buchwald, G., Friebel, A., Galan, J. E., Hardt, W. D., Wittinghofer, A., and Scheffzek, K. Structural basis for the reversible activation of a Rho protein by the bacterial toxin SopE. (2002) *Embo J* **21**, 3286-3295

63. Hasegawa, H., Kiyokawa, E., Tanaka, S., Nagashima, K., Gotoh, N., Shibuya, M., Kurata, T., and Matsuda, M. DOCK180, a major CRK-binding protein, alters cell

morphology upon translocation to the cell membrane. (1996) *Mol Cell Biol* **16**, 1770-1776

64. Kiyokawa, E., Hashimoto, Y., Kobayashi, S., Sugimura, H., Kurata, T., and Matsuda, M. Activation of Rac1 by a Crk SH3-binding protein, DOCK180. (1998) *Genes Dev* **12**, 3331-3336

65. Wu, Y. C., Tsai, M. C., Cheng, L. C., Chou, C. J., and Weng, N. Y. C. elegans CED-12 acts in the conserved crkII/DOCK180/Rac pathway to control cell migration and cell corpse engulfment. (2001) *Dev Cell* **1**, 491-502

66. Zhou, Z., Caron, E., Hartwieg, E., Hall, A., and Horvitz, H. R. The C. elegans PH domain protein CED-12 regulates cytoskeletal reorganization via a Rho/Rac GTPase signaling pathway. (2001) *Dev Cell* **1**, 477-489

67. Gumienny, T. L., Brugnera, E., Tosello-Trampont, A. C., Kinchen, J. M., Haney, L. B., Nishiwaki, K., Walk, S. F., Nemergut, M. E., Macara, I. G., Francis, R., Schedl, T., Qin, Y., Van Aelst, L., Hengartner, M. O., and Ravichandran, K. S. CED-12/ELMO, a novel member of the CrkII/Dock180/Rac pathway, is required for phagocytosis and cell migration. (2001) *Cell* **107**, 27-41

68. Brugnera, E., Haney, L., Grimsley, C., Lu, M., Walk, S. F., Tosello-Trampont, A. C., Macara, I. G., Madhani, H., Fink, G. R., and Ravichandran, K. S. Unconventional Rac-GEF activity is mediated through the Dock180-ELMO complex. (2002) *Nat Cell Biol* **4**, 574-582

69. Nishihara, H., Kobayashi, S., Hashimoto, Y., Ohba, F., Mochizuki, N., Kurata, T., Nagashima, K., and Matsuda, M. Non-adherent cell-specific expression of DOCK2, a member of the human CDM-family proteins. (1999) *Biochim Biophys Acta* **1452**, 179-187

70. Kashiwa, A., Yoshida, H., Lee, S., Paladino, T., Liu, Y., Chen, Q., Dargusch, R., Schubert, D., and Kimura, H. Isolation and characterization of novel presenilin binding protein. (2000) *J Neurochem* **75**, 109-116

71. Namekata, K., Enokido, Y., Iwasawa, K., and Kimura, H. MOCA induces membrane spreading by activating Rac1. (2004) *J Biol Chem* **279**, 14331-14337

72. Meller, N., Irani-Tehrani, M., Kiosses, W. B., Del Pozo, M. A., and Schwartz, M. A. Zizimin1, a novel Cdc42 activator, reveals a new GEF domain for Rho proteins. (2002) *Nat Cell Biol* **4**, 639-647

73. Cote, J. F., and Vuori, K. Identification of an evolutionarily conserved superfamily of DOCK180-related proteins with guanine nucleotide exchange activity. (2002) *J Cell Sci* **115**, 4901-4913

74. Cote, J. F., Motoyama, A. B., Bush, J. A., and Vuori, K. A novel and evolutionarily conserved PtdIns(3,4,5)P3-binding domain is necessary for DOCK180 signalling. (2005) *Nat Cell Biol* **7**, 797-807

75. Kanai, A., Ihara, S., Ohdaira, T., Shinohara-Kanda, A., Iwamatsu, A., and Fukui, Y. Identification of DOCK4 and its splicing variant as PIP3 binding proteins. (2008) *IUBMB Life* **60**, 467-472

76. Meller, N., Merlot, S., and Guda, C. CZH proteins: a new family of Rho-GEFs. (2005) *J Cell Sci* **118**, 4937-4946

77. Cote, J. F., and Vuori, K. In vitro guanine nucleotide exchange activity of DHR-2/DOCKER/CZH2 domains. (2006) *Methods Enzymol* **406**, 41-57

78. Premkumar, L., Bobkov, A. A., Patel, M., Jaroszewski, L., Bankston, L. A., Stec, B., Vuori, K., Cote, J. F., and Liddington, R. C. Structural basis of membrane-targeting by the Dock180 family of Rho-family guanine exchange factors (RhoGEFs). (2010) *J Biol Chem*

79. Yang, J., Zhang, Z., Roe, S. M., Marshall, C. J., and Barford, D. Activation of Rho GTPases by DOCK exchange factors is mediated by a nucleotide sensor. (2009) *Science* **325**, 1398-1402

80. Tosello-Trampont, A. C., Brugnera, E., and Ravichandran, K. S. Evidence for a conserved role for CRKII and Rac in engulfment of apoptotic cells. (2001) *J Biol Chem* **276**, 13797-13802

81. Albert, M. L., Kim, J. I., and Birge, R. B. alphavbeta5 integrin recruits the CrkII-Dock180-rac1 complex for phagocytosis of apoptotic cells. (2000) *Nat Cell Biol* **2**, 899-905

82. Akakura, S., Singh, S., Spataro, M., Akakura, R., Kim, J. I., Albert, M. L., and Birge, R. B. The opsonin MFG-E8 is a ligand for the alphavbeta5 integrin and triggers DOCK180-dependent Rac1 activation for the phagocytosis of apoptotic cells. (2004) *Exp Cell Res* **292**, 403-416

83. Kiyokawa, E., Hashimoto, Y., Kurata, T., Sugimura, H., and Matsuda, M. Evidence that DOCK180 up-regulates signals from the CrkII-p130(Cas) complex. (1998) *J Biol Chem* **273**, 24479-24484

84. Tu, Y., Kucik, D. F., and Wu, C. Identification and kinetic analysis of the interaction between Nck-2 and DOCK180. (2001) *FEBS Lett* **491**, 193-199

85. Park, D., Tosello-Trampont, A. C., Elliott, M. R., Lu, M., Haney, L. B., Ma, Z., Klibanov, A. L., Mandell, J. W., and Ravichandran, K. S. BAI1 is an engulfment receptor for apoptotic cells upstream of the ELMO/Dock180/Rac module. (2007) *Nature* **450**, 430-434

86. Santy, L. C., Ravichandran, K. S., and Casanova, J. E. The DOCK180/Elmo complex couples ARNO-mediated Arf6 activation to the downstream activation of Rac1. (2005) *Curr Biol* **15**, 1749-1754

87. White, D. T., McShea, K. M., Attar, M. A., and Santy, L. C. GRASP and IPCEF Promote ARF to Rac Signaling and Cell Migration by Coordinating the Association of ARNO/Cytohesin 2 with Dock180. (2009) *Mol Biol Cell*

88. Jarzynka, M. J., Hu, B., Hui, K. M., Bar-Joseph, I., Gu, W., Hirose, T., Haney, L. B., Ravichandran, K. S., Nishikawa, R., and Cheng, S. Y. ELMO1 and Dock180, a bipartite Rac1 guanine nucleotide exchange factor, promote human glioma cell invasion. (2007) *Cancer Res* **67**, 7203-7211

89. Laurin, M., Fradet, N., Blangy, A., Hall, A., Vuori, K., and Cote, J. F. The atypical Rac activator Dock180 (Dock1) regulates myoblast fusion in vivo. (2008) *Proc Natl Acad Sci U S A* **105**, 15446-15451

90. Li, X., Gao, X., Liu, G., Xiong, W., Wu, J., and Rao, Y. Netrin signal transduction and the guanine nucleotide exchange factor DOCK180 in attractive signaling. (2008) *Nat Neurosci* **11**, 28-35

91. Kobayashi, S., Shirai, T., Kiyokawa, E., Mochizuki, N., Matsuda, M., and Fukui, Y. Membrane recruitment of DOCK180 by binding to PtdIns(3,4,5)P3. (2001) *Biochem J* **354**, 73-78

92. Nishikimi, A., Fukuhara, H., Su, W., Hongu, T., Takasuga, S., Mihara, H., Cao, Q., Sanematsu, F., Kanai, M., Hasegawa, H., Tanaka, Y., Shibasaki, M., Kanaho, Y., Sasaki, T., Frohman, M. A., and Fukui, Y. Sequential regulation of DOCK2 dynamics by two phospholipids during neutrophil chemotaxis. (2009) *Science* **324**, 384-387

93. Yin, J., Haney, L., Walk, S., Zhou, S., Ravichandran, K. S., and Wang, W. Nuclear localization of the DOCK180/ELMO complex. (2004) *Arch Biochem Biophys* **429**, 23-29

94. Lu, M., Kinchen, J. M., Rossman, K. L., Grimsley, C., Hall, M., Sondek, J., Hengartner, M. O., Yajnik, V., and Ravichandran, K. S. A Steric-inhibition model for regulation of nucleotide exchange via the Dock180 family of GEFs. (2005) *Curr Biol* **15**, 371-377

95. Tachibana, M., Kiyokawa, E., Hara, S., Iemura, S., Natsume, T., Manabe, T., and Matsuda, M. Ankyrin repeat domain 28 (ANKRD28), a novel binding partner of DOCK180, promotes cell migration by regulating focal adhesion formation. (2009) *Exp Cell Res* **315**, 863-876

96. Nishihara, H., Maeda, M., Oda, A., Tsuda, M., Sawa, H., Nagashima, K., and Tanaka, S. DOCK2 associates with CrkL and regulates Rac1 in human leukemia cell lines. (2002) *Blood* **100**, 3968-3974

97. Fukui, Y., Hashimoto, O., Sanui, T., Oono, T., Koga, H., Abe, M., Inayoshi, A., Noda, M., Oike, M., Shirai, T., and Sasazuki, T. Haematopoietic cell-specific CDM family protein DOCK2 is essential for lymphocyte migration. (2001) *Nature* **412**, 826-831

98. Jiang, H., Pan, F., Erickson, L. M., Jang, M. S., Sanui, T., Kunisaki, Y., Sasazuki, T., Kobayashi, M., and Fukui, Y. Deletion of DOCK2, a regulator of the actin cytoskeleton in lymphocytes, suppresses cardiac allograft rejection. (2005) *J Exp Med* **202**, 1121-1130

99. Kunisaki, Y., Nishikimi, A., Tanaka, Y., Takii, R., Noda, M., Inayoshi, A., Watanabe, K., Sanematsu, F., Sasazuki, T., Sasaki, T., and Fukui, Y. DOCK2 is a

Rac activator that regulates motility and polarity during neutrophil chemotaxis. (2006) *J Cell Biol* **174**, 647-652

100. Janardhan, A., Swigut, T., Hill, B., Myers, M. P., and Skowronski, J. HIV-1 Nef binds the DOCK2-ELMO1 complex to activate rac and inhibit lymphocyte chemotaxis. (2004) *PLoS Biol* **2**, E6

101. Sanui, T., Inayoshi, A., Noda, M., Iwata, E., Stein, J. V., Sasazuki, T., and Fukui, Y. DOCK2 regulates Rac activation and cytoskeletal reorganization through interaction with ELMO1. (2003) *Blood* **102**, 2948-2950

102. Chen, Q., Kimura, H., and Schubert, D. A novel mechanism for the regulation of amyloid precursor protein metabolism. (2002) *J Cell Biol* **158**, 79-89

103. Chen, Q., Peto, C. A., Shelton, G. D., Mizisin, A., Sawchenko, P. E., and Schubert, D. Loss of modifier of cell adhesion reveals a pathway leading to axonal degeneration. (2009) *J Neurosci* **29**, 118-130

104. Chen, Q., Chen, T. J., Letourneau, P. C., Costa Lda, F., and Schubert, D. Modifier of cell adhesion regulates N-cadherin-mediated cell-cell adhesion and neurite outgrowth. (2005) *J Neurosci* **25**, 281-290

105. Caspi, E., and Rosin-Arbesfeld, R. A novel functional screen in human cells identifies MOCA as a negative regulator of Wnt signaling. (2008) *Mol Biol Cell* **19**, 4660-4674

106. Yajnik, V., Paulding, C., Sordella, R., McClatchey, A. I., Saito, M., Wahrer, D. C., Reynolds, P., Bell, D. W., Lake, R., van den Heuvel, S., Settleman, J., and Haber, D. A. DOCK4, a GTPase activator, is disrupted during tumorigenesis. (2003) *Cell* **112**, 673-684

107. Hiramoto, K., Negishi, M., and Katoh, H. Dock4 is regulated by RhoG and promotes Rac-dependent cell migration. (2006) *Exp Cell Res* **312**, 4205-4216

108. Yan, D., Li, F., Hall, M. L., Sage, C., Hu, W. H., Giallourakis, C., Upadhyay, G., Ouyang, X. M., Du, L. L., Bethea, J. R., Chen, Z. Y., Yajnik, V., and Liu, X. Z. An isoform of GTPase regulator DOCK4 localizes to the stereocilia in the inner ear and binds to harmonin (USH1C). (2006) *J Mol Biol* **357**, 755-764

109. Upadhyay, G., Goessling, W., North, T. E., Xavier, R., Zon, L. I., and Yajnik, V. Molecular association between beta-catenin degradation complex and Rac guanine exchange factor DOCK4 is essential for Wnt/beta-catenin signaling. (2008) *Oncogene* **27**, 5845-5855

110. Brazier, H., Stephens, S., Ory, S., Fort, P., Morrison, N., and Blangy, A. Expression profile of RhoGTPases and RhoGEFs during RANKL-stimulated osteoclastogenesis: identification of essential genes in osteoclasts. (2006) *J Bone Miner Res* **21**, 1387-1398

111. Omi, N., Kiyokawa, E., Matsuda, M., Kinoshita, K., Yamada, S., Yamada, K., Matsushima, Y., Wang, Y., Kawai, J., Suzuki, M., Hayashizaki, Y., and Hiai, H. Mutation of Dock5, a member of the guanine exchange factor Dock180 superfamily, in the rupture of lens cataract mouse. (2008) *Exp Eye Res* **86**, 828-834

112. Moore, C. A., Parkin, C. A., Bidet, Y., and Ingham, P. W. A role for the Myoblast city homologues Dock1 and Dock5 and the adaptor proteins Crk and Crk-like in zebrafish myoblast fusion. (2007) *Development* **134**, 3145-3153

113. Erickson, M. R., Galletta, B. J., and Abmayr, S. M. Drosophila myoblast city encodes a conserved protein that is essential for myoblast fusion, dorsal closure, and cytoskeletal organization. (1997) *J Cell Biol* **138**, 589-603

114. Wu, Y. C., and Horvitz, H. R. C. elegans phagocytosis and cell-migration protein CED-5 is similar to human DOCK180. (1998) *Nature* **392**, 501-504

115. Blelloch, R., Newman, C., and Kimble, J. Control of cell migration during Caenorhabditis elegans development. (1999) *Curr Opin Cell Biol* **11**, 608-613

116. Para, A., Krischke, M., Merlot, S., Shen, Z., Oberholzer, M., Lee, S., Briggs, S., and Firtel, R. A. Dictyostelium Dock180-related RacGEFs regulate the actin cytoskeleton during cell motility. (2009) *Mol Biol Cell* **20**, 699-707

117. Basu, D., Le, J., Zakharova, T., Mallery, E. L., and Szymanski, D. B. A SPIKE1 signaling complex controls actin-dependent cell morphogenesis through the heteromeric WAVE and ARP2/3 complexes. (2008) *Proc Natl Acad Sci U S A* **105**, 4044-4049

118. Grimsley, C. M., Kinchen, J. M., Tosello-Trampont, A. C., Brugnera, E., Haney, L. B., Lu, M., Chen, Q., Klingele, D., Hengartner, M. O., and Ravichandran, K. S. Dock180 and ELMO1 proteins cooperate to promote evolutionarily conserved Rac-dependent cell migration. (2004) *J Biol Chem* **279**, 6087-6097

119. Komander, D., Patel, M., Laurin, M., Fradet, N., Pelletier, A., Barford, D., and Cote, J. F. An alpha-helical extension of the ELMO1 pleckstrin homology domain mediates direct interaction to DOCK180 and is critical in Rac signaling. (2008) *Mol Biol Cell* **19**, 4837-4851

120. Lu, M., Kinchen, J. M., Rossman, K. L., Grimsley, C., deBakker, C., Brugnera, E., Tosello-Trampont, A. C., Haney, L. B., Klingele, D., Sondek, J., Hengartner, M. O., and Ravichandran, K. S. PH domain of ELMO functions in trans to regulate Rac activation via Dock180. (2004) *Nat Struct Mol Biol* **11**, 756-762

121. Meller, N., Irani-Tehrani, M., Ratnikov, B. I., Paschal, B. M., and Schwartz, M. A. The novel Cdc42 guanine nucleotide exchange factor, zizimin1, dimerizes via the Cdc42-binding CZH2 domain. (2004) *J Biol Chem* **279**, 37470-37476

122. Makino, Y., Tsuda, M., Ichihara, S., Watanabe, T., Sakai, M., Sawa, H., Nagashima, K., Hatakeyama, S., and Tanaka, S. Elmo1 inhibits ubiquitylation of Dock180. (2006) *J Cell Sci* **119**, 923-932

123. Katoh, H., and Negishi, M. RhoG activates Rac1 by direct interaction with the Dock180-binding protein Elmo. (2003) *Nature* **424**, 461-464

124. Handa, Y., Suzuki, M., Ohya, K., Iwai, H., Ishijima, N., Koleske, A. J., Fukui, Y., and Sasakawa, C. Shigella IpgB1 promotes bacterial entry through the ELMO-Dock180 machinery. (2007) *Nat Cell Biol* **9**, 121-128

125. Scott, M. P., Zappacosta, F., Kim, E. Y., Annan, R. S., and Miller, W. T. Identification of novel SH3 domain ligands for the Src family kinase Hck. Wiskott-Aldrich syndrome protein (WASP), WASP-interacting protein (WIP), and ELMO1. (2002) *J Biol Chem* **277**, 28238-28246

126. Yokoyama, N., deBakker, C. D., Zappacosta, F., Huddleston, M. J., Annan, R. S., Ravichandran, K. S., and Miller, W. T. Identification of tyrosine residues on ELMO1 that are phosphorylated by the Src-family kinase Hck. (2005) *Biochemistry* **44**, 8841-8849

127. Grimsley, C. M., Lu, M., Haney, L. B., Kinchen, J. M., and Ravichandran, K. S. Characterization of a novel interaction between ELMO1 and ERM proteins. (2006) *J Biol Chem* **281**, 5928-5937

128. Miyamoto, Y., Yamauchi, J., Sanbe, A., and Tanoue, A. Dock6, a Dock-C subfamily guanine nucleotide exchanger, has the dual specificity for Rac1 and Cdc42 and regulates neurite outgrowth. (2007) *Exp Cell Res* **313**, 791-804

129. Watabe-Uchida, M., John, K. A., Janas, J. A., Newey, S. E., and Van Aelst, L. The Rac activator DOCK7 regulates neuronal polarity through local phosphorylation of stathmin/Op18. (2006) *Neuron* **51**, 727-739

130. Yamauchi, J., Miyamoto, Y., Chan, J. R., and Tanoue, A. ErbB2 directly activates the exchange factor Dock7 to promote Schwann cell migration. (2008) *J Cell Biol* **181**, 351-365

131. Blasius, A. L., Brandl, K., Crozat, K., Xia, Y., Khovananth, K., Krebs, P., Smart, N. G., Zampolli, A., Ruggeri, Z. M., and Beutler, B. A. Mice with mutations of Dock7 have generalized hypopigmentation and white-spotting but show normal neurological function. (2009) *Proc Natl Acad Sci U S A* **106**, 2706-2711

132. Ruusala, A., and Aspenstrom, P. Isolation and characterisation of DOCK8, a member of the DOCK180-related regulators of cell morphology. (2004) *FEBS Lett* **572**, 159-166

133. Takahashi, K., Kohno, T., Ajima, R., Sasaki, H., Minna, J. D., Fujiwara, T., Tanaka, N., and Yokota, J. Homozygous deletion and reduced expression of the DOCK8 gene in human lung cancer. (2006) *Int J Oncol* **28**, 321-328

134. Saelee, P., Wongkham, S., Puapairoj, A., Khuntikeo, N., Petmitr, S., Chariyalertsak, S., Sumethchotimaytha, W., and Karalak, A. Novel PNLIPRP3 and DOCK8 gene expression and prognostic implications of DNA loss on chromosome 10q25.3 in hepatocellular carcinoma. (2009) *Asian Pac J Cancer Prev* **10**, 501-506

135. Griggs, B. L., Ladd, S., Saul, R. A., DuPont, B. R., and Srivastava, A. K. Dedicator of cytokinesis 8 is disrupted in two patients with mental retardation and developmental disabilities. (2008) *Genomics* **91**, 195-202

136. Zhang, Q., Davis, J. C., Lamborn, I. T., Freeman, A. F., Jing, H., Favreau, A. J., Matthews, H. F., Davis, J., Turner, M. L., Uzel, G., Holland, S. M., and Su, H. C. Combined immunodeficiency associated with DOCK8 mutations. (2009) *N Engl J Med* **361**, 2046-2055

137. Randall, K. L., Lambe, T., Johnson, A., Treanor, B., Kucharska, E., Domaschenz, H., Whittle, B., Tze, L. E., Enders, A., Crockford, T. L., Bouriez-Jones, T., Alston, D., Cyster, J. G., Lenardo, M. J., Mackay, F., Deenick, E. K., Tangye, S. G., Chan, T. D., Camidge, T., Brink, R., Vinuesa, C. G., Batista, F. D., Cornall, R. J., and Goodnow, C. C. Dock8 mutations cripple B cell immunological synapses, germinal centers and long-lived antibody production. (2009) *Nat Immunol* **10**, 1283-1291

138. Yelo, E., Bernardo, M. V., Gimeno, L., Alcaraz-Garcia, M. J., Majado, M. J., and Parrado, A. Dock10, a novel CZH protein selectively induced by interleukin-4 in human B lymphocytes. (2008) *Mol Immunol* **45**, 3411-3418

139. Kuramoto, K., Negishi, M., and Katoh, H. Regulation of dendrite growth by the Cdc42 activator Zizimin1/Dock9 in hippocampal neurons. (2009) *J Neurosci Res* **87**, 1794-1805

140. Detera-Wadleigh, S. D., Liu, C. Y., Maheshwari, M., Cardona, I., Corona, W., Akula, N., Steele, C. J., Badner, J. A., Kundu, M., Kassem, L., Potash, J. B., Gibbs, R., Gershon, E. S., and McMahon, F. J. Sequence variation in DOCK9 and heterogeneity in bipolar disorder. (2007) *Psychiatr Genet* **17**, 274-286

141. Nishikimi, A., Meller, N., Uekawa, N., Isobe, K., Schwartz, M. A., and Maruyama, M. Zizimin2: a novel, DOCK180-related Cdc42 guanine nucleotide exchange factor expressed predominantly in lymphocytes. (2005) *FEBS Lett* **579**, 1039-1046

142. Gadea, G., Sanz-Moreno, V., Self, A., Godi, A., and Marshall, C. J. DOCK10-mediated Cdc42 activation is necessary for amoeboid invasion of melanoma cells. (2008) *Curr Biol* **18**, 1456-1465

143. Lin, Q., Yang, W., Baird, D., Feng, Q., and Cerione, R. A. Identification of a DOCK180-related guanine nucleotide exchange factor that is capable of mediating a positive feedback activation of Cdc42. (2006) *J Biol Chem* **281**, 35253-35262

144. Kwofie, M. A., and Skowronski, J. Specific recognition of Rac2 and Cdc42 by DOCK2 and DOCK9 guanine nucleotide exchange factors. (2008) *J Biol Chem* **283**, 3088-3096

145. Peifer, M., Berg, S., and Reynolds, A. B. A repeating amino acid motif shared by proteins with diverse cellular roles. (1994) *Cell* **76**, 789-791

146. Mizuno, T., Kaibuchi, K., Yamamoto, T., Kawamura, M., Sakoda, T., Fujioka, H., Matsuura, Y., and Takai, Y. A stimulatory GDP/GTP exchange protein for smg p21 is active on the post-translationally processed form of c-Ki-ras p21 and rhoA p21. (1991) *Proc Natl Acad Sci U S A* **88**, 6442-6446

147. Yamamoto, T., Kaibuchi, K., Mizuno, T., Hiroyoshi, M., Shirataki, H., and Takai, Y. Purification and characterization from bovine brain cytosol of proteins that regulate the GDP/GTP exchange reaction of smg p21s, ras p21-like GTP-binding proteins. (1990) *J Biol Chem* **265**, 16626-16634

148. Hiraoka, K., Kaibuchi, K., Ando, S., Musha, T., Takaishi, K., Mizuno, T., Asada, M., Menard, L., Tomhave, E., Didsbury, J., Snyderman, R., and Takai, Y. Both stimulatory and inhibitory GDP/GTP exchange proteins, smg GDS and rho GDI, are active on multiple small GTP-binding proteins. (1992) *Biochem Biophys Res Commun* **182**, 921-930

149. Chuang, T. H., Xu, X., Quilliam, L. A., and Bokoch, G. M. SmgGDS stabilizes nucleotide-bound and -free forms of the Rac1 GTP-binding protein and stimulates GTP/GDP exchange through a substituted enzyme mechanism. (1994) *Biochem J* **303 (Pt 3)**, 761-767

150. Yaku, H., Sasaki, T., and Takai, Y. The Dbl Oncogene Product as a GDP/GTP Exchange Protein for the Rho Family: Its Properties in Comparison with Those of Smg GDS. (1994) *Biochem Biophys Res Commun* **198**, 811-817

151. Isomura, M., Kaibuchi, K., Yamamoto, T., Kawamura, S., Katayama, M., and Takai, Y. Partial purification and characterization of GDP dissociation stimulator (GDS) for the rho proteins from bovine brain cytosol. (1990) *Biochem Biophys Res Commun* **169**, 652-659

152. Kaibuchi, K., Mizuno, T., Fujioka, H., Yamamoto, T., Kishi, K., Fukumoto, Y., Hori, Y., and Takai, Y. Molecular cloning of the cDNA for stimulatory GDP/GTP exchange protein for smg p21s (ras p21-like small GTP-binding proteins) and

characterization of stimulatory GDP/GTP exchange protein. (1991) *Mol Cell Biol* **11**, 2873-2880

153. Orita, S., Kaibuchi, K., Kuroda, S., Shimizu, K., Nakanishi, H., and Takai, Y. Comparison of kinetic properties between two mammalian ras p21 GDP/GTP exchange proteins, ras guanine nucleotide-releasing factor and smg GDP dissociation stimulation. (1993) *J Biol Chem* **268**, 25542-25546

154. Sasaki, T., Kato, M., Nishiyama, T., and Takai, Y. The Nucleotide Exchange Rates of Rho and Rac Small GTP-Binding Proteins Are Enhanced to Different Extents by Their Regulatory Protein Smg GDS. (1993) *Biochem Biophys Res Commun* **194**, 1188-1193

155. Hutchinson, J. P., and Eccleston, J. F. Mechanism of nucleotide release from Rho by the GDP dissociation stimulator protein. (2000) *Biochemistry* **39**, 11348-11359

156. Kawamura, S., Kaibuchi, K., Hiroyoshi, M., Hata, Y., and Takai, Y. Stoichiometric interaction of smg p21 with its GDP/GTP exchange protein and its novel action to regulate the translocation of smg p21 between membrane and cytoplasm. (1991) *Biochem Biophys Res Commun* **174**, 1095-1102

157. Nakanishi, H., Kaibuchi, K., Orita, S., Ueno, N., and Takai, Y. Different functions of Smg GDP dissociation stimulator and mammalian counterpart of yeast Cdc25. (1994) *J Biol Chem* **269**, 15085-15091

158. Vikis, H. G., Stewart, S., and Guan, K.-L. SmgGDS displays differential binding and exchange activity towards different Ras isoforms. (2002) *Oncogene* **21**, 2425-2432

159. Kawamura, M., Kaibuchi, K., Kishi, K., and Takai, Y. Translocation of Ki-ras P21 between Membrane and Cytoplasm by smg GDS. (1993) *Biochem Biophys Res Commun* **190**, 832-841

160. Iouzalen, N., Camonis, J., and Moreau, J. Identification and characterization in Xenopus of XsmgGDS, a RalB binding protein. (1998) *Biochem Biophys Res Commun* **250**, 359-363

161. Vithalani, K. K., Parent, C. A., Thorn, E. M., Penn, M., Larochelle, D. A., Devreotes, P. N., and De Lozanne, A. Identification of darlin, a Dictyostelium protein with Armadillo-like repeats that binds to small GTPases and is important

for the proper aggregation of developing cells. (1998) *Mol Bio of the cell* **9**, 3095-3106

162. Hiroyoshi, M., Kaibuchi, K., Kawamura, S., Hata, Y., and Takai, Y. Role of the C-terminal region of smg p21, a ras p21-like small GTP-binding protein, in membrane and smg p21 GDP/GTP exchange protein interactions. (1991) *J Biol Chem* **266**, 2962-2969

163. Hori, Y., Kikuchi, A., Isomura, M., Katayama, M., Miura, Y., Fujioka, H., Kaibuchi, K., and Takai, Y. Post-translational modifications of the C-terminal region of the rho protein are important for its interaction with membranes and the stimulatory and inhibitory GDP/GTP exchange proteins. (1991) *Oncogene* **6**, 515-522

164. Shirataki, H., Kaibuchi, K., Hiroyoshi, M., Isomura, M., Araki, S., Sasaki, T., and Takai, Y. Inhibition of the action of the stimulatory GDP/GTP exchange protein for smg p21 by the geranylgeranylated synthetic peptides designed from its C-terminal region. (1991) *J Biol Chem* **266**, 20672-20677

165. Hutchinson, J. P., Rittinger, K., and Eccleston, J. F. Purification and characterization of guanine nucleotide dissociation stimulator protein. (2000) *Methods Enzymol* **325**, 71-82

166. Kotani, K., Kikuchi, A., Doi, K., Kishida, S., Sakoda, T., Kishi, K., and Takai, Y. The functional domain of the stimulatory GDP/GTP exchange protein (smg GDS) which interacts with the C-terminal geranylgeranylated region of rap1/Krev-1/smg p21. (1992) *Oncogene* **7**, 1699-1704

167. Lanning, C. C., Ruiz-Velasco, R., and Williams, C. L. Novel mechanism of the co-regulation of nuclear transport of SmgGDS and Rac1. (2003) *J Biol Chem* **278**, 12495-12506

168. Kawamura, S., Kaibuchi, K., Hiroyoshi, M., Fujioka, H., Mizuno, T., and Takai, Y. Inhibition of the action of a stimulatory GDP/GTP exchange protein for smg p21 by acidic membrane phospholipids. (1991) *Jpn J Cancer Res* **82**, 758-761

169. Hata, Y., Kaibuchi, K., Kawamura, S., Hiroyoshi, M., Shirataki, H., and Takai, Y. Enhancement of the actions of smg p21 GDP/GTP exchange protein by the protein kinase A-catalyzed phosphorylation of smg p21. (1991) *J Biol Chem* **266**, 6571-6577

170. Yoshida, Y., Kawata, M., Miura, Y., Musha, T., Sasaki, T., Kikuchi, A., and Takai, Y. Microinjection of smg/rap1/Krev-1 p21 into Swiss 3T3 cells induces DNA synthesis and morphological changes. (1992) *Mol Cell Biol* **12**, 3407-3414

171. Shimizu, K., Kawabe, H., Minami, S., Honda, T., Takaishi, K., Shirataki, H., and Takai, Y. SMAP, an Smg GDS-associating protein having arm repeats and phosphorylated by Src tyrosine kinase. (1996) *J Biol Chem* **271**, 27013-27017

172. Teng, J., Rai, T., Tanaka, Y., Takei, Y., Nakata, T., Hirasawa, M., Kulkarni, A. B., and Hirokawa, N. The KIF3 motor transports N-cadherin and organizes the developing neuroepithelium. (2005) *Nat Cell Biol* **7**, 474-482

173. Haraguchi, K., Hayashi, T., Jimbo, T., Yamamoto, T., and Akiyama, T. Role of the kinesin-2 family protein, KIF3, during mitosis. (2006) *J Biol Chem* **281**, 4094-4099

174. Brown, C. L., Maier, K. C., Stauber, T., Ginkel, L. M., Wordeman, L., Vernos, I., and Schroer, T. A. Kinesin-2 is a motor for late endosomes and lysosomes. (2005) *Traffic* **6**, 1114-1124

175. Jimbo, T., Kawasaki, Y., Koyama, R., Sato, R., Takada, S., Haraguchi, K., and Akiyama, T. Identification of a link between the tumour suppressor APC and the kinesin superfamily. (2002) *Nat Cell Biol* **4**, 323-327

176. Shin, E.-Y., Lee, C.-S., Cho, T. G., Kim, Y. G., Song, S., Juhnn, Y.-S., Park, S. C., Manser, E., and Kim, E.-G. betaPak-interacting exchange factor-mediated Rac1 activation requires smgGDS guanine nucleotide exchange factor in basic fibroblast growth factor-induced neurite outgrowth. (2006) *J Biol Chem* **281**, 35954-35964

177. Strassheim, D., Porter, R. A., Phelps, S. H., and Williams, C. L. Unique in Vivo Associations with SmgGDS and RhoGDI and Different Guanine Nucleotide Exchange Activities Exhibited by RhoA, Dominant Negative RhoAAsn-19, and Activated RhoAVal-14. (2000) *J Biol Chem* **275**, 6699-6702

178. Williams, C. L. The polybasic region of Ras and Rho family small GTPases: a regulator of protein interactions and membrane association and a site of nuclear localization signal sequences. (2003) *Cellular Signalling* **15**, 1071-1080

179. Thill, R., Campbell, W. B., and Williams, C. L. Identification and characterization of the unique guanine nucleotide exchange factor, SmgGDS, in vascular smooth muscle cells. (2008) *J Cell Biochem* **104**, 1760-1770

180. Takakura, A., Miyoshi, J., Ishizaki, H., Tanaka, M., Togawa, A., Nishizawa, Y., Yoshida, H., Nishikawa, S., and Takai, Y. Involvement of a small GTP-binding protein (G protein) regulator, small G protein GDP dissociation stimulator, in antiapoptotic cell survival signaling. (2000) *Mol Biol Cell* **11**, 1875-1886

181. Fujioka, H., Kaibuchi, K., Kishi, K., Yamamoto, T., Kawamura, M., Sakoda, T., Mizuno, T., and Takai, Y. Transforming and c-fos promoter/enhancer-stimulating activities of a stimulatory GDP/GTP exchange protein for small GTP-binding proteins. (1992) *J Biol Chem* **267**, 926-930

182. Tew, G. W., Lorimer, E. L., Berg, T. J., Zhi, H., Li, R., and Williams, C. L. SmgGDS regulates cell proliferation, migration, and NF-kappaB transcriptional activity in non-small cell lung carcinoma. (2008) *J Biol Chem* **283**, 963-976

183. Zhi, H., Yang, X. J., Kuhnmuench, J., Berg, T., Thill, R., Yang, H., See, W. A., Becker, C. G., Williams, C. L., and Li, R. SmgGDS is up-regulated in prostate carcinoma and promotes tumour phenotypes in prostate cancer cells. (2009) *Pathol J* **217**, 389-397

184. Hussey, D. J., Nicola, M., Moore, S., Peters, G. B., and Dobrovic, A. The (4;11)(q21;p15) translocation fuses the NUP98 and RAP1GDS1 genes and is recurrent in T-cell acute lymphocytic leukemia. (1999) *Blood* **94**, 2072-2079

185. Letunic, I., Copley, R. R., Schmidt, S., Ciccarelli, F. D., Doerks, T., Schultz, J., Ponting, C. P., and Bork, P. SMART 4.0: towards genomic data integration. (2004) *Nucleic Acids Res* **32**, D142-144

186. Kelley, L. A., MacCallum, R. M., and Sternberg, M. J. Enhanced genome annotation using structural profiles in the program 3D-PSSM. (2000) *J Mol Biol* **299**, 499-520

187. Lupas, A., Van Dyke, M., and Stock, J. Predicting coiled coils from protein sequences. (1991) *Science* **252**, 1162-1164

188. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J.

D., Gibson, T. J., and Higgins, D. G. Clustal W and Clustal X version 2.0. (2007) *Bioinformatics* **23**, 2947-2948

189. McGuffin, L. J., Bryson, K., and Jones, D. T. The PSIPRED protein structure prediction server. (2000) *Bioinformatics* **16**, 404-405

190. Ouali, M., and King, R. D. Cascaded multiple classifiers for secondary structure prediction. (2000) *Protein Sci* **9**, 1162-1176

191. Pollastri, G., Przybylski, D., Rost, B., and Baldi, P. Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles. (2002) *Proteins* **47**, 228-235

192. Karplus, K., Barrett, C., and Hughey, R. Hidden Markov models for detecting remote protein homologies. (1998) *Bioinformatics* **14**, 846-856

193. Stols, L., Gu, M., Dieckman, L., Raffen, R., Collart, F. R., and Donnelly, M. I. A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site. (2002) *Protein Expr Purif* **25**, 8-15

194. Rojas, R. J., Kimple, R. J., Rossman, K. L., Siderovski, D. P., and Sondek, J. Established and emerging fluorescence-based assays for G-protein function: Ras-superfamily GTPases. (2003) *Comb Chem High Throughput Screen* **6**, 409-418

195. Meller, N., Westbrook, M. J., Shannon, J. D., Guda, C., and Schwartz, M. A. Function of the N-terminus of zizimin1: autoinhibition and membrane targeting. (2008) *Biochem J* **409**, 525-533

196. Ikonomou, L., Schneider, Y. J., and Agathos, S. N. Insect cell culture for industrial production of recombinant proteins. (2003) *Appl Microbiol Biotechnol* **62**, 1-20

197. Gotoh, T., Miyazaki, Y., Sato, W., Kikuchi, K., and Bentley, W. E. Proteolytic activity and recombinant protein production in virus-infected Sf-9 insect cell cultures supplemented with carboxyl and cysteine protease inhibitors. (2001) *J Biosci Bioeng* **92**, 248-255

198. Narumiya, S., and Yasuda, S. Rho GTPases in animal cell mitosis. (2006) *Curr Opin Cell Biol* **18**, 199-205

199. Snyder, J. T., Worthylake, D. K., Rossman, K. L., Betts, L., Pruitt, W. M., Siderovski, D. P., Der, C. J., and Sondek, J. Structural basis for the selective activation of Rho GTPases by Dbl exchange factors. (2002) *Nat Struct Biol* **9**, 468-475

200. Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. (2003) *Nucleic Acids Res* **31**, 3784-3788

201. Xing, Y., Takemaru, K., Liu, J., Berndt, J. D., Zheng, J. J., Moon, R. T., and Xu, W. Crystal structure of a full-length beta-catenin. (2008) *Structure* **16**, 478-487

202. Eisenberg, D., Luthy, R., and Bowie, J. U. VERIFY3D: assessment of protein models with three-dimensional profiles. (1997) *Methods Enzymol* **277**, 396-404

203. Soundararajan, M., Turnbull, A., Fedorov, O., Johansson, C., and Doyle, D. A. RhoB can adopt a Mg2+ free conformation prior to GEF binding. (2008) *Proteins* **72**, 498-505

204. Mazhab-Jafari, M. T., Marshall, C. B., Smith, M., Gasmi-Seabrook, G. M., Stambolic, V., Rottapel, R., Neel, B. G., and Ikura, M. Real-time NMR study of three small GTPases reveals that fluorescent mant-tagged nucleotides alter hydrolysis and exchange kinetics. (2009) *J Biol Chem*

205. Walter, T. S., Meier, C., Assenberg, R., Au, K. F., Ren, J., Verma, A., Nettleship, J. E., Owens, R. J., Stuart, D. I., and Grimes, J. M. Lysine methylation as a routine rescue strategy for protein crystallization. (2006) *Structure* **14**, 1617-1622

206. Goldschmidt, L., Cooper, D. R., Derewenda, Z. S., and Eisenberg, D. Toward rational protein crystallization: A Web server for the design of crystallizable protein variants. (2007) *Protein Sci* **16**, 1569-1576

207. Erickson, J. W., and Cerione, R. A. Structural elements, mechanism, and evolutionary convergence of Rho protein-guanine nucleotide exchange factor complexes. (2004) *Biochemistry* **43**, 837-842

208. Corbett, K. D., and Alber, T. The many faces of Ras: recognition of small GTP-binding proteins. (2001) *Trends Biochem Sci* **26**, 710-716