Elucidating G protein signaling and ubiquitin conjugation in *Entamoeba histolytica*

Dustin Eli Bosch

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 School of Medicine (Pharmacology)

Chapel Hill
2013

Approved by:

David P. Siderovski, PhD

T. Kendall Harden, PhD

Brian Kuhlman, PhD

Henrik G. Dohlman, PhD

Scott E. Plevy, MD

ABSTRACT

DUSTIN ELI BOSCH: Elucidating G protein signaling and ubiquitin conjugation in
Entamoeba histolytica
(Under the direction of Dr. David Peter Siderovski)

The intestinal parasite *Entamoeba histolytica* is responsible for an estimated 50 million infections and 100,000 deaths per year worldwide. The causative agent of amoebic colitis and systemic amoebiasis is spread primarily through contaminated food and drinking water sources. Although reasonably effective treatments have long been available for invasive amoebiasis, imperfect patient response rates, drug side effects, and concern for emerging drug resistance all warrant exploration of new pharmacological targets in *E. histolytica*. This work describes structural, biochemical, and cell biological explorations of heterotrimeric G protein and Rho family GTPase signaling and ubiquitination in *E. histolytica*.

The heterotrimeric G protein subunits EhGα1, EhGβ1 and either EhGγ1/2 assembled a typical nucleotide-dependent heterotrimer. Overexpression of wildtype or dominant negative EhGα1 in *E. histolytica* altered chemotactic migration, Matrigel transmigration, host cell attachment, and cell killing. Transcriptomic studies linked EhGα1 expression and altered transcription and secretion of virulence factors. EhGα1 is distinct from the conventional mammalian Gα subfamilies, as revealed by sequence comparison and a crystal structure, but shares functionality with mammalian $Gα_{12/13}$ in engaging an RGS-RhoGEF effector. EhRGS-RhoGEF is apparently autoinhibited, as indicated by structural obstruction

of its catalytic domain in the inactive state. However, co-expression with constitutively active EhGα1 and EhRacC lead to EhRGS-RhoGEF activation in cells.

The Rho family GTPase EhRho1 lacks a signature Rho insert helix and sensitivity to C3 exoenzyme. Crystal structures of EhRho1 indicate unique nucleotide-contacting residues that confer fast intrinsic nucleotide exchange activity. However, EhRho1 functions like its homologs in engaging a diaphanous-related formin effector to regulate actin polymerization. EhFormin1 is autoinhibited and is activated by Rho GTPase binding. A crystal structure of the EhRho1/EhFormin1 complex indicates similarity to human RhoC/mDia1, despite an absent secondary binding site at the Rho insert helix.

Multiple ubiquitin-proteasome pathway genes were differentially transcribed upon perturbed EhGα1 expression. EhUbiquitin was activated by the E1 enzyme EhUba1 and conjugated by the E2 enzyme EhUbc5, indicating a conserved ubiquitination cascade in *E. histolytica*. Crystal structures of EhUbiquitin and EhUbc5 suggested potentially unique polyubiquitin linkages, but a likely conserved mode of ubiquitin chain elongation.

I dedicate my work:

To my wife Samantha for unconditional support, encouragement, and motivation,

as well as preservation of my sanity.

And to my family for past guidance and continued support.

ACKNOWLEDGEMENTS

I am grateful for and indebted to my dissertation advisor and mentor, Dr. David Siderovski, who has most impacted my development as a scientist. His approachability and generosity with both time and resources has been a foundation for my successful graduate training experience, as well as those of other former trainees. I have come to expect from him a cogent word of advice at times when I need guidance, and unfettered freedom to pursue and struggle with my own scientific ideas. Despite his career transitions, David has maintained my training at UNC as a priority, giving freely of his after-hours time to exchange emails and revise manuscripts. From him I have learned the rarity and value of great mentorship in both scientific training and career development, an example that I will strive to emulate in the future.

I also thank other members of the Siderovski lab who have shared their skills and experience and made the lab a fantastic place to work: Brian Buckley, Stephanie Hutsell, Dr. Staci Cohen, Dr. Emily Oestreich, Dr. Patrick Giguere, Dr. Genevieve Laroche, and Dr. Vince Setola. I particularly thank former graduate student Dr. Adam Kimple for his investment in my initial training, his example of success, and for many fruitful collaborations.

I thank fellow members of the UNC MD/PhD training program, in particular, Will Jeck, Erin Steinbach, Jeff Federspiel, Tricia Lenhart, Tom Jarrett, Isaac Chan, and Chris Dibble for their friendship and mutual support along our extensive training experience. The

TABLE OF CONTENTS

LIST OF TABLES

Table:

LIST OF FIGURES

Figure:

xviii

ABBREVIATIONS

Å	Angstrom
A (Ala)	alanine
AlF_4^-	Aluminum tetrafluoride (or tetrafluoroaluminate) ion
AMP	adenosine monophosphate
APS	Advanced Photon Source
A. thaliana	Arabidopsis thaliana
ATP	adenosine triphosphate
BiFC	bimolecular fluorescence complementation
C (Cys)	cysteine
Ca^{2+}	calcium ion
cAMP	cyclic adenosine monophosphate
C. botulinum	Clostridium botulinum
cDNA	complementary DNA
C. elegans	Caenorhabditis elegans
CHO	Chinese Hamster Ovary cells
ConA	concanavalin A
Ct	threshold cycle number
C-terminus	carboxyl terminus
CTX	cholera toxin
D (Asp)	aspartic acid
DAD	Diaphanous autoinhibitory domain
D. discoideum	Dictyostelium discoideum
DH	Dbl homology domain
DID	Diaphanous inhibitory domain
D. melanogaster	Drosophila melanogaster

DMEM	Dulbecco's modified eagles medium
DNA	Deoxyribonucleic acid
DRF	Diaphanous-related formin
DTT	dithiothreitol
DUB	deubiquitinating enzyme
E (Glu)	glutamic acid
E1 (Uba)	ubiquitin activating enzyme
E2 (Ubc)	ubiquitin conjugating enzyme
E3	ubiquitin ligating enzyme
EDTA	ethylenediaminetetraacetic acid
E. dispar	Entamoeba dispar
E. histolytica	Entamoeba histolytica
E. invadens	Entamoeba invadens
E. moshkovskii	Entamoeba moshkovskii
E. terrapinae	Entmaoeba terrapinae
ELISA	Enzyme-linked immunosorbent assay
ER	endoplasmic reticulum
F (Phe)	phenylalanine
F-actin	filamentous actin
FBS	fetal bovine serum
FH1	formin homology 1 domain
FH2	formin homology 2 domain
FH3	formin homology 3 domain
FPKM	fragments per kilobase of exon per million fragments mapped
Fmoc	fluoromethoxy-carbonyl
FP	Fluorescence polarization
FRET	fluorescence resonance energy transfer

g	acceleration by gravity
G (Gly)	glycine
G-protein	Guanine nucleotide binding protein
GAP	GTPase accelerating protein
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GBD	GTPase binding domain
GDI	guanine nucleotide dissociation inhibitor
GDP	guanosine-5'-diphosphate
GEF	guanine nucleotide exchange factor
GFP	green fluorescence protein
GGL	G-gamma like
GoLoco	Gi/o-Loco interaction motif
GPCR	G-protein coupled receptor
GRK	G-protein receptor kinase
GST	glutathione S transferrase
GTP	guanosine-5'-triphosphate
GTPase	GTP hydrolase
$G\alpha$	alpha subunit of heterotrimeric G protein
$G\alpha_i$	"inhibitory" G protein of adenylyl cyclase
$G\alpha_o$	"other"; most abundant G protein in bovine brain
$G\alpha_q$	"queer"; G protein that is not a substrate for ADP-ribosylation
$G\alpha_s$	"stimulatory" G protein of adenylyl cyclase
$G\alpha_t$	"transducin"; abundant G protein in retinal photoreceptor cells
$G\beta$	beta subunit of heterotrimeric G protein
$G\beta\gamma$	beta/gamma subunits of heterotrimeric G protein
$G\gamma$	gamma subunit of heterotrimeric G protein
H (His)	histidine

HA	hemagluttinin epitope tag
HECT	homology to E6AP C-terminal domain
HEK	Human Embryonic Kidney cells
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His_6	hexahistidine epitope tag
HRP	horse radish peroxidase
I (Ile)	isoleucine
IP_3	inositol 1,4,5-trisphosphate
IPTG	isopropyl-beta-D-thiogalactopyranoside
IQF	internally quenched fluorescence
K (Lys)	lysine
K_D	dissociation constant
kD (kDa)	kilodalton
K_m	Michaelis-Menton constant
k_{obs}	observed rate constant
k_{off}	dissociation rate constant
k_{on}	association rate contant
L (Leu)	leucine
LARG	leukemia associated RhoGEF
LPA	lysophosphatidic acid
M (Met)	methionine
M	molar
MAPK	mitogen activated protein kinase
Mg^{2+}	magnesium ion
min	minutes
MLY	N-dimethyl lysine
mRNA	messenger RNA

N (Asn)	asparagine
NMR	nuclear magnetic resonance
N-terminus	amino terminus
Na^+	sodium ion
NaCl	sodium chloride
NHS	N-Hydroxysuccinimide
Ni^{2+}	nickel ion
nm	nanometers
nM	nanomolar
NTA	nickel-nitrilotriacetic acid
ORF	open reading frame
P (Pro)	proline
PAK	p21-activated kinase
PBD	p21 binding domain
PBS	phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PDZ	PSD-95/Dlg/ZO-1 homology domain
PEG	polyethylene glycol
PH	Pleckstrin homology domain
PHMB	p-hydroxy-mercuribenzoic acid
PKA	protein kinase A
PLC	phospholipase-C
P-loop	phosphate binding loop
PP_i	pyrophosphate
PPV	pre-phagosomal vacuole
PTX	pertussis toxin

Q (Gln)	glutamine
R (Arg)	arginine
RBD	Ras binding domain
rgRGS	RhoGEF RGS-like domain
RGS	regulator of G-protein signaling
RhoGEF	Rho family GTPase guanine nucleotide exchange factor
RING	really interesting new gene related domain
RMSD (r.m.s.d.)	root mean square deviation
RNA	ribonucleic acid
ROCK	Rho activated kinase
RT-PCR	reverse transcriptase PCR
RU	resonance units
S (Ser)	serine
S200	sephadex 200 gel exclusion column
SAD	single-wavelength anomalous dispersion
SAXS	small angle X-Ray scattering
S. cerevisiae	Saccharomyces cerevisiae
SDS	sodium dodecyl sulfate
SDS/PAGE	SDS polyacrylamide gel electrophoresis
S.E.M.	standard error of the mean
SeMet	selenomethioine protein derivative
SPR	surface plasmon resonance
T (Thr)	threonine
TBS	tris-buffered saline
TBS-T	TBS with 0.001% tween-20
TCA	tricholoroacetic acid
tet	tetracycline

TEV	tobacco etch virus
TLS	translation/libration/screw parameters
TRIS	tris(hydroxymethyl)aminomethane
Ubi (Ub)	ubiquitin
UFD	ubiquitin fold domain
ULM	ubiquitin-like modifier
UNC	University of North Carolina at Chapel Hill
V (Val)	valine
V_{max}	maximum reaction velocity
W (Trp)	tryptophan
w/v	weight-to-volume
w/w	weight-to-weight
Y (Tyr)	tyrosine
YFP	yellow fluorescent protein

CHAPTER 1

GENERAL INTRODUCTION[1]

1.1   OVERVIEW

The parasite Entamoeba histolytica causes amoebic colitis and systemic amoebiasis. Among the known amoebic factors contributing to pathogenesis are signaling pathways involving heterotrimeric and Ras superfamily G proteins. Here, we review the current knowledge of the roles of heterotrimeric G protein subunits, Ras, Rho, and Rab GTPase families in E. histolytica pathogenesis, as well as of their downstream signaling effectors and nucleotide cycle regulators. Heterotrimeric G protein signaling likely modulates amoebic motility and attachment to and killing of host cells, in part through activation of an RGS-RhoGEF effector. Rho family GTPases, as well as RhoGEFs and Rho effectors (formins and PAKs) regulate the dynamic actin cytoskeleton of E. histolytica and associated pathogenesis-related cellular processes, such as migration, invasion, phagocytosis, and evasion of the host immune response by surface receptor capping. A remarkably large family of 91 Rab GTPases plays multiple roles in a complex amoebic vesicular trafficking system required for phagocytosis and pinocytosis and secretion of known virulence factors, such as amoebapores and cysteine proteases. Although much remains to be discovered, recent studies of G protein signaling in E. histolytica have enhanced our understanding of parasitic pathogenesis and suggested possible targets for pharmacological manipulation.

---

[1] Bosch, D.E. and Siderovski, D.P. (2013) G Protein Signaling in the Parasite *Entamoeba histolytica*. *Exp. Mol. Med.* (accepted review manuscript)

1.2 Entamoeba histolytica CAUSES AMOEBIC COLITIS AND SYSTEMIC AMOEBIASIS

1.2.1 Epidemiology, disease sequelae, and current treatment options

The parasite *Entamoeba histolytica* is the causative agent of infectious amoebic colitis and systemic amoebiasis [1]. The worldwide prevalence of *E. histolytica* infection is not precisely known, with the most recent published estimates (World Health Organization, 1997 [2]) being approximately 50 million infections and 100,000 deaths, annually. Epidemiological estimates have been historically complicated by limitations of diagnostic tests, as well as difficulty in differentiating *E. histolytica* from the morphologically similar, but typically nonpathogenic related *Entamoeba* species, *E. dispar* and *E. moshkovskii* [3]. However, more recently developed antigen detection and PCR-based modalities with improved sensitivity and specificity have allowed more accurate regional estimations of *E. histolytica* infections [4, 5]. The prevalence of *E. histolytica* infection is particularly high among susceptible populations with limited access to clean water. For instance, a study of preschool-aged children in Bangladesh revealed annual infections in 40-50% of subjects [6], a profile of Orang Asli ethnic groups in Malaysia found an overall prevalence of *E. histolytica* positive stool samples to be 15-20% [7], and *E. histolytica* was detected by PCR in 10-15% of a rural Mexican population [8]. The prevalence of antibodies specific for *E. histolytica* in sera of a Chinese population varied from 0.5 to 14%, depending on geographical location [9]. An interrelationship between host nutritional status and susceptibility to *E. histolytica* infection has also recently begun to emerge (reviewed in [10]). Although *E. histolytica* infection is

relatively rare in developed countries, such as the United States, it does occur among travelers, immigrants, and select susceptible subpopulations [11, 12]. Furthermore, outbreaks of *E. histolytica* have occurred due to contaminated municipal water supplies, for example [13].

The life cycle of *E. histolytica* consists of an interchange between an encysted form and the motile, pathogenic trophozoite form. *Entamoeba histolytica* cysts, shed in the feces of infected human hosts, are transmitted primarily by ingestion of contaminated water or food [1]. Excystation occurs in the small intestine, and the resultant *E. histolytica* trophozoites may then colonize the large intestine while evading the host immune response [3]. Although the majority of *E. histolytica* infections are asymptomatic, trophozoites can penetrate the intestinal mucous barrier, resulting in colitis [1]. Amoebic colitis is characterized by trophzoite-mediated killing of intestinal epithelial cells and responding immune cells, as well as local tissue destruction [14]. In rare cases, *E. histolytica* trophozoites can enter the blood stream and spread systemically, giving rise to abscesses, primarily in the liver and less frequently in the lungs and brain [3]. Although systemic amoebiasis requires prior intestinal infection, amoebic liver abscesses can develop in the absence of symptomatic colitis [14, 15] and are known to appear months or years following exposure [16]. Thus, treatment is recommended for patients with *E. histolytica* infection, even in the absence of symptomatic disease [3].

Nitroimidazoles, such as metronidazole, are the current best drugs for the treatment of invasive amoebiasis [3]. Approximately 90% of patients with mild or moderate amoebic colitis respond to nitroimidazole therapy, although persistent intestinal infection often requires additional treatment with paromomycin or diloxanide furoate for complete

eradication [3]. However, a significant fraction of patients with *E. histolytica* infection do not respond to nitroimidazoles, and relatively rare side effects such as allergic reactions, neuropathies, and additional gastrointestinal symptoms can also affect treatment tolerance [17]. Resistance of *E. histolytica* infection to nitroimidazoles and paromomycin has not yet emerged as a major limitation to treatment; however, numerous examples of antibiotic resistance in other microorganisms warrants further exploration of alternative pharmacological therapeutics [18]. A recent study identified auranofin, an FDA-approved rheumatoid arthritis drug, as a potent inhibitor of *E. histolytica* thioredoxin reductase and demonstrated its protective effects in a mouse model of amoebic colitis [19]. Other classes of compounds have also recently been pursued as nanomolar-potency inhibitors of *E. histolytica* growth in culture [20, 21]. Despite existing effective therapies, *E. histolytica* infection and associated disease remains endemic in many parts of the world, particularly in areas with contaminated drinking water and food sources [6, 8]. Problems with sanitation implementation and access to appropriate therapeutics could potentially be circumvented by the development of an *E. histolytica* vaccine, and efforts toward this goal are ongoing (*e.g.* [22]).

1.2.2 Parasite factors in pathogenesis

A number of *E. histolytica* molecular components have been thoroughly established as contributors to pathogenesis. During initial intestinal colonization, *E. histolytica* adheres to the colonic mucin layer primarily through a galactose-inhibitable lectin, known as the Gal/GalNAc lectin (reviewed in [23]). The trimeric surface protein is also a dominant factor in parasite attachment to host cells and subsequent tissue destruction, and functions

interdependently with the dynamic actin cytoskeleton of *E. histolytica* [23]. Trophozoites also secrete pore-forming peptides known as 'amoebapores' that assemble within host cell membranes to trigger cell death (reviewed in [24]). A relatively large family of *E. histolytica*-encoded cysteine proteases also contributes to host cell killing, as well as degradation of the extracellular matrix during invasive amoebic infection and evasion of the host immune response through proteolysis of immunoglobulins and complement (reviewed in [25]). Many regulators of the actin-rich cytoskeleton within *E. histolytica* are also emerging as contributors to pathogenesis-related processes, such as phagocytosis of host cells, trophozoite motility and tissue invasion, and shedding of host antibodies by surface receptor capping (reviewed in [26-29]).

1.3 HETEROTRIMERIC G PROTEINS AND RAS SUPERFAMILY GTPases

Sequencing of the complete *E. histolytica* genome [30] and genome-wide expression studies (*e.g.* [31]) have revealed large numbers of putative cell signaling molecules expressed in the single-celled parasite, including a substantial family of >300 kinases [32]. Also prominent within the *E. histolytica* are genes encoding heterotrimeric G protein subunits (Gα, Gβ, and Gγ) and a large number of small, ~21 kDa G proteins belonging to the Ras superfamily [30]. Gα subunits and Ras GTPases are molecular switches and cellular signaling nodes that bind guanine nucleotides (GTP or GDP) through highly conserved, nucleotide-interacting sequencing motifs [33, 34]. As mammalian G proteins are known to be master regulators of cellular functions spanning cell division and proliferation, cytoskeletal dynamics, vesicular trafficking, and specific responses to extracellular cues [33, 35], it is likely that *E. histolytica* homologs are similarly important for trophozoite biology and pathogenicity. G protein

signaling pathways are also notable for amenability to pharmacological manipulation; particularly, heterotrimeric G protein signaling *via* G protein-coupled receptors (GPCRs) is the target of approximately one third of all currently FDA-approved drugs [36, 37].

1.4 REGULATION OF THE GUANINE NUCLEOTIDE CYCLE

1.4.1 Heterotrimeric G proteins

A Gα subunit in the inactive, GDP-bound state forms a heterotrimer with the obligate Gβγ dimer (Figure 1.1A). A seven-transmembrane G protein-coupled receptor, when activated by an extracellular ligand, engages the heterotrimer and catalyzes the release of GDP from the Gα subunit [38]. Thus the GPCR is a guanine nucleotide exchange factor (GEF) for the Gα subunit, promoting GDP release and subsequent binding of GTP, which is present in a higher concentration than GDP in the cytoplasm [34]. Nucleotide exchange is accompanied by structural rearrangement of three switch regions in the Ras-like domain of the Gα subunit, resulting primarily from nucleotide-binding pocket interactions with the γ-phosphoryl group of GTP [39]. The activated Gα·GTP separates from the Gβγ dimer, and both components are free to signal through downstream effectors [34]. Mammalian Gα subtypes engage different effectors: $G\alpha_s$ activates, while $G\alpha_{i/o}$ inhibits, cyclic AMP generation by adenylyl cyclase; $G\alpha_q$ stimulates phospholipase Cβ activity and subsequent release of intracellular calcium stores; and $G\alpha_{12/13}$ signaling leads to Rho GTPase activation through RhoGEFs [34, 40]. Signaling is terminated by the intrinsic GTPase activity of the Gα subunit, leading to release of free phosphate and repeated formation of the Gαβγ heterotrimer (Figure 1.1A). Gα subunit-mediated GTP hydrolysis, and thus signal termination, is accelerated a by family of GTPase accelerating proteins (GAPs) known as 'regulators of G protein signaling' (RGS

proteins) [41]. RGS proteins do not directly contribute to the GTP hydrolysis reaction, but instead stabilize the Gα switch regions to allow for efficient hydrolysis [42]. Some Gα subunit effectors also enhance GTPase activity; particularly, phospholipase Cβ serves as a GAP for Gα_q, and the G$\alpha_{12/13}$ subfamily RGS-RhoGEF effectors possess a GTPase-accelerating domain (the rgRGS domain) with distant homology to RGS proteins [40, 43]. An additional class of Gα regulators is the GoLoco motif protein family, members of which serve as guanine nucleotide dissociation inhibitors (GDIs) by binding directly to G$\alpha \cdot$GDP and preventing nucleotide release [44].

1.4.2 Ras superfamily GTPases

The nucleotide cycle of Ras superfamily G proteins and its regulators closely parallel that of heterotrimeric G proteins. Inactive, GDP-bound Ras GTPases are activated by guanine nucleotide exchange factors (GEFs) in a process that involves structural rearrangement of two switch regions within the G protein to promote release of GDP and the Mg^{2+} cofactor (Figure 1.1B) [33, 45]. Following binding of GTP, activated Ras superfamily GTPases engage a host of downstream effectors. In contrast to heterotrimeric G proteins, the intrinsic GTPase activity of Ras superfamily members is typically very slow. Thus, Ras superfamily-specific GAPs truly 'activate' GTP hydrolysis (rather than merely accelerate hydrolysis as is the case with Gα GAPs) by contributing directly to the reaction, as typified by the "arginine finger" of p120GAP [46, 47]. In another distinct difference with Gα subunits, Ras superfamily GTPases typically possess a C-terminal cysteine residue that is isoprenylated in cells by specific lipid moiety transferases, a posttranslational modification that promotes membrane association [33]. GDIs associated with Ras superfamily GTPases slow nucleotide

exchange and utilize an isoprenyl group binding site to extract GTPases from, and shuttle them between, cellular membranes [48, 49].

1.5 HETEROTRIMERIC G PROTEIN SIGNALING IN E. histolytica

Prior to completion of the *E. histolytica* genome-sequencing project [30], indirect evidence for heterotrimeric G protein signaling components existing within *E. histolytica* accumulated in the literature, but specific genes and associated protein products had not been identified. Studies on the effects of histamine and serotonin, typical G protein coupled receptor agonists, on *E. histolytica* trophozoites revealed alterations in pathogenicity and phagocytic activity, as well as enhancement of virulence in a mouse model [50-53], suggesting the possible presence of a hormone-sensing G protein signaling pathway within *E. histolytica*. Exposure of *E. histolytica* to fibronectin fragments resulted in actin cytoskeleton rearrangements, as well as changes in intracellular calcium and cAMP levels [54-56], raising the possibility of fibronectin-responsive $G\alpha_q$, $G\alpha_s$, and/or $G\alpha_{i/o}$ signaling in trophozoites. Additional indirect evidence arose from studies utilizing cholera toxin (CTX) and pertussis toxin (PTX), factors know to ADP-ribosylate and activate $G\alpha_s$ or inhibit $G\alpha_{i/o}$ signaling, respectively. Both CTX and PTX were seen to ADP-ribosylate multiple proteins of diverse molecular weights in trophozoite lysates, and cholera toxin treatment lead to increased cAMP formation in both cytoplasmic and cell membrane preparations, as well as increased adhesion to a fibronectin-coated surface [57]. Studies in the related species *Entamoeba invadens* further suggested the possibility of heterotrimeric G protein signaling in *Entamoeba*. The catecholamines epinephrine and norepinephrine, classic GPCR agonists in mammals, were found to promote *E. invadens* encystation at high nanomolar or low micromolar concentrations, although a

traditional concentration-response pattern was not observed [58]. The authors hypothesized the presence of a β1 adrenergic receptor-like entity on trophozoite cell surfaces, as further supported by radioligand binding with a specific antagonist. Furthermore, chromatography techniques identified catecholamines within *E. histolytica* extracts, suggesting a potential autocrine G protein signaling loop [58]. Additional studies implied that CTX or PTX treatment, as well as the adenylyl cyclase-stimulating compound forskolin could also promote cAMP accumulation in and encystation of *E. invadens*, while application of an adenylyl cyclase inhibitor was reported to have opposite effects [59]. Together with epinephrine-induced binding of GTPγS on trophozoite membranes, these findings were suggestive of an adrenergic signal transduction cascade involving $G\alpha_s$- and/or $G\alpha_{i/o}$-like proteins with opposing regulatory effects on an adenylyl cyclase.

However, the sequenced *E. histolytica* genome [30], as well as those of *E. dispar* and *E. invadens*, have revealed the presence of two putative Gα subunits, a single Gβ subunit, and at least two Gγ subunits [60, 61]. Absent from the genome are clear homologs to mammalian phospholipase Cβ, as well as G protein-regulated adenylyl cyclases or cyclic nucleotide phosphodiesterase [30]. Thus, although exposure of *E. histolytica* to stimuli such as fibronectin and catecholamines may lead to cAMP accumulation or increased intracellular calcium levels, it is unlikely that these effects are mediated by traditional $G\alpha_s$/adenylyl cyclase, $G\alpha_{i/o}$/adenylyl cyclase, or $G\alpha_q$/phospholipase Cβ signaling pathways. Also, we have been unable to identify within the *E. histolytica* genome clear homologs of adrenergic, histamine, and serotonin receptors (unpublished data and [60]), suggesting that the functional effects of these biogenic amines on trophozoites may not be mediated by traditional GPCR/heterotrimeric G protein signaling.

Analysis of both the sequence and structure of the Gα subunit EhGα1 revealed a lack of homology to mammalian Gα subfamilies, including $G\alpha_s$ and $G\alpha_{i/o}$ [60]. This finding, together with a lack of the C-terminal cysteine required for ADP ribosylation by PTX, suggests that EhGα1 is unlikely to be specifically be modified by bacterial toxins [60]. The observed effects of CTX and PTX treatment on *Entamoeba* trophozoites might instead result from off-target effects, a hypothesis supported by CTX- and PTX-mediated ADP ribosylation of multiple proteins of diverse molecular weights in *E. histolytica* trophozoite lysates [57]. Despite its lack of phylogenetic relationship to any particular mammalian Gα subfamily [60], EhGα1 shares functional similarity with $G\alpha_{12/13}$ subunits in engaging and contributing to the activation of an RGS-RhoGEF effector (Figure 1.2) [62]. An evolutionary origin of *E. histolytica* heterotrimeric G protein signaling independent from but functionally convergent with that of mammalian $G\alpha_{12/13}$/RGS-RhoGEF pathways is suggested by multiple factors, including sequence divergence of EhGα1, the canonical nature of its interaction with the EhRGS-RhoGEF RGS domain (*i.e.* as opposed to the rgRGS domain found in mammalian RhoGEFs), and the structural features of the autoinhibited EhRGS-RhoGEF [60, 62]. Expression of constitutively active EhGα1 and EhRacC mutants, together with the effector EhRGS-RhoGEF, leads to Rho family GTPase activation in *Drosophila* S2 cells [62], suggesting that heterotrimeric G protein and Rho family GTPase signaling pathways communicate in *E. histolytica* (Figure 1.2). However, no specific Rho family GTPase has yet been identified as an EhRGS-RhoGEF substrate. Overexpression of either wild type EhGα1 or a dominant negative, constitutively EhGβγ-bound mutant has opposing effects on trophozoite migration, invasion, and host cell attachment and killing, suggesting that heterotrimeric G protein signaling modulates multiple pathogenesis-related behaviors [60].

Perturbation of EhGα1 expression also leads to significant changes in the *E. histolytica* transcriptome and altered the secretion of cytotoxic cysteine proteases [60], suggesting a possible functional overlap with Rab family GTPases (see below). Overexpression of EhRGS-RhoGEF has similar effects on trophozoite function when compared to the dominant negative EhGα1, consistent with its function as an EhGα1 GAP (also demonstrated *in vitro*) and, thus, a negative regulator of heterotrimeric G protein signaling in the context of its overexpression [60, 62]. Nucleotide exchange is rate-limiting in the EhGα1 nucleotide cycle [60], as seen in mammalian Gα subunits, suggesting that GEF activity is needed for signal activation. Yet the *E. histolytica* genome lacks homologs of non-receptor GEFs for heterotrimeric G proteins such as Ric-8 and GIV [30, 63], leading to the hypothesis that *E. histolytica* may express one or more GPCRs (*i.e.* a putative cell surface-spanning, EhGα1-directed GEF; Figure 1.2). Although a *bona fide* heterotrimeric G protein couple receptor has not yet been identified in this organism, one or more receptor/ligand pairs would provide valuable tools for manipulating G protein signaling in *E. histolytica* and also potentially serve as a candidate drug discovery target [36, 60].

A second, putative Gα subunit (AmoebaDB acc. no. EHI_186910) exhibits a unique domain structure, with an N-terminal Gα-like fold easily identifiable despite substantial sequence divergence from mammalian Gα subunits, and a C-terminal PP2C-related phosphatase domain [61]. The Gα-like region lacks determinants for CTX- or PTX-mediated ribosylation (as does EhGα1); furthermore, this putative Gα subunit lacks the otherwise very well-conserved nucleotide binding motifs shared among all G proteins, suggesting a lack of nucleotide binding (unpublished data and [64]). This apparently expressed protein awaits

functional assessment of its Gα-like domain and its unique relationship to the adjacent phosphatase domain.

EhGβ1 dimerized with one of two *E. histolytica* Gγ subunits when expressed in mammalian cells, and the EhGβγ dimer in turn bound EhGα1 in a nucleotide state-selective fashion [60]. Gβγ subunits also frequently engage downstream effectors, even when the associated Gα subunits lack a major known effector, as seen in the case of *Arabidopsis thaliana* sugar sensing and yeast phermone signaling [65, 66]. Signaling downstream of EhGβγ is a distinct possibility for *E. histolytica* and may contribute to the phenotypic effects of perturbed EhGα1 expression [60]; however, no EhGβγ effectors have yet been identified.

1.6   RAS SUPERFAMILY GTPases IN E. histolytica

The *E. histolytica* genome encodes a remarkably large number of small GTPases for a single-celled parasite (>170 annotated in AmoebaDB, [67]), suggesting a prominent role for Ras superfamily G protein signaling. The Ras superfamily can be divided into the Ras subfamily, typically regulating cell proliferation and survival; the Rho family that regulates actin organization, the cell cycle, and gene expression; the Ran family, implicated primarily in nucleocytoplasmic transport; and the Rab and Arf families, known as regulators of vesicular transport and trafficking (reviewed in [33]). Ten Ras proteins and two related Rap homologs have been described in *E. histolytica* [68, 69], although the complete set of Ras homolog genes has not been described since completion of the *E. histolytica* genome sequencing project. At least 20 Rho family GTPases, including Rho, Rac, and Cdc42 homologs are transcribed by *E. histolytica* trophozoites [70-72]. The Rab family is the most numerous small G protein group described in *E. histolytica*, with 91 annotated genes [73]. Although not

yet described in the literature, putative Ran and Arf family GTPases also exist in the *E. histolytica* genome [67]. While a small fraction of *E. histolytica* Ras superfamily GTPases has been investigated, the extent of functional redundancy, signaling specificity, and nucleotide cycle regulation among small G proteins remain largely unknown. Given the poor genetic tractability of *E. histolytica* trophozoites, investigations of G protein signaling in this organism have largely been limited to overexpression studies. While overexpression is certainly an informative genetic perturbation, it should be noted that overexpressed G proteins, or nucleotide cycle-impaired mutants thereof, are subject to potential mislocalization and non-physiological functions.

1.6.1 Ras family GTPases

An initial study in *E. histolytica* trophozoites identified two Ras genes and two related Rap genes, as well as a single protein that apparently cross-reacted with a mammalian anti-Ras antibody [68]. Ras family GTPases in mammals and yeast are isoprenylated with either a geranylgeranyl or a farnesyl group at the characteristic C-terminal CaaX motif, where "a" is an aliphatic amino acid and the final residue is predictive of either geranylgeranylation or farnesylation [33]. Expression of EhRap2, EhRas1, and CaaX motif mutants thereof, in mammalian reticulocytes revealed that *E. histolytica* Ras GTPases can be isoprenylated, but that their CaaX motif sequences are less predictive of the specific isoprenyl group added than mammalian counterparts [74]. An *E. histolytica* farnesyltransferase (EhFT), consisting of two subunits, was later cloned and shown to farnesylate human H-Ras and EhRas4, to the exclusion of three other *E. histolytica* Ras isoforms, indicating a distinct CaaX motif selectivity for isoprenylation [69]. Recombinant EhFT is resistant to mammalian farnesyl

transferase inhibitors, precluding their use as tools in studying Ras GTPase function in *E. histolytica* trophozoites. Ras GTPases and related signaling machinery have been the targets of much pharmaceutical development effort, given the centrality of oncogenic Ras signaling to cellular proliferation and survival in many human malignancies [75]. However, no studies of perturbed Ras signaling in *E. histolytica* have yet emerged. Similarly, putative regulators of Ras nucleotide cycling (*e.g.* GEFs and GAPs) and candidate Ras effectors are currently understudied in *E. histolytica*.

1.6.2 Rho family GTPases

Entamoeba histolytica possesses a highly dynamic, actin-rich cytoskeleton that participates in many pathogenesis-related processes (reviewed in [28]) as well as two major actin-associated myosins (reviewed in [29, 76]). Remarkably rapid actin remodeling is apparent in trophozoite motility [77], a process regulated by extracellular matrix interactions [78] as well as self-generated chemokines [79]. Cytoskeletal remodeling is also intimately associated with *E. histolytica* phagocytosis [26] and surface receptor capping [27]. As master regulators of the actin cytoskeleton, as well as cell division and transcription in mammals, Rho GTPases and their associated proteins have been a focus of intense investigation in *E. histolytica*.

The first identified Rho family GTPase in *E. histolytica* was EhRho1, also later referred to as EhRhoA1 (Figure 1.3) [70]. As a homolog of human RhoA, EhRho1 was a natural candidate substrate for the Rho-inhibiting C3 exoenzyme from *Clostridum botulinum*, a protein whose ectopic expression in *E. histolytica* trophozoites leads to ribosylation of an ~25 kDa protein and reduces both proliferation and host cell killing [80]. However, recombinant EhRho1 was later found not to be a substrate for C3 exoenzyme [81], but

instead glucosylated *in vitro* by both *C. difficile* toxin B and *C. novyi* α-toxin [82]. However, use of these two *Clostridium* toxins to study EhRho1 function *in vivo* is impaired by a lack of trophozoite membrane permeability [82]. A structural study of EhRho1 has more recently highlighted its conserved conformational difference between the GDP- and GTP-bound states, as well as its distinct lack of a "Rho insert helix" -- a structural feature that differentiates all other Rho family GTPases from the greater Ras superfamily [72]. EhRho1 also differs from its homologs at a key nucleotide-binding residue, a feature found to confer rapid intrinsic nucleotide exchange, but not constitutive activity [72, 81]. However, EhRho1 does exhibit a signature activity of other Rho family GTPases; expression of a constitutively active mutant in human cells promotes actin stress fiber formation [72]. Activated EhRho1·GTP binds a Diaphanous-related formin effector protein, EhFormin1, in contrast to multiple other *E. histolytica* Rho family GTPases (Figure 1.3) [72, 83]. EhFormin1 is known to modulate actin polymerization, to be autoinhibited by an N- to C-terminal intramolecular interaction like its well-studied mammalian homologs [84], and to be specifically activated by EhRho1·GTP [83]. A recent crystal structure of the EhRho1·GTPγS/EhFormin1 complex revealed a similar mode of interaction compared to mammalian RhoC/mDia1. However, the *E. histolytica* complex lacks a secondary binding site involving the Rho insert helix. Structure-based mutagenesis also yielded insights into specificity requirements for Rho GTPase/effector pairings [83]. EhFormin1 (also called EhDia) belongs to a family of eight *E. histolytica* formin proteins, three of which were Diaphanous-related (*i.e.* containing tandem Rho GTPase binding domains (GBDs) and formin homology three domains (FH3s)) [85, 86]. Overexpressed EhFormin1 in trophozoites localizes to pseudopodia, the microtubular assembly in the nucleus, and cytoplasmic F-actin structures in response to serum [86].

Furthermore, EhFormin1- and EhFormin2-overexpressing amoebae exhibit cell division defects, with an increased number of nuclei per cell and increased average DNA content per nucleus [86], suggesting that EhRho1/EhFormin1 signaling may be involved in actin polymerization in pseudopodia and/or trophozoite cell division (Figure 1.3).

EhRho1 has also been implicated in signaling downstream of lysophosphatidic acid (LPA), an agent that promotes actin polymerization and associated F-actin structures, alters concanavalin A (ConA)-induced surface receptor capping, increases migration and invasion, and modulates erythrophagocytosis in *E. histolytica* trophozoites [54, 87]. LPA treatment (on the order of 10 μM concentration) has been reported to promote EhRho1 activation in *E. histolytica*, as measured by a GST-Rhotekin Rho binding domain (RBD) pull-down assay [54, 87]. However, we and others have been unable to observe nucleotide-specific interaction between GST-Rhotekin RBD and either epitope-tagged EhRho1 expressed in cells or purified recombinant EhRho1 (unpublished data), suggesting that EhRho1 binding observed in other studies [54, 87] may be the result of non-specific interactions, or that the employed anti-EhRho1 antibody may cross-react with one or more other *E. histolytica* Rho family GTPases. LPA-induced EhRho1 activation has also been assessed by co-immunoprecipitation with a human antigen-derived anti-Rho kinase 2 (ROCK-2) antibody [54]; however no ROCK homologs in *E. histolytica* have yet been described or are apparent in the genome [67].

A number of Rac homologs are also expressed in *E. histolytica* [71], including EhRacA. Overexpression of a constitutively active EhRacA(G12V) in trophozoites leads to delayed cell division, as well as defects in phagocytosis of bacteria, human erythrocytes, and mucin-coated beads, and alterations in ConA-stimulated receptor capping [88]. Furthermore,

16

EhRacA was seen to specifically engage the p21-binding domain (PBD) of the p21-activated kinase EhPAK2, both in amoebic lysates and in the context of purified recombinant proteins (Figure 1.3) [89]. PAKs are effectors for canonical Rho family GTPases, and their serine/threonine kinase activities and/or localizations are modulated by binding of activated G proteins (reviewed in [90]). Trophozoites engineered to overexpress the kinase domain of EhPAK2, but not the full-length protein or the N-terminal regulatory region, exhibit defects in collagen matrix invasion, surface receptor capping, and cytokinesis [89]. Phenotypic overlap between EhRacA(G12V) and EhPAK2 kinase domain strains suggests a role for EhRacA/EhPAK2 signaling in surface receptor capping and regulation of cell division.

EhRacG has also been identified as a contributor to pathogenesis-related functions in *E. histolytica*. Overexpression of constitutively active EhRacG(G12V) in trophozoites leads to formation of a minor population of giant multinucleated cells, indicating a likely cytokinesis defect [91]. Filamentous actin arrangements and surface receptor capping are also altered with EhRacG(G12V) expression, and electron microscopy observations suggest increased budding of membrane vesicles [91]. Endogenously expressed EhRacG is enriched in ConA-induced uroids, together with filamentous actin and myosin II, consistent with its regulatory role in surface receptor capping *via* modulation of the actin cytoskeleton [91].

Activated EhRacC was recently shown to directly engage the heterotrimeric G protein effector EhRGS-RhoGEF [62]. Expression of constitutively active EhRacC, together with constitutively active EhGα1 is required to achieve EhRGS-RhoGEF activation in *Drosophila* S2 cells [62], suggesting a convergence with heterotrimeric G protein signaling. However, the contributions of EhRacC to cellular processes in *E. histolytica* remain to be directly assessed.

A number of other putative Rho family GTPase effectors have been described in *E. histolytica*, although without unequivocally associated G proteins. For instance, two other diaphanous-related formins with GBD-FH3 domain tandems are encoded by the *E. histolytica* genome, in addition to the EhRho1 effector EhFormin1 [86]. Overexpressed EhFormin2 in trophozoites, like EhFormin1, is localized in pseudopodia and pinocytic and phagocytic vesicles, and results in mitosis and cytokinesis defects [86], suggesting some functional redundancy among diaphanous-related formins despite differences in their Rho-GTPase binding sites and, thus, likely differences in Rho activator specificity ([83] and unpublished data). A fourth GBD-FH3 tandem protein, the actin-binding EhNCABP166, has also been implicated as a modulator of phagocytosis, chemotactic migration, and possibly proliferation in trophozoites [92]. The small G protein specificity of the EhNACAP166 GBD-FH3 domain tandem has been investigated; however, these binding experiments were conducted with denatured Rho GTPases [92], and intact Rho tertiary structure is required for the typical Rho/GBD-FH3 association (*e.g.* [83]). Some of the seven identified PAK family members, in addition to the EhRacA effector EhPAK2, have also been studied in *E. histolytica*. EhPAK (also called EhPAK1) localizes to pseudopods during amoebic migration and to the uroid upon ConA-induced capping [93]. The N-terminus of EhPAK1 was found to bind human Rac1 with typical nucleotide specificity (*i.e.* dependent on the GTP-bound state) despite the lack of an identifiable PBD; trophozoites overexpressing the EhPAK1 kinase domain exhibit reduced migration, an increased number of membrane extensions, and an increased rate of erythrocyte phagocytosis [94]. EhPAK3 is also expressed in trophozoites, and both recombinant protein purified under denaturing conditions and protein immunoprecipitated from amoebic lysates exhibit apparent kinase activity [95].

Putative regulators of Rho family GTPase nucleotide cycling are also prominent in the *E. histolytica* genome [30], including ~70 Dbl homology (DH) domain-containing candidate RhoGEFs, ~70 encoded RhoGAP domain-containing proteins, and a single RhoGDI. Although no studies of RhoGAP proteins have yet emerged, they are likely to regulate pathogenesis-related functions like their associated GTPases. Recombinant purified EhRhoGDI binds EhRho1 in a nucleotide state- and isoprenylation-dependent fashion [72]. As the only apparent RhoGDI, it is likely that this protein also engages other inactive Rho GTPases in *E. histolytica* to impair nucleotide exchange and regulate their subcellular localization.

Better studied are a number of Dbl family RhoGEFs. For example, overexpression of EhGEF1 in trophozoites decreases total cellular filamentous actin, reduces amoebic migration, and alters killing of mammalian cells [96, 97]. *In vitro* nucleotide exchange assays indicate that EhGEF1 likely catalyzes exchange on EhRacG and EhRho1 (the latter illustrated in Figure 1.3A), although concentrations of GEF protein employed in these assays as well as a concentration-response analysis were not included in this report [97]. Later studies utilized structural homology models to predict EhGEF1 Dbl homology domain (DH) point mutations that impaired GEF activity toward EhRho1 and EhRacG, as indicated by maximal nucleotide analog fluorescence at a single time point [98]. However, kinetic analysis is a preferable measure of GEF activity, as maximal fluorescence readings are subject to artifacts due to differing specific activities of recombinant proteins, non-specific binding, fluctuations in instrumentation settings, and/or "buffer shifts" in fluorescence that vary among protein preparations. EhGEF1 small-molecule inhibitors have also been pursued, based on a docking analysis using a homology model to existing mammalian RhoGEF

structures (~50% or less sequence similarity) [99]. Five compounds were assessed for EhGEF1 inhibition by *in vitro* nucleotide exchange assays and found to be active at ~50-100 µM concentrations [99]. However, exchange kinetics were not assessed, a typical concentration-response pattern was not obtained, and direct binding of compounds to EhGEF1 (or potentially Rho GTPases) has not yet been demonstrated in these studies. Furthermore, the specificity of these potential pharmacological tools, for instance across other *E. histolytica* RhoGEFs, remains to be determined.

The armadillo-repeat containing EhGEF2 has been implicated in erythrocyte phagocytosis, trophozoite proliferation, and chemotaxis, based upon an *E. histolytica* strain engineered to overexpress a dominant negative point mutant [100]. Both the N-terminal and DH domain regions were seen to contribute to EhGEF2 membrane localization. EhGEF2 was also suggested to activate EhRacA-D, EhRacG-H, and EhCdc42 *in vitro* [100], although no kinetic analysis was provided in this report and the fluorescence time courses shown appear to be caused by buffer shifts upon GEF addition rather than a single exponential binding event *per se*. Which Rho substrates and dominant negative mutant-impaired signals are relevant for the observed *in vivo* effects are currently unknown.

The DH-PH domain tandem of a third Dbl family RhoGEF, EhGEF3, stimulates nucleotide exchange on EhRacA and EhRho1 *in vitro* [101]. Simultaneous EhGEF3 and EhRacA overexpression in *E. histolytica* leads to increased migration toward fibronectin, while a dominant negative EhGEF3 mutant has the opposite effect. Overexpressed EhGEF3, but not the dominant negative point mutant, also promotes EhRacA activation in trophozoites, as assessed by a GST-EhPAK2 PBD pulldown assay, suggesting a role for EhGEF3/EhRacA signaling in chemotactic migration (Figure 1.3) [101]. EhGEF3 and

EhRacA co-localize in caps induced by ConA treatment, suggesting a possible additional role in surface receptor capping [101].

Members of a family of eleven RhoGEFs in *E. histolytica* each contain a FYVE domain, known to associate with inositol phospholipids and to decorate early phagosomes in trophozoites [102]. A GFP-tagged mammalian FYVE domain overexpressed in trophozoites was observed to translocate to phagocytic cups and phagosomes during host cell phagocytosis [103]. One overexpressed FYVE domain-containing RhoGEF, EhFP4, is also recruited to phagocytosis-related structures [103], and overexpression of the isolated FYVE domain from EhP4 impairs trophozoite phagocytosis. Interaction of EhFP4 or its DH-PH domains with recombinant *E. histolytica* Rho family GTPases has been assessed with pull-down assays. EhFP4 was seen to interact with EhRacC, EhRacD, and two unnamed G proteins, although nucleotide state selectivity was not assessed, and the authors reported inability to detect nucleotide exchange activity [103]. Thus, it remains to be established whether these FYVE domain-containing RhoGEFs exhibit GEF activity and whether there is any functional interplay between their FYVE and DH-PH domains.

1.6.3 Rab family GTPases

The *E. histolytica* genome encodes a remarkable 91 Rab family G proteins, many of which are not clear homologs of mammalian Rabs, suggesting an unusually high degree of complexity underlying vesicular trafficking regulation in trophozoites [73, 104, 105]. Endosomes isolated from *E. histolytica* by magnetic fractionation are associated with virulence-associated cysteine protease activity, as well as enrichment of Rab GTPases, such as EhRab11 and potentially a Rab7 homolog [106]. The importance of phagocytosis and

pinocytosis in nutrient uptake by trophozoites, the secretion of virulence factors like amoebapores and cysteine proteases, as well as the critical role of membrane-associated proteins like the Gal/GalNAc lectin, all support the hypothesis that Rab-regulated vesicular trafficking is important for *E. histolytica* biology and pathogenesis [14, 23, 24].

EhRabA is localized to vesicles at steady state, but moves to the leading edge in motile cells and the membrane opposite ConA-induced caps, as well as to membrane extensions upon N-formyl peptide-induced polarization [107]. Expression of a dominant negative EhRabA mutant in trophozoites produces changes in cell morphology and polarization, impairs motility, and reduces host cell attachment and killing, but has no observable effect on pinocytosis or erythrophagocytosis [108]. Conversely, overexpression of constitutively active EhRabA perturbs erythrophagocytosis and leads to formation of large tubular organelles apparently derived from the endoplasmic reticulum (ER) [109]. Two subunits of the Gal/Gal-NAc lectin and a cysteine protease are mislocalized to the EhRabA-induced organelles, and similar effects are seen with brefeldin A treatment, suggesting that EhRabA regulates trafficking between the ER and Golgi apparatus [109].

EhRabB is one of the first identified and the most frequently studied Rab family GTPases in *E. histolytica*. Initial immunofluorescence studies localized endogenous EhRabB to cytoplasmic vesicles and noted its translocation to the plasma membrane and phagocytic cups during erythrophagocytosis [110, 111]. Poor phagocytosis in a mutant *E. histolytica* strain correlated with increased expression of EhRabB, as well as substantial sequence differences between this mutant EhRabB and wild-type EhRabB, providing further evidence for its involvement in phagocytosis [112], although a causal association was not established. EhRabB was also observed to be enriched at phagosomes in a proteomics study [113].

Overexpression of EhRabB in trophozoites leads to a small diminution of phagocytosis, while expression of a dominant negative mutant (N118I) leads to decreases in both phagocytosis and cell monolayer destruction [114]. Of particular interest, EhRabB(N118I)-expressing trophozoites do not form liver abscesses in a hamster model, while vector-transfected and wild type EhRabB-expressing amoebae do form such abscesses, establishing EhRabB signaling as likely important for pathogenesis [114]. EhRabB was reported to interact with the transmembrane protein EhGPCR-1 by yeast two-hybrid, although binding data were not shown in this study [115]. Despite its initial naming as a G protein coupled receptor, further sequence analysis has indicated that EhGPCR-1 is more likely a Wnt-binding factor, rather than a heterotrimeric G protein GEF *per se* [60].

Entamoeba histolytica also expresses Rab5- and Rab7-related G proteins [116, 117]. Overexpressed EhRab5 and EhRab7A localize to independent vesicular structures at steady state, but exposure of these overexpression-modified trophozoites to erythrocytes is seen to cause convergence of the two Rabs at a large "pre-phagosomal vacuole" (PPV), distinct from actual phagosomes [118]. Electron microscopy studies have identified small amoebapore-containing vesicles in the PPV, suggesting a role for EhRab5 and/or EhRab7A in delivering cytotoxic amoebapores to phagosomes [118]. Consistent with this hypothesis, overexpression of wild type EhRab5 enhances phagocytosis kinetics and amoebapore transport, while expression of either constitutively active or dominant negative EhRab5 mutants impairs PPV formation and phagocytosis [118]. EhRab7 also co-localizes with early endosomes [116], and overexpression of EhRab7A in trophozoites reveals its subcellular localization to lysosomes and an increased acidic cellular compartment, as well as decreased cellular cysteine protease activity [119]. A retromer-like complex of *E. histolytica* proteins is seen to engage

recombinant EhRab7A in a nucleotide-dependent fashion, primarily through the C-terminus of EhVps26, leading to the hypothesis that EhRab7A may contribute to retrograde transport from vacuoles and phagosomes to the Golgi apparatus [119]. An *E. histolytica* homolog of Rab8 has also been cloned, but no cellular functions have yet been established for this G protein other than its vesicular localization [120, 121].

EhRab11 exhibits a punctate distribution in trophozoites and moves to the cell periphery upon iron and serum starvation of trophozoites [122], in contrast to EhRab7 and EhRabA. Iron and serum starvation is associated with altered cytokinesis and increased detergent-resistant cells, but whether EhRab11 contributes to these phenotypes is unknown [122]. A related isoform, EhRab11B, also exhibits a vesicular distribution, and overexpression of EhRab11B in trophozoites leads to an increase in both intracellular and secreted cysteine proteases [123]. Amoeba overexpressing EhRab11B exhibit slightly increased exocytosis of a fluid-phase marker and more efficiently kill mammalian cells, an effect reversed by treatment with the cysteine protease inhibitor E64 [123]. These findings suggest that the *E. histolytica* Rab11 isoforms have non-redundant functions.

No studies of Rab GTPase nucleotide cycle regulators in *E. histolytica* have yet emerged. The *E. histolytica* genome encodes for ~20 proteins with DENN domains [67], known in mammals to serve as Rab GEFs, along with other structurally unrelated proteins [124]. Also present are ~50 Rab-GAP/TBC domain containing proteins and two putative Rab GDIs [67]. Examination of these likely Rab regulators may shed further light on signaling mechanisms contributing to *E. histolytica* pathogenicity, especially in the context of vesicular trafficking mechanics.

1.7   THE UBIQUITIN-PROTEASOME SYSTEM IN E. histolytica

Among a number of genes with altered transcription upon perturbed EhGα1 expression are components of the ubiquitin-proteasome system, including a ubiquitin gene itself [60]. EhUbiquitin was previously shown to differ substantially in sequence from the otherwise highly conserved homologs in other species, although EhUbiquitin expression complements ubiquitin gene deletion in yeast [125, 126]. The likely importance of the ubiquitin-proteasome pathway in *Entamoeba* is evidenced by the impairment of trophozoite growth and encystation upon treatment with three proteasome inhibitors, lactacystin, clasto-lactacystin beta-lactone, and MG-132 [127]. Following completion of the *E. histolytica* genome sequencing [30], a complete bioinformatic analysis identified activating enzymes (E1), conjugating enzymes (E2), and ligases (E3) for ubiquitin as well as most other ubiquitin-like modifiers (ULMs) [128]. A recent study described functional ubiquitin conjugation machinery, including E1, E2, and E3 enzymes. EhUba1 (E1) catalyzes ATP-dependent activation of EhUbiquitin, transfer to the conjugating enzyme (E2) EhUbc5, and selective interaction between EhUbc5 and the E3 ligase EhRING1 [129]. Crystal structures of EhUbiquitin reveal clustering of the non-conserved residues, including an extra surface lysine, on a single surface that is not utilized in most structurally characterized ubiquitin interactions [129]. Thus, *E. histolytica* has been shown to express a functional ubiquitination enzyme cascade that likely regulates protein stability and proteasomal degradation, among other functions [130].

1.8 CONCLUSION AND PERSPECTIVE

A number of G proteins have been implicated in the pathogenic processes of *E. histolytica*, particularly heterotrimeric G proteins, a number of cytoskeleton-associated Rho GTPases, and also Rab GTPases primarily involved in vesicular trafficking. Exploitation of known signaling pathways for pharmacological manipulation is attractive, both in the development of tools for interrogating the specific functions of G protein signaling in *E. histolytica*, and as a potential approach to the development of anti-amoebiasis therapeutics. A first step has been taken in developing small molecule inhibitors for EhGEF1 [99], and other *E. histolytica* RhoGEFs may be targetable given some prior success in inhibiting mammalian Rho GTPase activation (*e.g.* [131, 132]). Some Ras and Rho GTPase effectors, particularly kinases like the PAKs and members of the mitogen-activated protein kinase (MAPK) cascade members, have also proven tractable as pharmacological targets in humans (*e.g.* [133, 134]). However, the importance of Ras effectors and downstream kinases in *E. histolytica* pathogenesis has not yet been explored. Particularly promising for pharmacological development is the recently described heterotrimeric G protein signaling within *E. histolytica* [36, 60], although identification of a *bona fide* GPCR and ligand pair in *E. histolytica* remains a barrier at this time.

Aside from pharmacological goals, much remains to be discovered regarding modulation of the Ras superfamily GTPase functions in *E. histolytica*, particularly regarding nucleotide cycle regulators (GAPs, GEFs, and GDIs). Also unclear is the interplay among the well-populated small G protein families in *E. histolytica*, such as the >20 member Rho family and the 91 member Rab family. Are many of these GTPases redundant in function and regulation? How is GTPase specificity for effectors and nucleotide cycle regulators achieved,

given such large numbers of simultaneously expressed G proteins in a single cell? Further study of both heterotrimeric and small G proteins in *E. histolytica* will likely add to our understanding of parasite biology and pathogenicity, as well as signaling in other organisms.
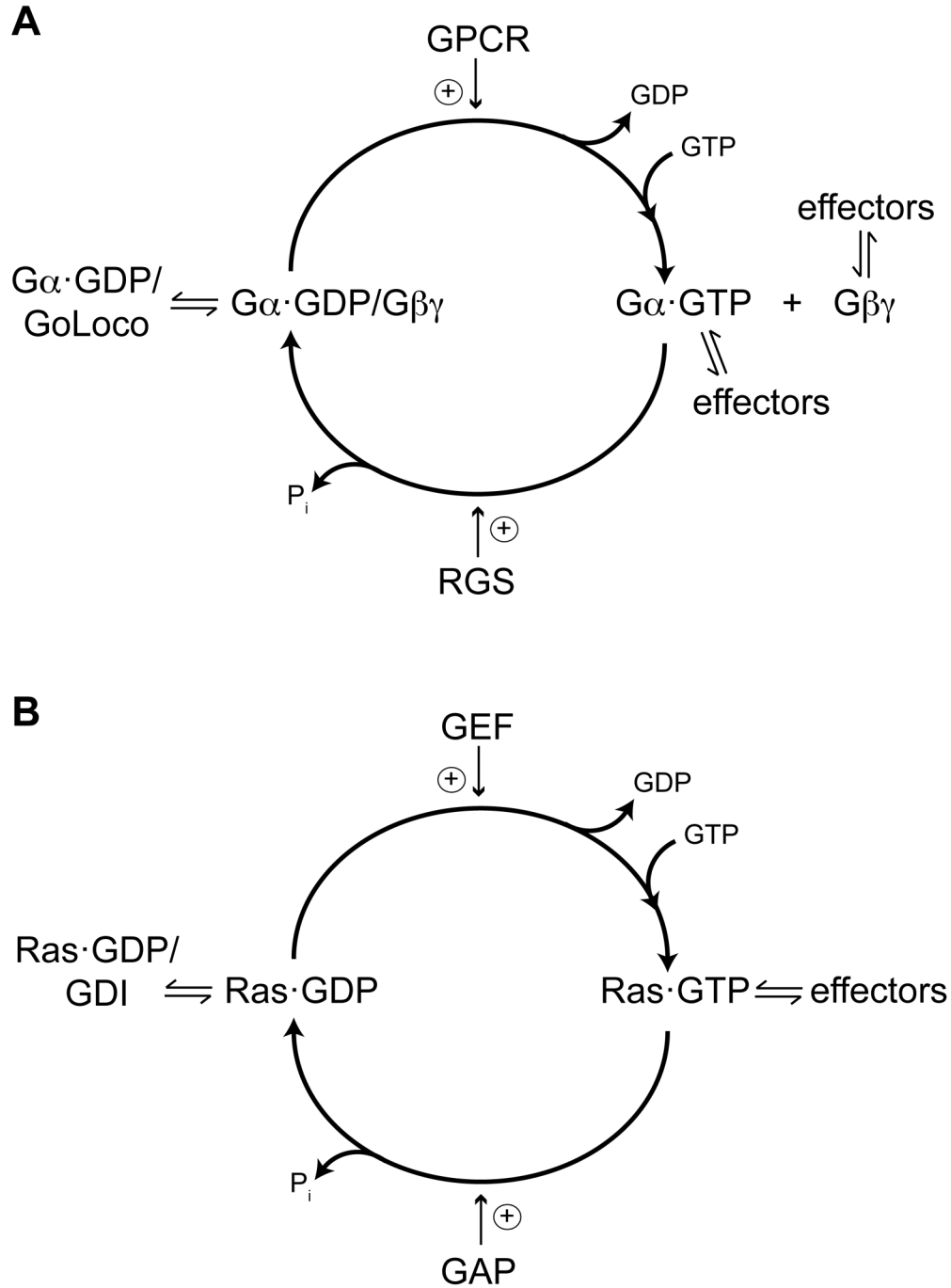
Figure 1.1. Nucleotide cycle regulation of heterotrimeric and Ras superfamily G proteins. (A) Gα subunits cycle between GDP- and GTP-bound states. GPCRs serve as GEFs for G protein heterotrimers, stimulating their release of GDP. Conversely, GoLoco motifs are GDIs that slow nucleotide exchange by Gα subunits. RGS proteins are GAPs for Gα subunits, promoting signal termination by both activated Gα subunits and free Gβγ. **(B)** The small G protein nucleotide cycle parallels that of heterotrimeric G proteins, with GEF-stimulated and GDI-inhibited nucleotide exchange as well as GAP-mediated activation of GTP hydrolysis.
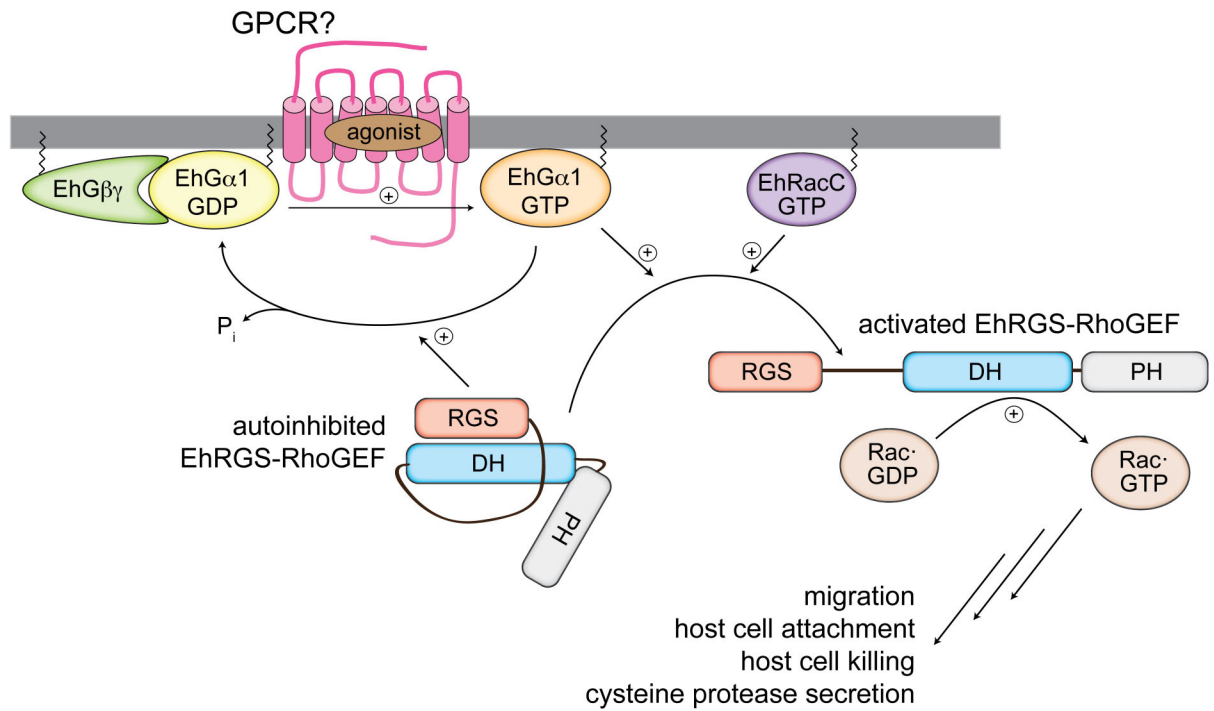
Figure 1.2. Model of heterotrimeric G protein signaling in *E. histolytica*. Activated EhGα1, together with GTP-bound EhRacC, engages the autoinhibited EhRGS-RhoGEF, leading to Rac GTPase signaling in *Drosophila* S2 cells [62], although no specific *E. histolytica* Rho family substrate for EhRGS-RhoGEF has yet been identified. Both EhGα1 and EhRGS-RhoGEF alter trophozoite migration, host cell attachment, and cell killing by altered cysteine protease secretion [60, 62]. An associated GPCR is postulated, but not yet established within this signaling pathway. Despite its name, the protein EhGPCR-1 is more likely a Wnt-binding factor than a ligand-activated heterotrimeric G protein GEF [60].
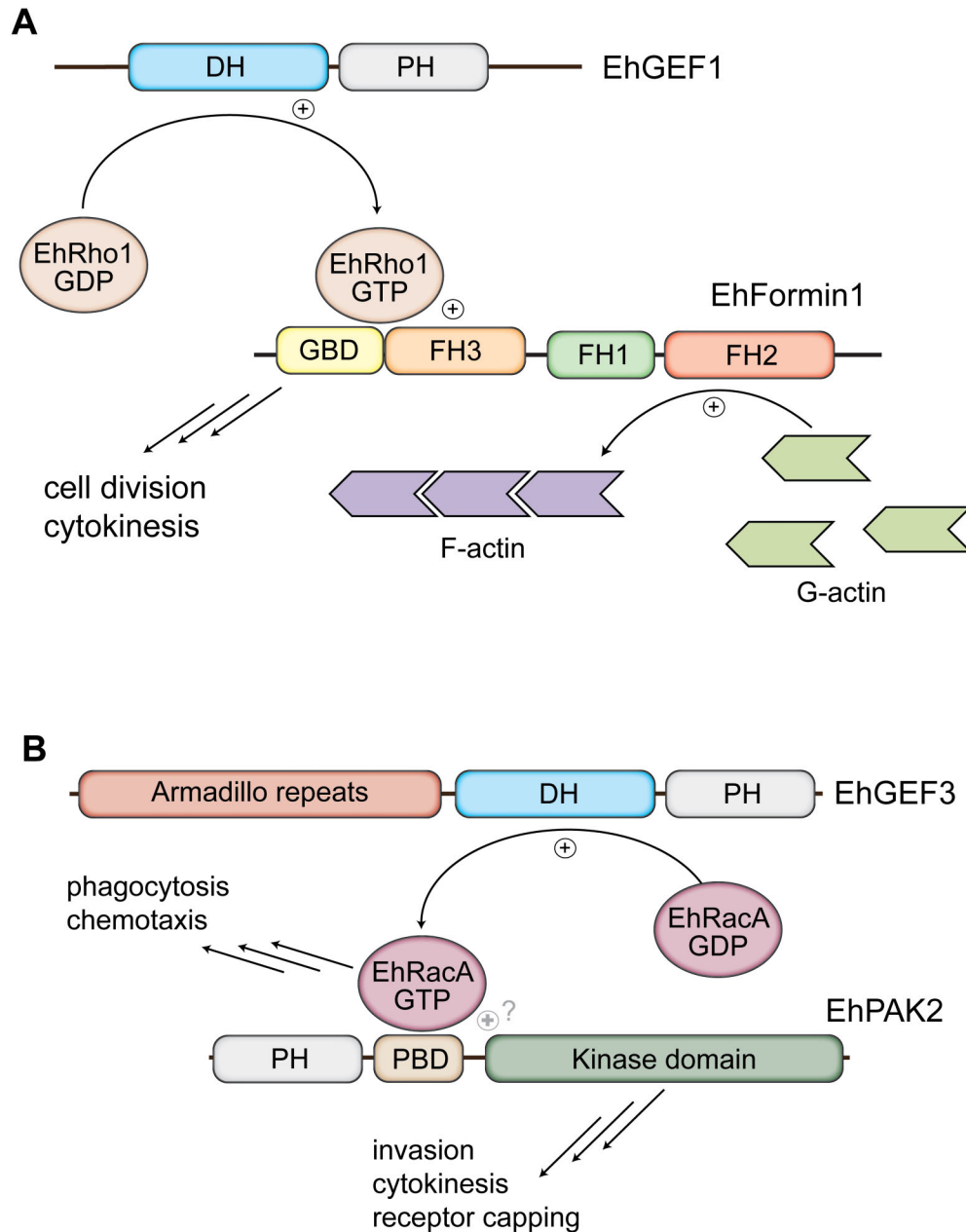
Figure 1.3. EhRho1 and EhRacA signaling modulate pathogenic behaviors in *E. histolytica*. (A) Nucleotide exchange on EhRho1 is known to be catalyzed by EhGEF1 in vitro [97]. EhRho1 engages the GBD-FH3 domain tandem of the diaphanous-related and autoinhibited EhFormin1 to modulate actin polymerization [83]. EhFormin1 has also been implicated in trophozoite proliferation and cytokinesis [86]. (B) EhRacA nucleotide exchange is known to be accelerated by EhGEF3 in vitro [101]. Constitutively active EhRacA perturbs phagocytosis and chemotaxis, as well as surface receptor capping in trophozoites [88]. EhRacA·GTP was also shown to bind EhPAK2, a likely effector whose kinase domain is implicated in collagen matrix invasion, cytokinesis and surface receptor capping [89]. Stimulation of EhPAK2 kinase activity by EhRacA is postulated, but has not yet been established.

1.9   REFERENCES

1. Ravdin, J.I., *Amoebiasis*. Vol. 2. 2000, London: Imperial College Press.

2. WHO, *WHO/PAHO/UNESCO report. A consultation with experts on amoebiasis. Mexico City, Mexico 28-29 January, 1997*. Epidemiol Bull, 1997. **18**(1): p. 13-4.

3. Haque, R., C.D. Huston, M. Hughes, E. Houpt, and W.A. Petri, Jr., *Amebiasis*. N Engl J Med, 2003. **348**(16): p. 1565-73.

4. Korpe, P.S., et al., *Evaluation of a rapid point-of-care fecal antigen detection test for Entamoeba histolytica*. Am J Trop Med Hyg, 2012. **86**(6): p. 980-1.

5. Haque, R., et al., *Diagnosis of amebic liver abscess and amebic colitis by detection of Entamoeba histolytica DNA in blood, urine, and saliva by a real-time PCR assay*. J Clin Microbiol, 2010. **48**(8): p. 2798-801.

6. Haque, R., et al., *Entamoeba histolytica infection in children and protection from subsequent amebiasis*. Infect Immun, 2006. **74**(2): p. 904-9.

7. Anuar, T.S., et al., *Molecular epidemiology of amoebiasis in Malaysia: Highlighting the different risk factors of Entamoeba histolytica and Entamoeba dispar infections among Orang Asli communities*. Int J Parasitol, 2012. **42**(13-14): p. 1165-75.

8. Ramos, F., et al., *High prevalence rate of Entamoeba histolytica asymptomatic infection in a rural Mexican community*. Am J Trop Med Hyg, 2005. **73**(1): p. 87-91.

9. Yang, B., et al., *Seroprevalence of Entamoeba histolytica infection in China*. Am J Trop Med Hyg, 2012. **87**(1): p. 97-103.

10. Verkerke, H.P., W.A. Petri, Jr., and C.S. Marie, *The dynamic interdependence of amebiasis, innate immunity, and undernutrition*. Semin Immunopathol, 2012. **34**(6): p. 771-85.

11. Hung, C.C., S.Y. Chang, and D.D. Ji, *Entamoeba histolytica infection in men who have sex with men*. Lancet Infect Dis, 2012. **12**(9): p. 729-36.

12. Nozaki, T., S. Kobayashi, T. Takeuchi, and A. Haghighi, *Diversity of clinical isolates of Entamoeba histolytica in Japan*. Arch Med Res, 2006. **37**(2): p. 277-9.

13. Barwick, R.S., et al., *Outbreak of amebiasis in Tbilisi, Republic of Georgia, 1998*. Am J Trop Med Hyg, 2002. **67**(6): p. 623-31.

14. Christy, N.C. and W.A. Petri, Jr., *Mechanisms of adherence, cytotoxicity and phagocytosis modulate the pathogenesis of Entamoeba histolytica*. Future Microbiol, 2011. **6**(12): p. 1501-19.

15. Aikat, B.K., S.R. Bhusnurmath, A.K. Pal, P.N. Chhuttani, and D.V. Datta, *The pathology and pathogenesis of fatal hepatic amoebiasis--A study based on 79 autopsy cases*. Trans R Soc Trop Med Hyg, 1979. **73**(2): p. 188-92.

16. Knobloch, J. and E. Mannweiler, *Development and persistence of antibodies to Entamoeba histolytica in patients with amebic liver abscess. Analysis of 216 cases*. Am J Trop Med Hyg, 1983. **32**(4): p. 727-32.

17. Lofmark, S., C. Edlund, and C.E. Nord, *Metronidazole is still the drug of choice for treatment of anaerobic infections*. Clin Infect Dis, 2010. **50 Suppl 1**: p. S16-23.

18. Bansal, D., N. Malla, and R.C. Mahajan, *Drug resistance in amoebiasis*. Indian J Med Res, 2006. **123**(2): p. 115-8.

19. Debnath, A., et al., *A high-throughput drug screen for Entamoeba histolytica identifies a new lead and target*. Nat Med, 2012. **18**(6): p. 956-60.

20. Wani, M.Y., et al., *Synthesis and in vitro evaluation of novel tetrazole embedded 1,3,5-trisubstituted pyrazoline derivatives as Entamoeba histolytica growth inhibitors*. Eur J Med Chem, 2012. **54**: p. 845-54.

21. Mushtaque, M., F. Avecilla, and A. Azam, *Synthesis, characterization and structure optimization of a series of thiazolidinone derivatives as Entamoeba histolytica inhibitors*. Eur J Med Chem, 2012. **55**: p. 439-48.

22. Abd Alla, M.D., et al., *Efficacy of a Gal-lectin subunit vaccine against experimental Entamoeba histolytica infection and colitis in baboons (Papio sp.)*. Vaccine, 2012. **30**(20): p. 3068-75.

23. Petri, W.A., Jr., R. Haque, and B.J. Mann, *The bittersweet interface of parasite and host: lectin-carbohydrate interactions during human invasion by the parasite Entamoeba histolytica*. Annu Rev Microbiol, 2002. **56**: p. 39-64.

24. Leippe, M. and R. Herbst, *Ancient weapons for attack and defense: the pore-forming polypeptides of pathogenic enteric and free-living amoeboid protozoa*. J Eukaryot Microbiol, 2004. **51**(5): p. 516-21.

25. Que, X. and S.L. Reed, *Cysteine proteinases and the pathogenesis of amebiasis*. Clin Microbiol Rev, 2000. **13**(2): p. 196-206.

26. Voigt, H. and N. Guillen, *New insights into the role of the cytoskeleton in phagocytosis of Entamoeba histolytica*. Cell Microbiol, 1999. **1**(3): p. 195-203.

27. Tavares, P., P. Sansonetti, and N. Guillen, *The Interplay Between Receptor Capping and Cytoskeleton Remodeling in Entamoeba histolytica*. Archives of Medical Research, 2000. **31**: p. S140-142.

28. Meza, I., P. Talamas-Rohana, and M.A. Vargas, *The cytoskeleton of Entamoeba histolytica: structure, function, and regulation by signaling pathways*. Arch Med Res, 2006. **37**(2): p. 234-43.

29. Labruyere, E. and N. Guillen, *Host tissue invasion by Entamoeba histolytica is powered by motility and phagocytosis*. Arch Med Res, 2006. **37**(2): p. 253-8.

30. Loftus, B., et al., *The genome of the protist parasite Entamoeba histolytica*. Nature, 2005. **433**(7028): p. 865-8.

31. Gilchrist, C.A., et al., *Impact of intestinal colonization and invasion on the Entamoeba histolytica transcriptome*. Mol Biochem Parasitol, 2006. **147**(2): p. 163-76.

32. Anamika, K., A. Bhattacharya, and N. Srinivasan, *Analysis of the protein kinome of Entamoeba histolytica*. Proteins, 2008. **71**(2): p. 995-1006.

33. Wennerberg, K., K.L. Rossman, and C.J. Der, *The Ras superfamily at a glance*. J Cell Sci, 2005. **118**(Pt 5): p. 843-6.

34. Oldham, W.M. and H.E. Hamm, *Heterotrimeric G protein activation by G-protein-coupled receptors*. Nat Rev Mol Cell Biol, 2008. **9**(1): p. 60-71.

35. Wennerberg, K. and C.J. Der, *Rho-family GTPases: it's not only Rac and Rho (and I like it)*. J Cell Sci, 2004. **117**(Pt 8): p. 1301-12.

36. Gilchrist, A., *Modulating G-protein-coupled receptors: from traditional pharmacology to allosterics*. Trends Pharmacol Sci, 2007. **28**(8): p. 431-7.

37. Rask-Andersen, M., M.S. Almen, and H.B. Schioth, *Trends in the exploitation of novel drug targets*. Nat Rev Drug Discov, 2011. **10**(8): p. 579-90.

38. Rasmussen, S.G., et al., *Crystal structure of the beta2 adrenergic receptor-Gs protein complex*. Nature, 2011. **477**(7366): p. 549-55.

39. Coleman, D.E., et al., *Structures of active conformations of Gi alpha 1 and the mechanism of GTP hydrolysis*. Science, 1994. **265**(5177): p. 1405-12.

40. Aittaleb, M., C.A. Boguth, and J.J. Tesmer, *Structure and function of heterotrimeric G protein-regulated Rho guanine nucleotide exchange factors*. Mol Pharmacol, 2010. **77**(2): p. 111-25.

41. Kimple, A.J., D.E. Bosch, P.M. Giguere, and D.P. Siderovski, *Regulators of G-protein signaling and their Galpha substrates: promises and challenges in their use as drug discovery targets*. Pharmacol Rev, 2011. **63**(3): p. 728-49.

42. Tesmer, J.J., D.M. Berman, A.G. Gilman, and S.R. Sprang, *Structure of RGS4 bound to AlF4--activated G(i alpha1): stabilization of the transition state for GTP hydrolysis*. Cell, 1997. **89**(2): p. 251-61.

43. Waldo, G.L., et al., *Kinetic scaffolding mediated by a phospholipase C-beta and Gq signaling complex*. Science, 2010. **330**(6006): p. 974-80.

44. Kimple, R.J., M.E. Kimple, L. Betts, J. Sondek, and D.P. Siderovski, *Structural determinants for GoLoco-induced inhibition of nucleotide release by Galpha subunits*. Nature, 2002. **416**(6883): p. 878-81.

45. Worthylake, D.K., K.L. Rossman, and J. Sondek, *Crystal structure of Rac1 in complex with the guanine nucleotide exchange region of Tiam1*. Nature, 2000. **408**(6813): p. 682-8.

46. Scheffzek, K., et al., *The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants*. Science, 1997. **277**(5324): p. 333-8.

47. Gavriljuk, K., et al., *Catalytic mechanism of a mammalian Rab{middle dot}RabGAP complex in atomic detail*. Proc Natl Acad Sci U S A, 2012. **109**(52): p. 21348-53.

48. Garcia-Mata, R., E. Boulter, and K. Burridge, *The 'invisible hand': regulation of RHO GTPases by RHOGDIs*. Nat Rev Mol Cell Biol, 2011. **12**(8): p. 493-504.

49. Scheffzek, K., I. Stephan, O.N. Jensen, D. Illenberger, and P. Gierschik, *The Rac-RhoGDI complex and the structural basis for the regulation of Rho proteins by RhoGDI*. Nat Struct Biol, 2000. **7**(2): p. 122-6.

50. Khan, M.A. and P.C. Sen, *Demonstration of histamine receptors on the surface of Entamoeba histolytica*. Indian J Exp Biol, 1989. **27**(1): p. 96-7.

51. Khan, M.A. and P.C. Sen, *Effect of histamine and serotonin on phagocytic activity of Entamoeba histolytica*. Indian J Exp Biol, 1988. **26**(9): p. 730-1.

52. Khan, M.A. and P.C. Sen, *Modulation of pathogenicity of Entamoeba histolytica by histamine*. Indian J Med Res, 1988. **88**: p. 225-7.

53. Khan, M.A., P.C. Sen, and B. Mishra, *Enhancement of virulence of Entamoeba histolytica by histamine in vitro*. Indian J Exp Biol, 1990. **28**(4): p. 376-7.

54. Rios, A., et al., *Participation of Rho, ROCK-2, and GAP activities during actin microfilament rearrangements in Entamoeba histolytica induced by fibronectin signaling*. Cell Biology International, 2008. **32**: p. 984-1000.

55. Manning-Cela, R., A. Pina, and I. Meza, *cAMP levels and up-regulation of actin mRNA in Entamoeba histolytica*. Arch Med Res, 1997. **28 Spec No**: p. 134-5.

56.   Carbajal, M.E., R. Manning-Cela, A. Pina, E. Franco, and I. Meza, *Fibronectin-induced intracellular calcium rise in Entamoeba histolytica trophozoites: effect on adhesion and the actin cytoskeleton*. Exp Parasitol, 1996. **82**(1): p. 11-20.

57.   Soid-Raggi, L.G., M.E. Torres-Marquez, and I. Meza, *Entamoeba histolytica: identification of functional Gs and Gi proteins as possible signal transduction elements in the interaction of trophozoites with fibronectin*. Exp Parasitol, 1998. **90**(3): p. 262-9.

58.   Coppi, A., S. Merali, and D. Eichinger, *The enteric parasite Entamoeba uses an autocrine catecholamine system during differentiation into the infectious cyst stage*. J Biol Chem, 2002. **277**(10): p. 8083-90.

59.   Frederick, J. and D. Eichinger, *Entamoeba invadens contains the components of a classical adrenergic signaling system*. Mol Biochem Parasitol, 2004. **137**(2): p. 339-43.

60.   Bosch, D.E., et al., *Heterotrimeric G-protein signaling is critical to pathogenic processes in Entamoeba histolytica*. PLoS Pathog, 2012. **8**(11): p. e1003040.

61.   Anantharaman, V., S. Abhiman, R.F. de Souza, and L. Aravind, *Comparative genomics uncovers novel structural and functional features of the heterotrimeric GTPase signaling system*. Gene, 2011. **475**(2): p. 63-78.

62.   Bosch, D.E., et al., *Structural Determinants of RGS-RhoGEF Signaling Critical to Entamoeba histolytica Pathogenesis*. Structure, 2013. **21**(1): p. 65-75.

63.   Garcia-Marcos, M., P. Ghosh, and M.G. Farquhar, *GIV is a nonreceptor GEF for G alpha i with a unique motif that regulates Akt signaling*. Proc Natl Acad Sci U S A, 2009. **106**(9): p. 3178-83.

64.   Wittinghofer, A. and I.R. Vetter, *Structure-function relationships of the G domain, a canonical switch motif*. Annu Rev Biochem, 2011. **80**: p. 943-71.

65.   Johnston, C.A., M.D. Willard, A.J. Kimple, D.P. Siderovski, and F.S. Willard, *A sweet cycle for Arabidopsis G-proteins: Recent discoveries and controversies in plant G-protein signal transduction*. Plant Signal Behav, 2008. **3**(12): p. 1067-76.

66.   Slessareva, J.E. and H.G. Dohlman, *G protein signaling in yeast: new components, new connections, new compartments*. Science, 2006. **314**(5804): p. 1412-3.

67.   Aurrecoechea, C., et al., *AmoebaDB and MicrosporidiaDB: functional genomic resources for Amoebozoa and Microsporidia species*. Nucleic Acids Res, 2011. **39**(Database issue): p. D612-9.

68.   Shen, P.S., A. Lohia, and J. Samuelson, *Molecular cloning of ras and rap genes from Entamoeba histolytica*. Mol Biochem Parasitol, 1994. **64**(1): p. 111-20.

69.   Kumagai, M., A. Makioka, T. Takeuchi, and T. Nozaki, *Molecular cloning and characterization of a protein farnesyltransferase from the enteric protozoan parasite Entamoeba histolytica*. J Biol Chem, 2004. **279**(3): p. 2316-23.

70.   Lohia, A. and J. Samuelson, *Molecular cloning of a rho family gene of Entamoeba histolytica*. Mol Biochem Parasitol, 1993. **58**(1): p. 177-80.

71.   Lohia, A. and J. Samuelson, *Heterogeneity of Entamoeba histolytica rac genes encoding p21rac homologues*. Gene, 1996. **173**(2): p. 205-8.

72.   Bosch, D.E., E.S. Wittchen, C. Qiu, K. Burridge, and D.P. Siderovski, *Unique structural and nucleotide exchange features of the Rho1 GTPase of Entamoeba histolytica*. J Biol Chem, 2011. **286**(45): p. 39236-46.

73.   Nakada-Tsukui, K., Y. Saito-Nakano, A. Husain, and T. Nozaki, *Conservation and function of Rab small GTPases in Entamoeba: annotation of E. invadens Rab and its use for the understanding of Entamoeba biology*. Exp Parasitol, 2010. **126**(3): p. 337-47.

74.   Shen, P.S., J.C. Sanford, and J. Samuelson, *Entamoeba histolytica: isoprenylation of p21ras and p21rap in vitro*. Exp Parasitol, 1996. **82**(1): p. 65-8.

75.   Baines, A.T., D. Xu, and C.J. Der, *Inhibition of Ras for cancer treatment: the search continues*. Future Med Chem, 2011. **3**(14): p. 1787-808.

76.   Marion, S., C. Laurent, and N. Guillen, *Signalization and cytoskeleton activity through myosin IB during the early steps of phagocytosis in Entamoeba histolytica: a proteomic approach*. Cell Microbiol, 2005. **7**(10): p. 1504-18.

77.   Maugis, B., et al., *Dynamic instability of the intracellular pressure drives bleb-based motility*. J Cell Sci, 2010. **123**(Pt 22): p. 3884-92.

78.   Meza, I., *Extracellular matrix-induced signaling in Entamoeba histolytica: its role in invasiveness*. Parasitol Today, 2000. **16**(1): p. 23-8.

79.   Zaki, M., N. Andrew, and R.H. Insall, *Entamoeba histolytica cell movement: a central role for self-generated chemokines and chemorepellents*. Proc Natl Acad Sci U S A, 2006. **103**(49): p. 18751-6.

80.   Godbold, G.D. and B.J. Mann, *Cell killing by the human parasite Entamoeba histolytica is inhibited by the Rho-inactivating C3 exoenzyme*. Molecular & Biochemical Parasitology, 2000. **108**: p. 147-151.

81.   Godbold, G.D., K.D. Corbett, and B.J. Mann, *A Rho-like small GTPase of Entamoeba histolytica contains an unusual amino acid residue in a conserved GDP-stabilization region and is not a substrate for C3 exoenzyme*. Exp Parasitol, 2002. **101**(2-3): p. 107-10.

82.   Majumder, S., G. Schmidt, A. Lohia, and K. Aktories, *EhRho1, a RhoA-Like GTPase of Entamoeba histolytica, Is Modified by Clostridial Glucosylating Cytotoxins*. Applied and Environmental Microbiology, 2006. **72**(12): p. 7842-7848.

83.   Bosch, D.E., B. Yang, and D.P. Siderovski, *Entamoeba histolytica Rho1 regulates actin polymerization through a divergent, diaphanous-related formin*. Biochemistry, 2012. **51**(44): p. 8791-801.

84.   Chesarone, M.A., A.G. DuPage, and B.L. Goode, *Unleashing formins to remodel the actin and microtubule cytoskeletons*. Nat Rev Mol Cell Biol, 2010. **11**(1): p. 62-74.

85.   Ganguly, A. and A. Lohia, *The diaphanous protein from Entamoeba histolytica controls cell motility and cytokinesis*. Arch Med Res, 2000. **31**(4 Suppl): p. S137-9.

86.   Majumder, S. and A. Lohia, *Entamoeba histolytica encodes unique formins, a subset of which regulates DNA content and cell division*. Infect Immun, 2008. **76**(6): p. 2368-78.

87.   Franco-Barraza, J., et al., *Rho Signaling in Entamoeba histolytica Modulates Actomyosin-Dependent Activities Stimulated During Invasive Behavior*. Cell Motility and the Cytoskeleton, 2006(63): p. 117-131.

88.   Ghosh, S.K. and J. Samuelson, *Involvement of p21racA, phosphoinositide 3-kinase, and vacuolar ATPase in phagocytosis of bacteria and erythrocytes by Entamoeba histolytica: suggestive evidence for coincidental evolution of amebic invasiveness*. Infect Immun, 1997. **65**(10): p. 4243-9.

89.   Arias-Romero, L.E., et al., *EhPAK2, a novel p21-activated kinase, is required for collagen invasion and capping in Entamoeba histolytica*. Mol Biochem Parasitol, 2006. **149**(1): p. 17-26.

90.   Kumar, A., et al., *PAK thread from amoeba to mammals*. J Cell Biochem, 2009. **107**(4): p. 579-85.

91.   Guillen, N., P. Boquet, and P. Sansonetti, *The small GTP-binding protein RacG regulates uroid formation in the protozoan parasite Entamoeba histolytica*. J Cell Sci, 1998. **111 (Pt 12)**: p. 1729-39.

92.   Campos-Parra, A.D., N.A. Hernandez-Cuevas, R. Hernandez-Rivas, and M. Vargas, *EhNCABP166: a nucleocytoplasmic actin-binding protein from Entamoeba histolytica*. Mol Biochem Parasitol, 2010. **172**(1): p. 19-30.

93.   Labruyere, E., V. Galy, P. Sansonetti, and N. Guillen, *Distribution of a potential p21-activated serine/threonine kinase (PAK) in Entamoeba histolytica*. Arch Med Res, 2000. **31**(4 Suppl): p. S128-30.

94.   Labruyere, E., C. Zimmer, V. Galy, J.-C. Olivo-Marin, and N. Guillen, *EhPAK, a member of the p21-activated kinase family, is involved in the control of Entamoeba histolytica migration and phagocytosis*. Journal of Cell Science, 2003. **116**: p. 61-71.

95.   Dutta, S., A. Sardar, D. Ray, and S. Raha, *Molecular and functional characterization of EhPAK3, a p21 activated kinase from Entamoeba histolytica*. Gene, 2007. **402**(1-2): p. 57-67.

96.   Santiago-Resendiz, M. and M. Vargas-Mejia, *The tail domain of Entamoeba histolytica GEF: a guanine nucleotide exchange factor*. Arch Med Res, 2000. **31**(4 Suppl): p. S301-2.

97.   Aguilar-Rojas, A., et al., *Entamoeba histolytica: inhibition of cellular functions by overexpression of EhGEF1, a novel Rho/Rac guanine nucleotide exchange factor*. Exp Parasitol, 2005. **109**(3): p. 150-62.

98.   Hernandez-Cuevas, N.A., et al., *Entamoeba histolytica EhGEF1 structure and mutational analysis: New specific residues critical for function*. Molecular & Biochemical Parasitology, 2009. **164**: p. 118-125.

99.   Hernandez-Cuevas, N.A., R. Hernandez-Rivas, A. Sosa-Peinado, A. Rojo-Dominguez, and M. Vargas, *Identification and evaluation of inhibitors of the EhGEF1 protein from Entamoeba histolytica*. J Mol Recognit, 2011. **24**(6): p. 935-44.

100.   Gonzalez De la Rosa, C.H., et al., *EhGEF2, a Dbl-RhoGEF from Entamoeba histolytica has atypical biochemical properties and participates in essential cellular processes*. Mol Biochem Parasitol, 2007. **151**(1): p. 70-80.

101.   Arias-Romero, L.E., et al., *EhGEF3, a Novel Dbl Family Member, Regulates EhRacA Activation During Chemotaxis and Capping in Entamoeba histolytica*. Cell Motility and the Cytoskeleton, 2007. **64**: p. 390-404.

102.   Powell, R.R., et al., *Entamoeba histolytica: FYVE-finger domains, phosphatidylinositol 3-phosphate biosensors, associate with phagosomes but not fluid filled endosomes*. Exp Parasitol, 2006. **112**(4): p. 221-31.

103.   Nakada-Tsukui, K., H. Okada, B.N. Mitra, and T. Nozaki, *Phosphatidylinositol-phosphates mediate cytoskeletal reorganization during phagocytosis via a unique modular protein consisting of RhoGEF/DH and FYVE domains in the parasitic protozoon Entamoeba histolytica*. Cell Microbiol, 2009. **11**(10): p. 1471-91.

104.   Saito-Nakano, Y., M. Nakazawa, Y. Shigeta, T. Takeuchi, and T. Nozaki, *Identification and characterization of genes encoding novel Rab proteins from Entamoeba histolytica*. Mol Biochem Parasitol, 2001. **116**(2): p. 219-22.

105.   Saito-Nakano, Y., B.J. Loftus, N. Hall, and T. Nozaki, *The diversity of Rab GTPases in Entamoeba histolytica*. Exp Parasitol, 2005. **110**(3): p. 244-52.

106.   Temesvari, L.A., E.N. Harris, S.L. Stanley, Jr., and J.A. Cardelli, *Early and late endosomal compartments of Entamoeba histolytica are enriched in cysteine proteases, acid phosphatase and several Ras-related Rab GTPases*. Mol Biochem Parasitol, 1999. **103**(2): p. 225-41.

107.   Welter, B.H. and L.A. Temesvari, *A unique Rab GTPase, EhRabA, of Entamoeba histolytica, localizes to the leading edge of motile cells*. Mol Biochem Parasitol, 2004. **135**(2): p. 185-95.

108.   Welter, B.H., R.R. Powell, M. Leo, C.M. Smith, and L.A. Temesvari, *A unique Rab GTPase, EhRabA, is involved in motility and polarization of Entamoeba histolytica cells*. Mol Biochem Parasitol, 2005. **140**(2): p. 161-73.

109.   Welter, B.H. and L.A. Temesvari, *Overexpression of a mutant form of EhRabA, a unique Rab GTPase of Entamoeba histolytica, alters endoplasmic reticulum morphology and localization of the Gal/GalNAc adherence lectin*. Eukaryot Cell, 2009. **8**(7): p. 1014-26.

110.   Rodriguez, M.A., et al., *An Entamoeba histolytica rab-like encoding gene and protein: function and cellular location*. Mol Biochem Parasitol, 2000. **108**(2): p. 199-206.

111.   Rodriguez, M.A. and E. Orozco, *Characterization of the EhRabB recombinant protein of Entamoeba histolytica*. Arch Med Res, 2000. **31**(4 Suppl): p. S171-2.

112.   Guzman-Medrano, R., et al., *Entamoeba histolytica: alterations in EhRabB protein in a phagocytosis deficient mutant correlate with the Entamoeba dispar RabB sequence*. Exp Parasitol, 2005. **110**(3): p. 259-64.

113.   Marion, S. and N. Guillen, *Genomic and proteomic approaches highlight phagocytosis of living and apoptotic human cells by the parasite Entamoeba histolytica*. Int J Parasitol, 2006. **36**(2): p. 131-9.

114.   Juarez-Hernandez, L.J., et al., *Entamoeba histolytica: The over expression of a mutated EhRabB protein produces a decrease of in vitro and in vivo virulence*. Exp Parasitol, 2012.

115.   Picazarri, K., J.P. Luna-Arias, E. Carrillo, E. Orozco, and M.A. Rodriguez, *Entamoeba histolytica: identification of EhGPCR-1, a novel putative G protein-coupled receptor that binds to EhRabB*. Exp Parasitol, 2005. **110**(3): p. 253-8.

116.   Welter, B.H., R.C. Laughlin, and L.A. Temesvari, *Characterization of a Rab7-like GTPase, EhRab7: a marker for the early stages of endocytosis in Entamoeba histolytica*. Mol Biochem Parasitol, 2002. **121**(2): p. 254-64.

117.   Saito-Nakano, Y., et al., *Identification and characterizaton of a Rab5 homologue in Entamoeba histolytica*. Arch Med Res, 2000. **31**(4 Suppl): p. S155-6.

118. Saito-Nakano, Y., T. Yasuda, K. Nakada-Tsukui, M. Leippe, and T. Nozaki, *Rab5-associated vacuoles play a unique role in phagocytosis of the enteric protozoan parasite Entamoeba histolytica*. J Biol Chem, 2004. **279**(47): p. 49497-507.

119. Nakada-Tsukui, K., Y. Saito-Nakano, V. Ali, and T. Nozaki, *A retromerlike complex is a novel Rab7 effector that is involved in the transport of the virulence factor cysteine protease in the enteric protozoan parasite Entamoeba histolytica*. Mol Biol Cell, 2005. **16**(11): p. 5294-303.

120. Juarez, P., R. Sanchez-Lopez, M.A. Ramos, R.P. Stock, and A. Alagon, *Rab8 as a molecular model of vesicular trafficking to investigate the latter steps of the secretory pathway in Entamoeba histolytica*. Arch Med Res, 2000. **31**(4 Suppl): p. S157-9.

121. Juarez, P., et al., *Characterization of the Ehrab8 gene, a marker of the late stages of the secretory pathway of Entamoeba histolytica*. Mol Biochem Parasitol, 2001. **116**(2): p. 223-8.

122. McGugan, G.C., Jr. and L.A. Temesvari, *Characterization of a Rab11-like GTPase, EhRab11, of Entamoeba histolytica*. Mol Biochem Parasitol, 2003. **129**(2): p. 137-46.

123. Mitra, B.N., Y. Saito-Nakano, K. Nakada-Tsukui, D. Sato, and T. Nozaki, *Rab11B small GTPase regulates secretion of cysteine proteases in the enteric protozoan parasite Entamoeba histolytica*. Cell Microbiol, 2007. **9**(9): p. 2112-25.

124. Barr, F. and D.G. Lambright, *Rab GEFs and GAPs*. Curr Opin Cell Biol, 2010. **22**(4): p. 461-70.

125. Wostmann, C., D. Liakopoulos, A. Ciechanover, and T. Bakker-Grunwald, *Characterization of ubiquitin genes and -transcripts and demonstration of a ubiquitin-conjugating system in Entamoeba histolytica*. Mol Biochem Parasitol, 1996. **82**(1): p. 81-90.

126. Wostmann, C., E. Tannich, and T. Bakker-Grunwald, *Ubiquitin of Entamoeba histolytica deviates in six amino acid residues from the consensus of all other known ubiquitins*. FEBS Lett, 1992. **308**(1): p. 54-8.

127. Makioka, A., M. Kumagai, H. Ohtomo, S. Kobayashi, and T. Takeuchi, *Effect of proteasome inhibitors on the growth, encystation, and excystation of Entamoeba histolytica and Entamoeba invadens*. Parasitol Res, 2002. **88**(5): p. 454-9.

128. Arya, S., G. Sharma, P. Gupta, and S. Tiwari, *In silico analysis of ubiquitin/ubiquitin-like modifiers and their conjugating enzymes in Entamoeba species*. Parasitol Res, 2012. **111**(1): p. 37-51.

129. Bosch, D.E. and D.P. Siderovski, *Structural determinants of ubiquitin conjugation in Entamoeba histolytica*. J Biol Chem, 2012.

130.   Schulman, B.A. and J.W. Harper, *Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways*. Nat Rev Mol Cell Biol, 2009. **10**(5): p. 319-31.

131.   Shutes, A., et al., *Specificity and mechanism of action of EHT 1864, a novel small molecule inhibitor of Rac family small GTPases*. J Biol Chem, 2007. **282**(49): p. 35666-78.

132.   Evelyn, C.R., et al., *High-throughput screening for small-molecule inhibitors of LARG-stimulated RhoA nucleotide binding via a novel fluorescence polarization assay*. J Biomol Screen, 2009. **14**(2): p. 161-72.

133.   Rusconi, P., E. Caiola, and M. Broggini, *RAS/RAF/MEK inhibitors in oncology*. Curr Med Chem, 2012. **19**(8): p. 1164-76.

134.   Zhao, Z.S. and E. Manser, *Do PAKs make good drug targets?* F1000 Biol Rep, 2010. **2**: p. 70.

CHAPTER 2

HETEROTRIMERIC G-PROTEIN SIGNALING IS CRITICAL TO PATHOGENIC PROCESSES IN Entamoeba histolytica[1]

2.1   OVERVIEW

Heterotrimeric G-protein signaling pathways are vital components of physiology, and many are amenable to pharmacologic manipulation. Here, we identify functional heterotrimeric G protein subunits in *Entamoeba histolytica*, the causative agent of amoebic colitis. The *E. histolytica* Gα subunit EhGα1 exhibits conventional nucleotide cycling properties and is seen to interact with EhGβγ dimers and a candidate effector, EhRGS-RhoGEF, in typical, nucleotide-state-selective fashions. In contrast, a crystal structure of EhGα1 highlights unique features and classification outside of conventional mammalian Gα subfamilies. *E. histolytica* trophozoites overexpressing wildtype EhGα1 in an inducible manner exhibit an enhanced ability to kill host cells that may be wholly or partially due to enhanced host cell attachment. EhGα1-overexpressing trophozoites also display enhanced transmigration across a Matrigel barrier, an effect that may result from altered baseline migration. Inducible expression of a dominant negative EhGα1 variant engenders the converse phenotypes. Transcriptomic studies reveal that modulation of pathogenesis-related trophozoite behaviors by perturbed heterotrimeric G-protein expression includes transcriptional regulation of

---

[1] Bosch, D.E., Kimple, A.J., Muller, R.E., Giguere, P.M., Machius, M., Willard, F.S., Temple, B.R., and Siderovski, D.P. (2012) Heterotrimeric G-protein Signaling Is Critical to Pathogenic Processes in *Entamoeba histolytica*. *PLoS Pathogens*. 8(11): e1003040.

virulence factors and altered trafficking of cysteine proteases. Collectively, our studies suggest that *E. histolytica* possesses a divergent heterotrimeric G-protein signaling axis that modulates key aspects of cellular processes related to the pathogenesis of this infectious organism.

2.2   INTRODUCTION

GTP-binding proteins (G-proteins) are important transducers of cellular signaling [1]. Heterotrimeric G-proteins are composed of three distinct subunits (Gα, Gβ, and Gγ) and typically coupled to seven-transmembrane domain, G-protein coupled receptors (GPCRs). Gα binds guanine nucleotide while Gβ and Gγ form an obligate heterodimer [1]. Conventionally, Gα forms a high-affinity binding site for G$\beta\gamma$ when Gα is in its inactive GDP-bound state. Activated receptor acts as a guanine nucleotide exchange factor (GEF) for Gα, releasing GDP and allowing subsequent GTP binding. The binding of GTP causes a conformational change in three flexible "switch" regions within Gα, resulting in G$\beta\gamma$ dissociation. Gα·GTP and freed G$\beta\gamma$ independently activate downstream effectors, such as adenylyl cyclases, phospholipase C isoforms, and Rho-family guanine nucleotide exchange factors (RhoGEFs) to modulate levels of intracellular second messengers [1, 2]. 'Regulator of G-protein signaling' (RGS) proteins generally serve as inhibitors of Gα-mediated signaling [3]; however, one class of RGS protein, the RGS-RhoGEFs, serve as positive "effectors" for activated Gα signal transduction [2, 4].

Heterotrimeric G-protein signaling has provided a wealth of targets amenable to pharmacologic manipulation, most prevalent being the GPCR itself [5]. Heterotrimeric G-

proteins in mammals regulate processes as diverse as vision, neurotransmission, and vascular contractility [1, 5]. Heterotrimeric G-proteins in non-mammalian organisms also exhibit a wide range of functions; for example, pheromone and nutrient sensing in yeast [6], hydrophobic surface recognition in the rice blast fungus [7], and cellular proliferation and chemical gradient sensing in the slime mold *Dictyostelium discoideum* [8, 9].

Entamoeba histolytica causes an estimated 50 million infections and 100,000 deaths per year worldwide [10]. *E. histolytica* infection is endemic in countries with poor barriers between drinking water and sewage; however, outbreaks also occur among travelers and susceptible subpopulations in developed countries [11]. Upon cyst ingestion, the amoeba may colonize the human colon. Although the majority of infections are asymptomatic (e.g. ref [12]), a fraction results in symptomatic amoebic colitis. Migratory *E. histolytica* trophozoites attach to intestinal epithelial cells through a Gal/Gal-NAc lectin [13]. Amoebae subsequently kill host cells through a number of mechanisms, including secretion of cell-perforating amoebapores [14, 15] and release of cytotoxic cysteine proteases [16].

E. histolytica has been studied for more than 50 years, and some of the signaling pathways important for pathogenesis have been identified. Several transmembrane kinases have been implicated cellular proliferation, phagocytosis, and the establishment of intestinal infection [17-19]. Calcium signaling is also involved in phagocytosis; for instance, calcium binding protein 1 (EhCaBP1) modulates the actin cytoskeleton at phagocytic cups and, together with the EhC2PK kinase, is involved in phagosome maturation [20-23]. Rho family GTPases and their activating exchange factors are also involved in a variety of pathogenic processes, including migration, phagocytosis, and surface receptor capping [24-27]. The

related Rab family small GTPases control trafficking and maturation of cellular vesicles, and are implicated in processes such as phagocytosis and cysteine protease secretion [28-31].

However, many *E. histolytica* signaling components, and thus potential targets for therapeutic intervention, remain under-studied. For example, recent sequencing of the *E. histolytica* genome identified multiple potential cell signal transduction components; e.g., 307 putative protein kinases representing all seven eukaryotic kinase families have been identified, including receptor tyrosine kinases [19, 32]. In this paper, we describe genetic, structural, and biochemical data establishing the identity of *E. histolytica* heterotrimeric G-protein signal transduction components as well as their regulatory roles in pathogenesis-related behaviors of *E. histolytica*.

2.3.   EXPERIMENTAL PROCEDURES

2.3.1 Cloning of *E. histolytica* G-protein subunits

The open reading frame (ORF) of EhGα1 was amplified from *E. histolytica* genomic DNA (Dr. M. Vargas, Center of Investigation and Advanced Studies, Mexico City) by polymerase chain-reaction (PCR) using Phusion polymerase (New England BioLabs) and Invitrogen primers. Amplicons were subcloned using ligation-independent cloning [33] into a Novagen pET vector-based prokaryotic expression construct ("pET-His-LIC-C") to form N-terminal tobacco etch virus (TEV) protease-cleavable, hexahistidine-tagged fusions. Mutations were made using QuikChange site-directed mutagenesis (Stratagene). ORFs of EhGα1, EhGβ1, EhGγ1, and EhGγ2, codon-optimized for mammalian cells, were obtained from Geneart (Regensburg, Germany); EhGα1 with an internal FLAG epitope, DYKDDDK inserted after

His-83, was also obtained for co-immunoprecipitations. Sequences for EhGγ1 and EhGγ2, identified in genomic shotgun sequences were MSQQQLTRLLQEKERLMKNFERSKNLM KVSEACSDLVNFTKSKVDPFSPEFKDSNPWDKNNEGGCCALV and MSQQQLIRLLQ EKERLMKNFERSKNLMKVSEACSELVNFTKNKIDPFSPEFKDTNPWDKSSNAGCCSL M, respectively.

2.3.2 Protein purification

Human $G\alpha_{i1}$ was purified as described [34]. For hexahistidine-tagged EhGα1 and EhRGS-RhoGEF, B834 E. coli were grown to an OD_{600nm} of 0.7−0.8 at 37°C before induction with 700 μM or 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG), respectively, for 14-16 hours at 20°C. Cell pellets were resuspended in N1 buffer (for EhGα1: 50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 10 mM $MgCl_2$, 10 mM NaF, 30 μM AlCl3, 50 μM GDP, 30 mM imidazole, 5% (w/v) glycerol; for EhRGS-RhoGEF: 50 mM HEPES pH 8.0, 150 mM NaCl, 30 mM imidazole, 5% (w/v) glycerol) and lysed at 10,000 kPa using an Avestin Emulsiflex. Cleared lysates were applied to nickel-nitrilotriacetic acid resin (GE Healthcare), washed, and eluted with N1 buffer containing 300 mM imidazole. Eluted protein was resolved using a calibrated size exclusion column (GE Healthcare) with S200 buffer for EhGα1 (50 mM Tris-HCl pH 8.0, 250 mM NaCl, 5 mM DTT, 5% (w/v) glycerol, and 50 mM GDP) or EhRGS-RhoGEF (50 mM Tris-HCl pH 8.0, 250 mM NaCl, 5 mM DTT, 5% (w/v) glycerol). Comparison of global fold between wild type and E39K mutant EhRGS-RhoGEF was performed using circular dichroism as previously described [35].

For crystallization, the flexible N-terminal helix (22 residues) was removed from EhGα1; similar alterations in human Gα aided crystallization while not perturbing the

resulting structure [34]. For anomalous dispersion, a selenomethionine derivative was produced using selenomethionine-containing minimal media (Molecular Dimensions). Selenomethionine EhGα1ΔN was purified by nickel-NTA chromatography as above, TEV protease cleavage and anion exchange chromatography, as described for Gα_{i1} [34]. Prior to gel filtration, EhGα1ΔN was subjected to reductive alkylation with a dimethylamine-borane complex, exactly as described [36]. Selenomethionine EhRGS-RhoGEF lacking the first two and last two residues (a.a. 3-517) was purified by nickel-NTA chromatography as above, with TEV protease cleavage of the affinity tag and a second nickel-NTA chromatography step prior to gel filtration.

2.3.3 Crystallization and structure determination

Crystals of lysine-methylated selenomethionine EhGα1ΔN (lacking 22 N-terminal residues) were obtained by vapor diffusion from hanging drops at 18°C. EhGα1ΔN at 15 mg/mL in crystallization buffer (50 mM HEPES pH 6.5, 150 mM NaCl, 10 mM MgCl$_2$, 10 mM NaF, 30 μM AlCl$_3$, and 50 μM GDP) was mixed 1:1 and equilibrated against crystallization solution containing 1.5 M ammonium sulfate, 175 mM K/Na tartrate, and 100 mM sodium citrate pH 5.6. Hexagonal rod crystals grew to 300 x 40 x 20 μm over 21 days. EhGα1ΔN crystals displayed the symmetry of space group P2$_1$2$_1$2$_1$ (a = 56.3 Å, b = 56.9 Å, c = 229.8 Å, α = β = γ = 90°), with two monomers in the asymmetric unit. Prior to data collection, crystals were serially transferred for ~30 seconds to well solution supplemented with 25% (v/v) glycerol at 5% increments and plunged into liquid nitrogen.

Anomalous diffraction data were obtained at 0.97954 Å wavelength (selenium absorption peak) and 100K temperature at the GM/CA-CAT ID-D beamline (APS, Argonne

47

National Labs) and processed using HKL2000 [37]. Since the EhGα1 crystals were sensitive to radiation, a highly redundant dataset was assembled by combining partial datasets collected at seven points along a single rod-shaped crystal. Heavy atom searching, experimental phasing, and automated model building used Phenix AutoSol [38]. Heavy atom searching identified 26 of 26 possible sites, and refinement yielded an estimated Bayes correlation coefficient of 59.7 to 2.8 Å resolution. After density modification, the estimated Bayes correlation coefficient increased to 59.8. ~80% of the model was constructed automatically, and the remaining portion was built manually. The current model (Table 2.1) contains two EhGα1 monomers bound to GDP; AlF_4^- could not be located in the electron density map and thus does not seem to be incorporated, despite the presence of $AlCl_3$, $MgCl_2$, and NaF (AMF) in the crystallization buffer.

Refinement was carried out against peak anomalous data with Bijvoet pairs kept separate using phenix.refine [38] interspersed with manual model revisions using the program Coot [39] and consisted of conjugate-gradient minimization and calculation of individual atomic displacement and translation/libration/screw (TLS) parameters [40]. Residues that could not be identified in the electron density were: chain A, 19-21, 221-222, 304-317, 356-358; chain B, 19-22, 188-196, 219-224, 302-313, 354-358. The model exhibits excellent geometry as determined by MolProbity [41]. A Ramachandran analysis identified 95.2% favored, 4.8% allowed, and 0% disallowed residues. A refined molecular replacement solution of diffraction data to 3.3 Å from native, methionine-containing EhGα1 crystals showed no noticeable differences at this resolution. The selenomethionine EhGα1 data, of higher quality and resolution, were used for refinement of the current model. Coordinates and structure factors are deposited in the RCSB Protein Data Bank (id 3RKA).

Crystals of EhGα1 were obtained only in the presence of GDP·AMF, but no electron density was observed for the AlF_4^-, and thus and thus our structure is representative of the inactive, GDP-bound state. Part of switch 3 in both monomers of the asymmetric unit and most of switch 2 in chain B are disordered, which is characteristic of the Gα inactive state in the absence of Gβγ binding [42, 43]. Switch 2 forms a well-ordered helix in chain A (Fig. 2.9) with Trp-196 in a solvent-exposed position, also indicative of the inactive state. This conserved tryptophan rotates into a hydrophobic pocket upon activation by either GTP or AMF (Fig. 2.4, 2.8; [34]). Interestingly, the switch 2 region of EhGα1 occupies a position intermediate between that seen for human transducin in either the GDP- or AMF-bound forms (Fig. 2.8C,D; PDB 1TAG and 1TAD). The unique conformation of switch 2 in EhGα1 is likely a result of crystal contacts; Trp-196 and the N-dimethylated Lys-195 interact with a hydrophobic patch on the neighboring molecule (Fig. 2.8B). These crystal contacts may pull switch 2 away from the nucleotide pocket, potentially allowing AlF_4^- to dissociate from EhGα1 and thus accounting for its absence in the electron density map. Another example of a GDP-only bound Gα structure despite crystallization in the presence of GDP·AMF is human Gα$_{13}$ [42]. The authors hypothesize that low-pH crystallization conditions are unfavorable for AlF_4^- binding, which is corroborated by the study presented here as our EhGα1 crystals were grown at low pH (~4.5) as well.

2.3.4 qRT-PCR of *E. histolytica* gene transcription

RNA from ~1 x 10^6 *E. histolytica* HM1:IMSS trophozoites (Dr. William Petri Jr., UVa) was isolated by TRIzol Reagent (Invitrogen) and phenol extraction as per manufacturer's instruction. RT-PCR was conducted exactly as in [44], with and without added reverse

transcriptase (RT) to control for genomic DNA contamination, since the target *E. histolytica* transcripts lacked suitable intron/exon boundaries. The number of cycles until threshold (Ct) was determined using an ABI Prism 7700. The threshold cycle number of each gene in RT-negative control reactions was subtracted from the Ct value for each transcript in RT-containing reactions, yielding ΔCt.

For timecourse validation of RNA-seq results, RNA was isolated (RNeasy, Qiagen) from uninduced trophozoites (time zero) or trophozoites exposed to 5 μg/mL of tetracycline for 6, 12, or 24 hours. RT-PCR was conducted and threshold cycles determined exactly as for detection of signaling component transcripts, above. Changes in transcript levels were determined by the $2^{-\Delta\Delta Ct}$ method as described previously [44], with the time zero average being assigned a 100% value. Each time point represents duplicate RT-PCR reactions conducted on duplicate biological samples, and statistically significant difference from time zero determined by an unpaired, two-tailed Student's t-test.

qRT-PCR primers (5'->3' direction) for quantitation of *E. histolytica* gene expression were as follows: GAPDH, forward CATATTAAGGGAGGAGCTAAGA, reverse ATGCCTCAGTGTTAACTCCA, probe F-TCAGCCCCATCTGCTGATGCACCA-Q; EhGβ1, forward TTCGTATTAGAAGTTCATGGGT, reverse GGACACAATATTATCAAGACCAC, probe F-TGACATGTGCCTATGCCCCTTCTATG-Q; EhGα1, forward CCTAAAAGTAAAGAATTTACTACAG, reverse CATCTGCCCAAAGTGCTTCA, probe F-ACCCTGTTACTCTTCCATTTTCACCAG-Q; EhGγ1, forward AACTTCGAGAGATCAAAGAACT, reverse CAGGACTGAAAGGGTCTACT, probe F-TGAAAGTAAGTGGAGCATGTTCTGAATTAG-Q; amoebapore A, forward

GGAGCAGTTGATAAAGTAACTGA, reverse CCATATGAAACAATCTTTGTGCA, probe F-CACTCTGTGCTAAAGCAGATGGTCTTG-Q; cysteine protease (EHI_006920), forward AGAAGCACTAACTCCAGTAAAG, reverse CTTTCAAGAAATCCAATAGCAGC, probe F-AGCCCAACATGTCCCTCTTGAAAATTG-Q. "F" represents fluorescein, or 5'-tetrachloro-fluorescein (TET) in 18S, and "Q" represents the quencher, TAMRA.

2.3.5 Western blotting

Trophozoites in the logarithmic growth phase were harvested and washed three times in PBS. For assessment of amoebapore A expression, trophozoites were grown in the presence or absence of 5 μg/mL tetracycline for 24 hours prior to harvesting. Trophozoites were lysed in buffer containing 20 mM Tris pH 7.5, 100 mM NaCl, 5 mM $MgCl_2$, 1 mM EDTA, 1% (v/v) NP-40 alternative (Calbiochem), 0.25% deoxycholate, and protease inhibitors (Complete Mini, Roche), and insoluble fractions were removed by centrifugation. For assessment of FLAG-EhGα1 overexpression, soluble lysates were incubated overnight with anti-FLAG affinity gel (Sigma) prior to elution in SDS buffer and protein separation by SDS-PAGE. Proteins were transferred to nitrocellulose membranes and immunoblotted with an anti-FLAG HRP conjugate (M2 monoclonal antibody, Sigma), a polyclonal anti-amoebapore A (kind gift of Dr. M. Leippe, Kiel, Germany), or an HRP-conjugated anti-β actin (cat. # sc-47778, Santa Cruz).

2.3.6 Fluorescence complementation and co-immunoprecipitation

Yellow fluorescent protein (YFP) bimolecular fluorescence complementation was performed as described [45] with modifications below. Codon-optimized ORFs of EhGγ isoforms were subcloned as HA-tagged fusions to the N-terminal 159 amino acids of YFP-venus (pcDNA3.1-YFP$_N$; Dr. Nevin Lambert, MCG). The EhGβ1 ORF was subcloned as an HA-tagged fusion with a C-terminal fragment (residues 159-239) of YFP-venus (pcDNA3.1-YFP$_C$; also obtained from Dr. Lambert, along with control YFP$_N$-human Gγ2 and YFP$_C$-human Gβ1 fusions). 200,000 COS-7 cells per well in 6-well dishes were transfected with 1 μg DNA using FuGENE-6 as per manufacturer's directions. Empty pcDNA3.1 DNA was used to maintain a constant amount of total DNA per well. Forty-eight hours post-transfection, epifluorescence was observed using an Olympus IX70 microscope with Hamamatsu monochrome CCD camera. Digital images were imported into MATLAB 2007a and quantified as previously described [45]. Pixels with greater than 40 units of intensity were considered to be fluorescent, and the percentage of positive pixels was quantified. All experiments were repeated three times. Co-immunoprecipitation was performed using the YFP-fusion proteins as previously described [45].

2.3.7 Nucleotide binding, hydrolysis, and EhGα1 activation

Spontaneous GDP release, measured by [^{35}S]GTPγS incorporation, and [γ-^{32}P]GTP hydrolysis by single turnover assays were both quantified as previously described [34]. For GTPase acceleration assays, increasing concentrations of purified EhRGS-RhoGEF were added along with the hydrolysis-initiating magnesium. Real-time monitoring of EhGα1

tryptophan fluorescence (excitation 280 nm; emission 350 nm) was conducted as described for Gα_{i1} [34].

2.3.8 Evolutionary analysis

The protein sequences of Gα subunits from humans, *S. cerevisiae*, *A. thaliana*, *D. melanogaster*, and *D. discoideum* were aligned and an unrooted phylogram derived using T-coffee [46]. Percent amino acid sequence similarities of EhGα1 and S. cerevisiae GPA1 were calculated relative to each human Gα subunit, using a multiple sequence alignment, as described previously [47]. The Gα family of *Drosophila melanogaster* served as a positive control for subfamily classification.

2.3.9 Surface plasmon resonance

Optical detection of protein binding was conducted as described previously [48]. Briefly, His$_6$-tagged EhRGS-RhoGEF was immobilized on an NTA chip surface and increasing concentrations of wildtype EhGα1 and mutants were flowed over at 10 μL/s in various nucleotide states.

2.3.10 Trophozoite stable transfection

EhGα1 and EhGα1^{S37C} were subcloned with internal FLAG epitope tags into a tetracycline-inducible expression vector, described previously [49]. Axenic cultures were transfected by lipofection as previously described [50]. Briefly, amoebae at ~5 x 10^6/mL were suspended in medium 199 (Sigma) supplemented with 5.7 mM cysteine, 1 mM ascorbic acid, 25 mM HEPES (pH 6.9), 15 μg of DNA, and 30 μL of Superfect (Qiagen). After 3 hours at 37 °C,

trophozoites were transferred to TYI-S-33 medium overnight and selected for stable transfection with 10 μg/mL hygromycin over 3 weeks.

2.3.11 Trophozoite migration and Matrigel transmigration

Trophozoite migration assays were performed essentially as described previously [51]. Briefly, amoebae were grown in the presence or absence of 5 μg/mL tetracycline for 24 hours, harvested in log growth phase, suspended in serum free TYI growth medium, and 50,000 cells loaded in the upper chamber of a Transwell migration chamber (Costar, 8 μm pore size). The lower chamber contained growth medium with or without 15% adult bovine serum. Transwell plates were incubated at 37°C for 2 hr under anaerobic conditions (GasPak EZ, BD Biosciences). Matrigel transmigration assays were performed in similar fashion, except that Matrigel was first diluted to 5 mg/mL in serum free TYI growth medium, layered on the Transwell porous filter, and allowed to gel for 6 hr prior to assay initiation. Incubation time was also extended to 16 hr to allow penetration. Migrated trophozoites attached to the lower chamber wall were detached on ice, fixed, and counted. Each experiment was performed in triplicate and statistical significance among four independent experiments was determined by an unpaired, two-tailed Student's t-test.

2.3.12 Host cell attachment

Attachment of *E. histolytica* trophozoites to epithelial monolayers was assessed as previously described [18]. Chinese hamster ovary (CHO) cells were grown to confluency in 24-well plates, washed, and fixed in 4% paraformaldehyde for 30 minutes. Trophozoites (3×10^5) grown in the presence or absence of 5 μg/mL tetracycline for 24 hours were added to the

fixed monolayers in medium 199 supplemented with 5.7 mM cysteine, 1 mM ascorbic acid, and 25 mM HEPES (pH 6.9). After incubation at 37°C for 30 minutes, each well was washed gently two times with warm PBS to remove unattached trophozoites. Monolayer-attached trophozoites were detached on ice and quantified by counting with an inverted microscope. In similar experiments, trophozoites were labeled with carboxyfluorescein diacetate succinimidyl ester (CFDA-SE). Attached fluorescent trophozoites were counted in three microscopic fields at 10X magnification. Each experiment was performed in quadruplicate and statistical significance determined by an unpaired two-tailed Student's t-test.

2.3.13 Cell killing

Killing of mammalian cells (Jurkat) was assessed using the CytoTox-ONE membrane integrity assay (Promega). In 96-well plates, 5×10^5 Jurkat cells and/or 2.5×10^4 trophozoites, grown with or without 5 μg/mL tetracycline for 24 hours, were incubated at 37°C in 200 μL of medium 199 (Sigma) supplemented with 5.7 mM cysteine, 0.5% BSA, and 25 mM HEPES pH 6.8. After 2.5 hr, 50 μL of medium from each well was incubated with Cytotox reagent and a colorimetric measure of extracellular lactate dehydrogenase activity was obtained after 10 min. 0.5% Triton X-100 was used to define 100% host cell death. Each experiment was performed with five replicates and statistical significance among three independent experiments was determined by an unpaired two-tailed Student's t-test.

2.3.14 Whole transcriptome shotgun sequencing

Total RNA from 10^6 trophozoites each of the tetracycline-induced (5 μg/mL tetracycline for 24 hours) EhGα1^{wt} and EhGα1^{S37C} strains, as well as a tetracycline-free control, was isolated

using an RNeasy Mini Kit (Qiagen) per manufacturer's instructions. Duplicate RNA purifications and sequencing were obtained for each condition.

Quality of total RNA from each sample was estimated by automated electrophoresis (Bioanalyzer, Agilent). Libraries were constructed using TruSeq RNA library preparation kits (Illumina) according to manufacturer's recommendations; molarity was estimated by analysis of DNA concentration from fluorometer detection and DNA fragment size. Prepared libraries with equal molarity were pooled and used for multiplex sequencing reactions. Libraries were sequenced using 57 cycles in a single end Illumina flowcell v.3 on a HiSeq2000 instrument (Illumina) at the UNC High Throughput Sequencing Facility. Primary data analysis and demultiplexing was performed using a standard Illumina pipeline 1.8.2.

Resulting mRNA sequence reads were mapped to the annotated *Entamoeba histolytica* genome (AmoebaDB.org) using Bowtie v0.12.7 [52]. Between 12×10^6 and 32×10^6 reads were aligned for each sample. Aligned reads were further analyzed with Cufflinks v1.3.0 [53] and visualized using the Integrative Genomics Viewer (www.broadinstitute.org/igv/). Cuffdiff was used to determine differential expression by comparing relative transcript abundances between pairs of duplicate experiments: EhGα1^{wt} expression vs. tetracycline-free control, EhGα1^{S37C} expression vs. tetracycline-free control, and EhGα1^{wt} vs. EhGα1^{S37C} expression. Genes exhibiting statistically significant differential transcription were compiled and corresponding annotations retrieved using software from Dr. Chung-Chau Hon (Institut Pasteur) [54]. Transcripts that were either up- or down-regulated in both the induced EhGα1^{wt} and EhGα1^{S37C} strains were excluded from further analysis, because of potential transcriptional modulation due to tetracycline treatment. Functions of the associated proteins were inferred from prior *E. histolytica* studies, by similarity to

mammalian protein families, or from conserved domains of known function. All encoded proteins without annotated conservation and those with domains of unknown function were classified as "unknown".

2.3.15 Cysteine protease activity

Intracellular cysteine protease activity in amoebic lysates was assayed essentially as described previously [55]. Crude extracts of 10^6 trophozoites, grown with or without 5 µg/mL tetracycline for 24 hours, were obtained by lysing with 5 cycles of freeze-thaw. Total protein concentration was quantified by Bradford's method. 2 mg of azo dye-impregnated collagen (Sigma) with 100 µg of crude extract in 500 µL of protease activation buffer (100 mM Tris pH 7.0 and 10 mM $CaCl_2$) were incubated at 37°C for 18 hr, then terminated with 500 µL of 10% TCA. Samples were centrifuged to exclude intact collagen fibers, and supernatants collected for absorbance reading at 540 nm. In parallel experiments, the inhibitor p-hydroxy-mercuribenzoic acid (PHMB) was included at 1 mM to assess the fraction of specific cysteine protease activity. Residual protease activity (after PHMB treatment) was subtracted to determine total cysteine protease activity.

Extracellular cysteine protease activity was also assayed with azo-collagen as described above. However, 10^6 trophzoites were incubated at 37 °C for 3 hr in 500 µL PBS supplemented with 20 mM cysteine, 0.15 mM $CaCl_2$, and 0.5 mM $MgCl_2$, conditions known to sustain *E. histolytica* growth and allow cysteine protease secretion [56]. Following centrifugation, the cell-free conditioned medium was assayed for cysteine protease activity as above. Statistical significance was determined by an unpaired, two-tailed Student's t-test.

2.4. RESULTS

2.4.1 Identification of *E. histolytica* heterotrimeric G-protein subunits

By a BLAST sequence similarity search with human $G\alpha_{i1}$ (E-value cutoff of 10^{-30}), we identified a single gene in *E. histolytica* encoding a putative $G\alpha$ subunit (EhGα1; AmoebaDB EHI_140350) also present in the related *E. dispar*, *E. invadens*, *E. moshkovskii*, and *E. terrapinae*. One $G\beta$ subunit was also identified (AmoebaDB EHI_000240) by sequence similarity to human $G\beta1$ (E-value cutoff of 10^{-30}), termed EhGβ1 (Fig. 2.1B). As $G\beta$ subunits form obligate heterodimers with short $G\gamma$ polypeptides or $G\gamma$-like (GGL) domains [3], we also searched for putative $G\gamma$-encoding genes. Based on sequence similarity with *S. cerevisiae* Ste18 and *D. discoideum* gpgA, together with alignment of candidate protein sequences and identification of key functional residues, we identified two putative $G\gamma$-encoding genes named EhGγ1 and EhGγ2; these two open-reading frames (in the NCBI *E. histolytica* genomic contigs AAFB02000029.1 and AAFB02000157.1, respectively) each possess a C terminal CAAX-box that specifies isoprenylation in conventional $G\gamma$ subunits [57].

To determine whether these G-protein subunits are expressed in *E. histolytica*, we amplified trophozoite mRNA using quantitative RT-PCR. Transcripts of EhGα1, EhGβ1, and EhGγ1 were all detected, along with the housekeeping gene glyceraldehyde-3-phosphate dehydrogenase (GAPDH; AmoebaDB EHI_167320) (Fig. 2.2).

2.4.2 Functional assessments of *E. histolytica* G-protein subunits

To determine whether the identified EhGα1, EhGβ1, EhGγ1, and EhGγ2 subunits form conventional heterodimeric (G$\beta\gamma$) and heterotrimeric (Gα·GDP/G$\beta\gamma$) complexes,

bimolecular fluorescence complementation and co-immunoprecipitation assays were performed (Fig. 2.3A, B). The N-terminal half of yellow fluorescent protein (YFP$_N$) was fused to EhGγ1 and EhGγ2 open reading frames while the C-terminus of YFP (YFP$_C$) was fused to EhGβ1. Only when YFP$_C$ and YFP$_N$ are fused to interacting proteins will the fluorescent protein fold and function correctly [58], as shown with the human G-protein subunits Gβ1 and Gγ2 (Fig. 2.3A). Significant cellular fluorescence was observed only when EhGα1 was co-transfected with YFP$_C$-EhGβ1 and either YFP$_N$-EhGγ1 or YFP$_N$-EhGγ2 (Fig. 2.3B). As expected, co-expression of YFP$_C$ alone with any of the YFP$_N$-fusions did not yield measurable cellular fluorescence. EhGβ1/γ1 and EhGβ1/γ2 dimers were found to interact with EhGα1 only in the presence of GDP (and not GTPγS) (Fig. 2.3C), consistent with canonical Gα·GDP/Gβγ interaction selectivity.

To determine if *E. histolytica* Gα binds and hydrolyzes GTP, EhGα1 was purified from *E. coli*. Spontaneous nucleotide exchange (as measured by [^{35}S]GTPγS binding) was determined to be 0.27 min^{-1} and 0.064 min^{-1} at 30 °C for EhGα1 and human Gα$_{i1}$ respectively (Fig. 2.4A). The observed EhGα1 exchange rate is comparable to that of Gα$_o$ [59], one of the faster spontaneous exchangers among mammalian Gα subunits. EhGα1 exhibited an intrinsic GTP hydrolysis rate of 0.21 min^{-1} at 20 °C (Fig. 2.4B), comparable to rates previously observed for human Gα$_{i1}$ and Gα$_{i3}$ under the same conditions (e.g., ref. [60]).

Trp-196 in switch 2 of EhGα1 is universally conserved among Gα subunits (e.g., Fig. 2.1A) and translocates to a hydrophobic pocket upon Gα activation − an event which is easily measured as a dramatic change of intrinsic tryptophan fluorescence in select Gα subunits that lack multiple additional tryptophan residues (e.g., ref. [34]). Exposure to the

activating reagent AlF$_4^-$ and magnesium (AMF) increases tryptophan fluorescence (Fig. 2.4C), and thus EhGα1 appears to assume a similar, activated switch conformation as conventional Gα subunits. Since the measured rates of EhGα1 nucleotide exchange (0.27 min^{-1} at 30°C) and hydrolysis (0.21 min^{-1} at 20°C) were on the same order of magnitude, we tested whether hydrolysis was rate-limiting, as seen for the *A. thaliana* Gα protein, AtGPA1 [61]. While EhGα1 assumes an activated conformation upon exposure to the non-hydrolyzable GTP analog, GppNHp, as indicated by intrinsic tryptophan fluorescence, addition of hydrolyzable GTP was insufficient to activate EhGα1 (Fig. 2.4D). Thus, nucleotide exchange is the rate-limiting step in the steady-state nucleotide cycling of EhGα1, as for mammalian Gα subunits, indicating that activation likely relies on GEF-stimulated exchange.

2.4.3 EhGα1 functional mutants

To further characterize EhGα1 activation properties and provide tools for probing G-protein function in *E. histolytica* trophozoites, we mutated presumed key residues of the nucleotide-cycling function of EhGα1. Gln-189 in switch 2 (Fig. 2.1A) is predicted to coordinate the critical nucleophilic water responsible for γ-phosphoryl group hydrolysis [62]. Mutation of this residue to leucine in mammalian Gα subunits results in inability to hydrolyze GTP even in the presence of GTPase-accelerating proteins [59]. The corresponding EhGα1(Q189L) mutation abolished the ability of EhGα1 to hydrolyze GTP (Fig. 2.4E), suggesting a conserved role for the switch 2 Gln-189 residue in orienting the nucleophilic water. The Q189L mutant also exhibited a slower rate of 0.026 min^{-1} (95% C.I., 0.021-0.031 min^{-1}) for GTPγS binding compared to wildtype (Fig. 2.4F), likely due to the slow rate of GTP

dissociation in the absence of hydrolysis. Co-immunoprecipitation experiments demonstrated that EhGα1(Q189L) did not interact with EhGβ1/γ2 dimers when cell lysates were incubated with either GDP or GTPγS (Fig. 2.5), consistent with a state of constitutive activation.

In a mutagenesis screen [63], the mammalian Gα residue corresponding to Ser-37 of EhGα1, when mutated to cysteine, was identified as constitutively binding Gβγ irrespective of whether presented with GDP or GTP analogs. We hypothesized that we could create an EhGα1 variant that constitutively binds GDP by mutating Ser-37 to cysteine. The EhGα1(S37C) mutant showed no appreciable GTPγS binding (Fig. 2.4F), consistent with dominant negative behavior due to disrupted GTP/Mg^{2+} binding. Given that the EhGα1(S37C) mutant did not bind GTP, single turnover assays were not possible with this mutant. However, EhGα1(S37C) was observed to form a heterotrimer with EhGβ1/γ2 in the presence of either GDP or GTPγS (Fig. 2.5), consistent with dominant negative character.

2.4.4 Evolutionary analysis of EhGα1 and identification of a putative effector

In an attempt to identify the Gα subunit family that EhGα1 most closely resembles, we generated a phylogenetic tree comparing Gα subunits from multiple species (Fig. 2.6A) using MEGA5 [64]. EhGα1 is only distantly related to the metazoan Gα subunits, including the Gα$_{12/13}$ subfamily that couples to RGS-RhoGEFs. EhGα1 is most similar to *D. discoideum* Gα9, a Gα subunit involved in cellular proliferation [9], although low bootstrap values in the phylogram region surrounding EhGα1 indicate uncertain topology. EhGα1 also has similarity to *A. thaliana* GPA1 and the yeast Gα subunits, GPA1 and GPA2, the latter with roles in pheromone response and nutrient sensing, respectively [6]. The *A. thaliana* GPA1 regulates diverse processes, such as transpiration and cellular proliferation in response

61

to glucose [65, 66]. We also calculated sequence similarity between EhGα1 and an array of human Gα subunits based upon multiple sequence alignments. In calibrating this method, the five known Gα subunits of *Drosophila melanogaster* showed sequence similarity patterns allowing facile classification into each of the Gα subfamilies (Gα_s, G$\alpha_{i/o}$, Gα_q, G$\alpha_{12/13}$) (Fig. 2.7A); however, both EhGα1 and GPA1 from *Saccharomyces cerevisiae* exhibited low sequence similarities to each of the human Gα subfamilies (Fig. 2.7B). EhGα1 exhibits the lowest similarity to each mammalian Gα tested, implying a likely early evolutionary departure from an ancestral Gα.

The *E. histolytica* genome was found to encode an RGS domain-containing RhoGEF (AmoebaDB EHI_010670; named EhRGS-RhoGEF) with distant homology to the RGS-RhoGEF effectors of mammalian G$\alpha_{12/13}$ subunits; no other canonical Gα effector proteins, such as adenylyl cyclases or phospholipase Cβ isoforms, were identified. The transcript encoding EhRGS-RhoGEF was detected within trophozoite mRNA using quantitative RT-PCR (Fig. 2.2). Recombinant EhRGS-RhoGEF was therefore expressed and purified from *E. coli*; as measured by surface plasmon resonance, immobilized EhRGS-RhoGEF protein was found to bind EhGα1 selectively in its GDP·AlF$_4^-$ (AMF) nucleotide state (Fig. 2.6B). This selective binding is consistent with a putative EhGα1 effector function for EhRGS-RhoGEF, yet occurs in the absence of significant homology of EhGα1 to the mammalian G$\alpha_{12/13}$ subunits that interact with mammalian RGS RhoGEFs [2, 4].

2.4.5 A crystal structure of EhGα1

To gain better insight into the distant homology of EhGα1 versus other Gα subunits, we determined a crystal structure of EhGα1 bound to GDP by single-wavelength anomalous

dispersion (SAD) using data to 2.6 Å resolution (Table 2.1; Fig. 2.8). To obtain high-quality

diffracting crystals, we modified EhGα1 by removing its extended N-terminal helix (a.a. 1-

22) and subjecting it to reductive lysine methylation. Neither alteration perturbed the

nucleotide cycle or activation kinetics of EhGα1 (Fig. 2.4B,C). EhGα1 features the highly

conserved Ras-like and all-helical domain structure and nucleotide-binding pocket

characteristic of Gα subunits (Fig. 2.9A). The three switch regions are ordered in one of the

two monomers in the asymmetric unit, likely due to crystal contacts (Fig. 2.8B). EhGα1

exhibits a highly conserved mode of nucleotide interaction, including the dispositions of

residues Ser-37 and Gln-189 (Fig. 2.9B). The guanine ring is embraced by the conserved

NKxD motif (residues 254-257; Fig. 2.10), with the hydrophobic portion of Lys-255 packing

against the planar guanine ring. The phosphate-binding loop (P-loop) forms numerous polar

contacts with the α- and β-phosphoryl groups of GDP [62].

Unique to EhGα1 is the absence of an αB helix in the all-helical domain (Fig. 2.9A).

Although the segment between αA and αC (αA-αC loop) could be affected by crystal

packing, five prolines scattered throughout this loop (positions 84, 89, 99, 103, and 106; Fig.

2.1A) suggest this region likely also lacks helical structure in solution. GoLoco motif-

containing proteins are one of the few molecules that interact with the αB helix (e.g., ref.

[34]); not surprisingly, given the lack of a structurally-conserved binding site on EhGα1, the

E. histolytica genome does not seem to encode any GoLoco motifs. EhGα1 also harbors a

unique 16-residue insert in the Ras-like domain following the α4 helix (Figs. 2.9A, 2.1A). A

portion of this insert forms a short β-strand (here termed β7) that extends the six-stranded β-

sheet common to all heterotrimeric and Ras-family GTPases [67, 68], followed by a 15-

residue loop that is disordered in our crystal structure. This region of Gα is critical for

interaction with GPCRs as seen, e.g., in the crystal structure of the β2 adrenergic receptor/G$_s$ complex [69]. Because this region is important for receptor coupling and/or specificity, the existence of this insert in EhGα1 suggests a potentially unique GPCR-coupling mechanism in *E. histolytica*, but no receptor has yet been identified (see Discussion).

2.4.6 G-protein signaling perturbation modulates trophozoite migration, Matrigel transmigration, and host cell attachment and killing

To determine roles of heterotrimeric G-protein signaling in pathogenesis-related behaviors of *E. histolytica*, HM-1:IMSS trophozoites were stably transfected with tetracycline-inducible expression plasmids [49] encoding either wildtype EhGα1 or the dominant negative EhGα1^{S37C} (Fig. 2.11A). A strain expressing the constitutively active EhGα1^{Q189L} could not be established, potentially due to cellular toxicity; however, overexpression of wildtype EhGα1 is expected to result in a moderately higher basal level of signaling to downstream components. Overexpression of signaling components is subject to limitations, including the possibility that supra-physiological expression levels and/or protein mislocalization result in toxicity or other cellular effects not typically mediated by endogenous signaling. However, this approach is useful to suggest cellular processes that may be regulated by heterotrimeric G-protein signaling and to mimic the gross perturbation that may be achieved with pharmacological agents acting on this pathway. Immunofluorescence of overexpressed EhGα1 revealed a diffuse, cytoplasmic cellular distribution that did not differ significantly between the wild type and S37C mutant strains (Fig. 2.12A). Endogenous EhGα1 was not assessed due to a current lack of specific antibodies. To assess potential effects of Gα subunit overexpression on trophzoite growth

and viability, growth curves were assessed for the parent HM-1:IMSS, EhGα1^{wt}, and EhGα1^{S37C} strains in the presence and absence of tetracycline. No significant differences in growth or viability (>90% at all time points) were observed over three days, although trophozoites expressing EhGα1^{S37C} displayed a trend toward slower growth at day 3 (Fig. 2.12B). All subsequent cellular experiments were conducted following growth with or without tetracycline for 24 hours.

Trophozoite motility is related to the pathogenesis of amoebic colitis, likely contributing to tissue invasion [70, 71]. Tetracycline-induced EhGα1^{wt} overexpression increased migration in the absence of a serum stimulus while EhGα1^{S37C} expression reduced migration in the presence or absence of serum in Transwell migration assays (Fig. 2.11B), suggesting that perturbation of heterotrimeric G-protein signaling may regulate motility at baseline and potentially in response to serum factor stimuli. However, the reduced migration of the EhGα1^{S37C} strain in the presence of serum may be due to the lower baseline trophozoite motility, as observed in the absence of a serum stimulus, rather than due to specific heterotrimeric G-protein involvement in a signaling response to serum factors. Tetracycline treatment had no measurable effect on the migration of the HM-1:IMSS parent strain or trophozoites transfected with an empty expression vector (Fig. 2.13A).

E. histolytica invades the intestinal mucosa, giving rise to ulcers and, in rare cases, systemic amoebiasis [72, 73]. To assess migration across a barrier, transfected trophozoite strains were profiled by a Transwell assay, with upper and lower chambers separated by Matrigel. Induced expression of EhGα1^{wt} enhanced, but EhGα1^{S37C} reduced, Matrigel transmigration relative to uninduced controls (Fig. 2.11C), revealing a potential regulatory role for heterotrimeric G-protein signaling. Tetracycline treatment had no effect on the

transmigration of HM-1:IMSS or empty vector-transfected trophozoites (Fig. 2.13B). The effects of EhGα1^{wt} and EhGα1^{S37C} overexpression on Matrigel transmigration displayed the same trends seen for migration in the absence of serum (Fig. 2.11B). Thus, differential baseline migration rates may account for part or all of the observed differences in Matrigel transmigration.

E. histolytica trophozoites also attach to and kill host cells, including intestinal epithelium and responding immune cells. Host cell attachment, achieved primarily through a galactose-inhibitable lectin [13, 74], is required for subsequent cell killing. Trophozoites expressing EhGα1^{wt} displayed greater attachment to CHO cell monolayers than uninduced controls, and the opposite effect was seen in the EhGα1^{S37C} strain (Fig. 2.14A, 2.15). EhGα1^{wt} overexpression enhanced Jurkat cell killing, as assessed with a membrane integrity assay, while trophozoites expressing the dominant negative EhGα1^{S37C} were less cytotoxic (Fig. 2.14B). Tetracycline treatment had no effect on host cell attachment or killing by HM-1:IMSS or empty vector-transfected trophozoites (Fig. 2.13C, D). Thus, perturbation of heterotrimeric G-protein signaling also regulates host cell killing by E. histolytica. Similar patterns were observed in host cell attachment and cell killing assays; different degrees of attachment upon expression of EhGα1^{wt} or EhGα1^{S37C} may be partially or wholly responsible for the observed changes in contact-dependent cell killing.

2.4.7 Regulation of transcription by perturbed heterotrimeric G-protein signaling

To gain insight into potential mechanisms by which perturbation of EhGα1 expression controls pathogenesis-related behaviors in E. histolytica, RNA-seq was performed on mRNA isolated from trophozoites expressing EhGα1^{wt}, EhGα1^{S37C}, and uninduced

controls. To emphasize highly transcribed genes and eliminate potential transcriptional effects of tetracycline treatment, transcripts with a Fragments Per Kilobase of exon per Million fragments mapped (FPKM) value less than 10 and transcripts that were up- or down-regulated (in the same direction) in both EhGα1^{wt} and EhGα1^{S37C} samples (24 hour tetracycline treatment at 5 μg/mL) relative to uninduced (tetracycline-free) trophozoites were excluded. Twenty-one genes were differentially transcribed in opposite directions upon expression of either EhGα1^{wt} or EhGα1^{S37C} (Fig. 2.16A). Transcriptional changes of multiple genes were verified over a 24-hour time course by RT-PCR (Fig. 2.17). For instance, EhGβ1 was found to be more highly expressed in trophozoites expressing EhGα1^{S37C}. Analysis of putative functions for the differentially transcribed genes revealed a diversity of responses to altered heterotrimeric G-protein signaling (Fig. 2.16). Stress response-related transcripts, such as those encoding heat shock proteins, were exclusively down-regulated upon EhGα1^{wt} expression and up-regulated in the dominant negative EhGα1^{S37C} strain; conversely, numerous metabolic enzymes were selectively up-regulated following expression of EhGα1^{S37C}, suggesting that heterotrimeric G protein signaling may be involved in sensing and responding to vital extracellular nutrients.

Genes with known effects on *E. histolytica* pathogenesis were also differentially transcribed, as measured by RNA-seq. (Table 2.2). For example, the host cell lytic factor amoebapore C was up-regulated upon EhGα1wt expression, while the amoebapore A precursor was down-regulated by EhGα1^{S37C} (Table 2.2), consistent with the higher or lower cell killing efficiencies, respectively, of each strain (Fig. 2.14B) [14, 15, 70]. Down-regulation of amoebapore A upon expression of EhGα1^{S37C} was confirmed by RT-PCR at the transcriptional level, and by western blot at the protein level (Fig. 2.17; anti-amoebapore A

was a gift from Dr. M. Leippe, U. of Kiel, Germany). A number of cysteine proteases, known factors in both host cell killing and Matrigel transmigration [56], were differentially transcribed following expression of EhGα1^{S37C} (Table 2.1). The down-regulation of one cysteine protease (EHI_006920) was confirmed by RT-PCR (Fig. 2.17). Ten Rab family GTPases, known to regulate vesicular trafficking and cysteine protease secretion [31], as well as other putative secretion/trafficking proteins, were also differentially transcribed. Specifically, four cysteine protease binding factors (CBPFs), recently shown to modulate cysteine protease secretion [75], were down-regulated in trophozoites expressing EhGα1^{S37C} (Table 2.1). These transcriptional effects suggested that altered cysteine protease activity and/or secretion may be a mechanism by which perturbation of heterotrimeric G-protein signaling modulates Matrigel transmigration and host cell killing (Figs. 2.11C & 2.14B). To test this hypothesis, intracellular and secreted cysteine protease activities were each measured in the EhGα1^{wt} and EhGα1^{S37C} strains. EhGα1^{wt} expression increased extracellular and decreased intracellular cysteine protease activity, likely reflecting more efficient vesicular trafficking and secretion (Fig. 2.16C). In contrast, EhGα1^{S37C} expression resulted in a trend toward more intracellular protease activity, although not statistically significant (p = 0.07), and significantly less extracellular protease activity relative to uninduced control trophozoites, correlating with reduced Matrigel transmigration and cell killing by this strain (Figs. 2.11C & 2.14B).

2.5. DISCUSSION

Here we demonstrate that functional heterotrimeric G-protein subunits are encoded by the pathogen *Entamoeba histolytica*, including single Gα and Gβ subunits, and two Gγ subunits. Like their mammalian counterparts, EhGα1, EhGβ1, and EhGγ1/2 form a nucleotide state-dependent heterotrimer. EhGα1 binds and hydrolyzes GTP and its switch regions undergo a conserved conformational change. When in an activated state, EhGα1 is seen to engage a putative effector protein, namely an RGS domain-containing RhoGEF (EhRGS-RhoGEF). EhRGS-RhoGEF likely represents a functional signaling link between heterotrimeric G-proteins and Rho family GTPases in *E. histolytica*. Indeed, Rho GTPases and other Dbl family RhoGEFs in *E. histolytica* have been implicated in multiple processes important for pathogenesis-related processes such as actin reorganization during chemotaxis, surface receptor capping, cell killing, phagocytosis, and tissue destruction [24-27, 76].

The sequence of EhGα1 diverges from each of the mammalian Gα subunit subfamilies, including the $Gα_{12/13}$ subfamily that couples to RGS-RhoGEFs. Thus EhGα1 likely represents an early evolutionary departure from the metazoan Gα/RGS-RhoGEF signaling axis, or possibly a signaling pathway of similar function with an independent evolutionary origin. A search of publicly available genome sequences using SMART [77] identified the RGS and DH-PH domain combinations exclusively in metazoan species, with the only exception being the amoebazoans. This lack of RGS-RhoGEF related proteins in non-metazoan species suggests an independent origin of the *E. histolytica* Gα/RGS-RhoGEF interaction; however, we cannot rule out the possibility that a Gα/RGS-RhoGEF interaction arose early in evolutionary history, such as an ancestral Unikonta supergroup member (e.g.

69

[78]), and was later lost in fungal species, but retained in metazoans and amoebae. Among the species compared in this study, EhGα1 was found to be most similar in sequence to the *D. discoideum* Gα9, followed more distantly by *S. cerevisiae* GPA1 and GPA2, as well as *A. thaliana* GPA1. This set of Gα subunits is only loosely related by function, with *D. discoideum* Gα9 regulating cellular proliferation [9], while yeast GPA1 and GPA2 transduce signals in response to pheromones and nutrients, respectively [6]. A variety of downstream signaling machinery is utilized as well, with *S. cerevisiae* pheromone signaling occurring predominantly through Gβγ subunit effectors, while S.c. GPA2 engages an adenylyl cylase effector [6]. The current study clearly differentiates EhGα1 from these relatively similar Gα subunits on the sequence level, demonstrating interaction with an RGS-RhoGEF effector and no significant effect on cellular proliferation, but apparent roles in multiple pathogenesis-related processes of *E. histolytica*.

Perturbation of heterotrimeric G-protein signaling in *E. histolytica* trophozoites was observed to modulate migration, Matrigel transmigration, and host cell attachment and killing. Notably, trophozoite Matrigel transmigration is dependent on general migration to some degree, and host cell killing is dependent on attachment. Thus, the effects of heterotrimeric G-protein perturbation on Matrigel transmigration and host cell killing may be partially or wholly due to the alterations in migration and attachment, respectively. Induced expression of the dominant negative EhGα1^{S37C} impaired these pathogenic processes, suggesting that antagonizing G-protein signaling may reduce *E. histolytica* virulence. The complete mechanisms by which heterotrimeric G-proteins are linked to specific trophozoite behaviors remain to be elucidated. For instance, it is presently unclear which signaling cascades are utilized to effect transcriptional changes in response to perturbed EhGα1

expression. EhGα1 likely engages its RGS-RhoGEF effector, leading to activation of specific Rho GTPases, some of which are known to regulate cytoskeletal dynamics required for such processes as migration and Matrigel transmigration [24, 27, 76, 79]. EhGβγ may also engage as yet unidentified effectors, like its homologs in other species, leading to changes in pathogenic processes [1].

It is presently unclear how heterotrimeric G-protein signaling is activated in *E. histolytica*. Since nucleotide exchange is the rate-limiting step in the nucleotide cycle of EhGα1, an exchange factor, such as a GPCR, is likely required for high levels of EhGα1 activation. At this time, the only putative GPCR described is the Rab GTPase-binding protein EhGPCR-1 [80]. While it would be compelling to demonstrate receptor-mediated nucleotide exchange on EhGα1, our own bioinformatic analysis revealed that EhGPCR-1, while containing seven-transmembrane spanning regions, is more likely a conserved Wnt-binding factor required for Wnt secretion (as seen in C. elegans) [81]. Identification of a *bona fide* GPCR/ligand pair or other heterotrimeric G-protein activation mechanism in *E. histolytica* will provide powerful tools for further probing of the roles of heterotrimeric G-protein signaling in trophozoites.

A

Structural elements: β1 — P-loop — α1

Eh Gα1	MG--------------------CGQSKDNS-----IDKQIVEMAK--SKPITVMLLGSGESGKSTIAK
Dd Gα1	MG-----------NICGKPELGSPEEIKANQH-----INSLLKQARSKLEGEIKLLLLGAGESGKSTIAK
Sc GPA1	MG--------CTVSTQTIGDESDPFLQNKRANDVIEQSLQLEKQRDKNEIKLLLLGAGESGKSTVLK
Hs Gαi1	MG-----------CTLS----AEDKAAVERSK----MIDRNLREDGEKAAREVKLLLLGAGESGKSTIVK
Hs Gα13	MADFLPSRSVLSVCFPGCLLTSGEAEQQRKSK----EIDKCLSREKTYVKRLVKILLLGAGESGKSTFLK

Structural elements: αA

Eh Gα1	QLKILFGGGFPEQERATHKSSICSNVVTCMRTLIEQSAILNHPMKYQPKSEFTTEDP------------
Dd Gα1	QMKIIHLNGFNDEEKSSYKTIIYNNTVGSMRVLVNAAEELKIGISENNKEAASRISN------------
Sc GPA1	QLKLLHQGGFSHQERLQYAQVIWADAIQSMKILIIQARKLGIQLDCDDPI[insert]QEGDDQMFVRKTSRE
Hs Gαi1	QMKIIHEAGYSEEECKQYKAVVYSNTIQSIIAIIRAMGRLKIDFG-DSARADDAR----QLFV-----
Hs Gα13	QMRIIHGQDFDQRAREEFRPTIYSNVIKGMRVLVDAREKLHIPWGDNSNQQHGDKMMS----FDTRAPMA

Structural elements: αC — αD — αE — αF

Eh Gα1	----VTLPFSPELVGDVEALWA-DEGIQATYEESAKFQLPDCAKYLFENVKRIAMEDYVPTEEDLIHNR
Dd Gα1	DLGDHFNGVLTAELAQDIKALWA-DPGIQNTFQRSSEFQLNDSAAYYFDSIDRISQPLYLPSENDVLRSR
Sc GPA1	IQGGNRRNLIHEDIAKAIKQLWNNDKGIKQCFARSNEFQLEGSAAYYFDNIEKFASPNYVCTDEDILKGR
Hs Gαi1	LAGAAEEGFMTAELAGVIKRLWK-DSGVQACFNRSREYQLNDSAAYYLNDLDRIAQPNYIPTQQDVLRTR
Hs Gα13	AQGMVETRVFLQYLP-AIRALWA-DSGIQNAYDRRREFQLGESVKYFLDNLDKLGEPDYIPSQQDILLAR

Structural elements: Switch I — β2 — β3 — α2 Switch II — β4 — Switch III — α3

Eh Gα1	TKTTGIIHEYDFVVKDIPFHLIDVGGQRSERKKWVSFFSDVDCAIFVTSLAEYDMKLYEDGNTSRLTESIA
Dd Gα1	TKTTGIIETVFEIQNSTFRMVDVGGQRSERKKWMHCFQEVTAVIFCVALSEYDLKLYEDDTTNRMQESLK
Sc GPA1	IKTTGITETEFNIGSSKFKVLDAGGQRSERKKWIHCFEGITAVLFVLAMSEYDQMLFEDERVNRMHESIM
Hs Gαi1	VKTTGIVETHFTFKDLHFKMFDVGGQRSERKKWIHCFEGVTAIIFCVALSDYDLVLAEDEEMNRMHESMK
Hs Gα13	RPTKGIHEYDFEIKNVPFKMVDVGGQRSERKRWFECFDSVTSILFLVSSSEFDQVLMEDRLTNRLTESLN

Structural elements: β5 — αG — α4 — β7

Eh Gα1	VFKDIMTNEFLKGAVKLIFLNKMDLFEEKLTKVPLNTIFPEYTG-GDNAVMGAQYIQQLFTGKLQTEEMN
Dd Gα1	LFKEICNTKWFANTAMILFLNKRDIFSEKITKTPINTVCFKEYDG-PQTYEGCSEFIKQQFINQNENP-K-
Sc GPA1	LFDTLLNSKWFKDTPFILFLNKIDLFEEKVKSMPIRKYFPDYQGRVGDAEAGLKYFEKIFLSLNKTN---
Hs Gαi1	LFDSICNNKWFTDTSIILFLNKKDLFEEKIKKSPLTICYPEYAG-SNTYEEAAAYIQCQFEDLNKRKDT-
Hs Gα13	IFETIVNNRVFSNVSIILFLNKTDLLEEKVQIVSIKDYFLEFEGDPHCLRDVQKFLVECFRNKRRDQQQ-

Structural elements: β6 — α5

Eh Gα1	ISGADGTANIEGAVNEKVYTNPTNATDGSNIKRVFMLAVDVIMKNMAANGKMRPTK
Dd Gα1	---------------KSIYPHLTCATDTNNILVVFNAVKDIVLNLTLGEAGMIL--
Sc GPA1	---------------KPIYVKRTCATDTQTMKFVLSAVTDLIIQQNLKKIGII---
Hs Gαi1	---------------KEIYTHFTCATDTKNVQFVFDAVTDVIIKNNLKDCGLF---
Hs Gα13	---------------KPLYHHFTTAINTENIRLVFRDVKDTILHDNLKQLMLQ---

B

Eh Gβ1	MAEDLQAKIIQTRKEIDQIKEETRANREAKEDTTMKKFTSEIPGISRSD-HVKFKRGLRGHISKIYAMHW
Dd Gβ	MSSDISEKIQQARRDAESMKEQIRANRDVMNDTTLKTFTTRDLPGLPKMEGKIKVRRNLKGHLAKIYAMHW
Sc STE4	MS-----ELDQLRQEAEQLKNQIRDARKACADATLSQITNNIDPVGRIQ---MRTRRTLRGHLANKISDFRW
Hs Gβ1	MS-----ELDQLRQEAEQLKNQIRDARKACADATLSQITNNIDPVGRIQ---MRTRRTLRGHLAKIYAMHW

Eh Gβ1	ASDSTHLVSASQDGLMLWDGQTTNKIHCIRIRSSWVMTCAYAPSMKFVACGGLDNIVSIYYSIAPKSSSS
Dd Gβ	AEDNVHLVSASQDGKLLVWDGLTTNKVHAIPLRSSWVMTCAYSPTANFVACGGLDNICSIYNLR----SR
Sc STE4	SRDSKRILSASQDGFMLIWDSASGLKQNAIPLDSQWVLSCAISPSSTLVASAGLNNNCTIYRVS---KEN
Hs Gβ1	GTDSRLLVSASQDGKLIIWDSYTTNKVHAIPLRSSWVMTCAYAPSGNYVACGGLDNICSIYNLK----TR

Eh Gβ1	EPLNNTTAELAGHVGYISCMRYVDENRILTSSGDSTCCLWDVEQSAKIMDFKDHQADVMCVSVSPD----
Dd Gβ	EQPIRVCRELNSHTGYLSCCRFLNDRQIVTSSGDMTCILWDVENGTKITEFSDHNGDVMSVSVSPD----
Sc STE4	RVAQNVASIFKGHTCYISDIEFTDNAHIITASGDMTCALWDIPKAKRVREYSDHLGDVLALAIPEEPNSE
Hs Gβ1	EGNVRVSRELAGHTGYLSCCRFLDDNQIVTSSGDTTCALWDIETGQQTTTFTGHTGDVMSLSLAPD----

Eh Gβ1	--QNTFVSGACDSMAKLWDIRMENCVATFTGHDADINAIAFHPSGNAFITGSDDFSCKLFDIRADRELMN
Dd Gβ	--KNYFISGACDATAKLWDIRSGKCVQTFTGHEADINAVQYFPNGLSFGTGSDDASCRLFDIRADRELMQ
Sc STE4	NSSNTFASCGSDGYTYIWDSRSPSAVQSFYVNDSDINALRFFKDGMSIVAGSDNGAINMYDIRSDCSIAT
Hs Gβ1	--TRLFVSGACDASAKLWDVREGMCRQTFTGHESDINAICFFPNGNAFATGSDDATCRLFDLRADQELMT

Eh Gβ1	YS---------------SESMQH------------------------GVTSVAISSTGRYLFCGYDDLGCLWWD
Dd Gβ	YT--------------HDNILC------------------------GITSVGFSFSGRFLFAGYDDFTCNVWD
Sc STE4	FSLFRGYEERTPTPTYMAANMEYNTAQSPQTLKSTSSSYLDNQGVVSLDFSASGRLMYSCYTDIGCVVWD
Hs Gβ1	YS--------------HDNIIC------------------------GITSVSFSKSGRLLLAGYDDFNCNVWD

Eh Gβ1	VLKGDYITKLTGHENRVSCLGVSPDGYALCTGSWDSTLRIWAN---
Dd Gβ	TLKGERVLSLTGHGNRVSCLGVPTDGMALKIWA---
Sc STE4	VLKGEIVGKLEGHGGRVTGVRSSPDGLAVCTGSWDSTMKIWSPGYQ
Hs Gβ1	ALKADRAGVLAGHDNRVSCLGVTDDGMAVAS-----VLG-------

Figure 2.1. The genome of *Entamoeba histolytica* encodes heterotrimeric G-protein subunits. **(A)** A multiple sequence alignment of EhGα1 with selected Gα subunits from other species (Dd = *Dictyostelium discoideum*, Sc = *Saccharomyces cerevisiae*, Hs = *Homo sapiens*). The secondary structure information above the aligned sequences reflects the crystal structure of EhGα1 (this study), with naming adapted from human transducin (PDB 1TND). Residues mutated in this study are marked with black arrowheads, and gray bars indicate relative sequence identity. A 110-residue insert within Sc GPA1 (gray box) was omitted for clarity. Although a number of *E. histolytica* proteins are reportedly ADP-ribosylated by pertussis toxin [82], EhGα1 is not likely to be a substrate as it lacks the C-terminal cysteine ADP-ribosylation site shared among conventional Gα$_{i/o}$ subunits (e.g., Cys-351 in human Gα$_o$). Based on the sequence of the amino terminus of EhGα1, it is likely that this protein is myristoylated on its second residue (glycine) and palmitoylated on its third residue (cysteine) [57]. **(B)** EhGα1 is aligned with selected Gβ subunits in a fashion identical to panel A with secondary structure elements as found in transducin Gβγ (PDB 1TBG).
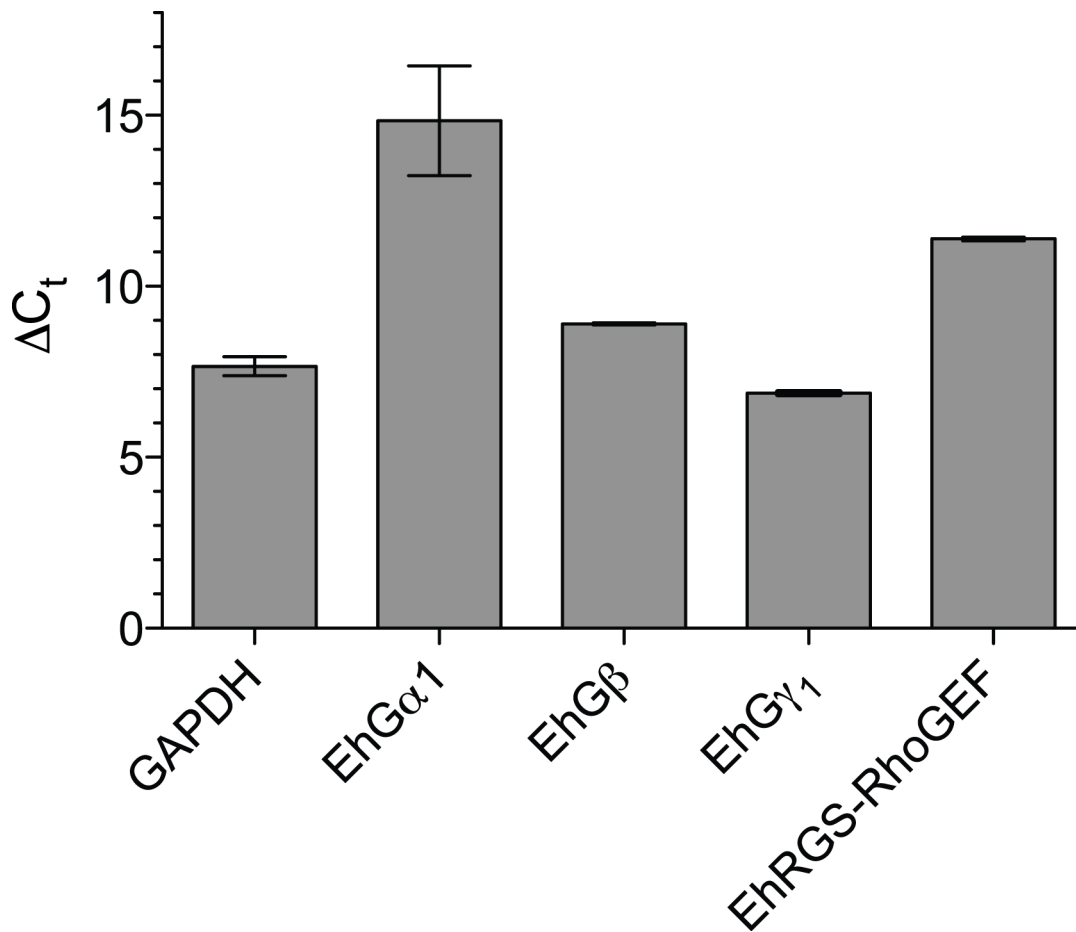
Figure 2.2. Heterotrimeric G-protein signaling components are expressed in *E. histolytica*. qRT-PCR amplification of RNA isolated from HM1 *E. histolytica* trophozoites (a kind gift of Dr. William Petri, Jr.) confirmed transcription of EhGα1, EhGβ1, EhGγ1, and EhRGS-RhoGEF genes. The basally expressed housekeeping gene GAPDH was included as a control. ΔC_t reflects the difference in threshold cycle relative to reactions lacking reverse transcriptase, used as a control for DNA contamination. Error bars represent standard error of the mean.
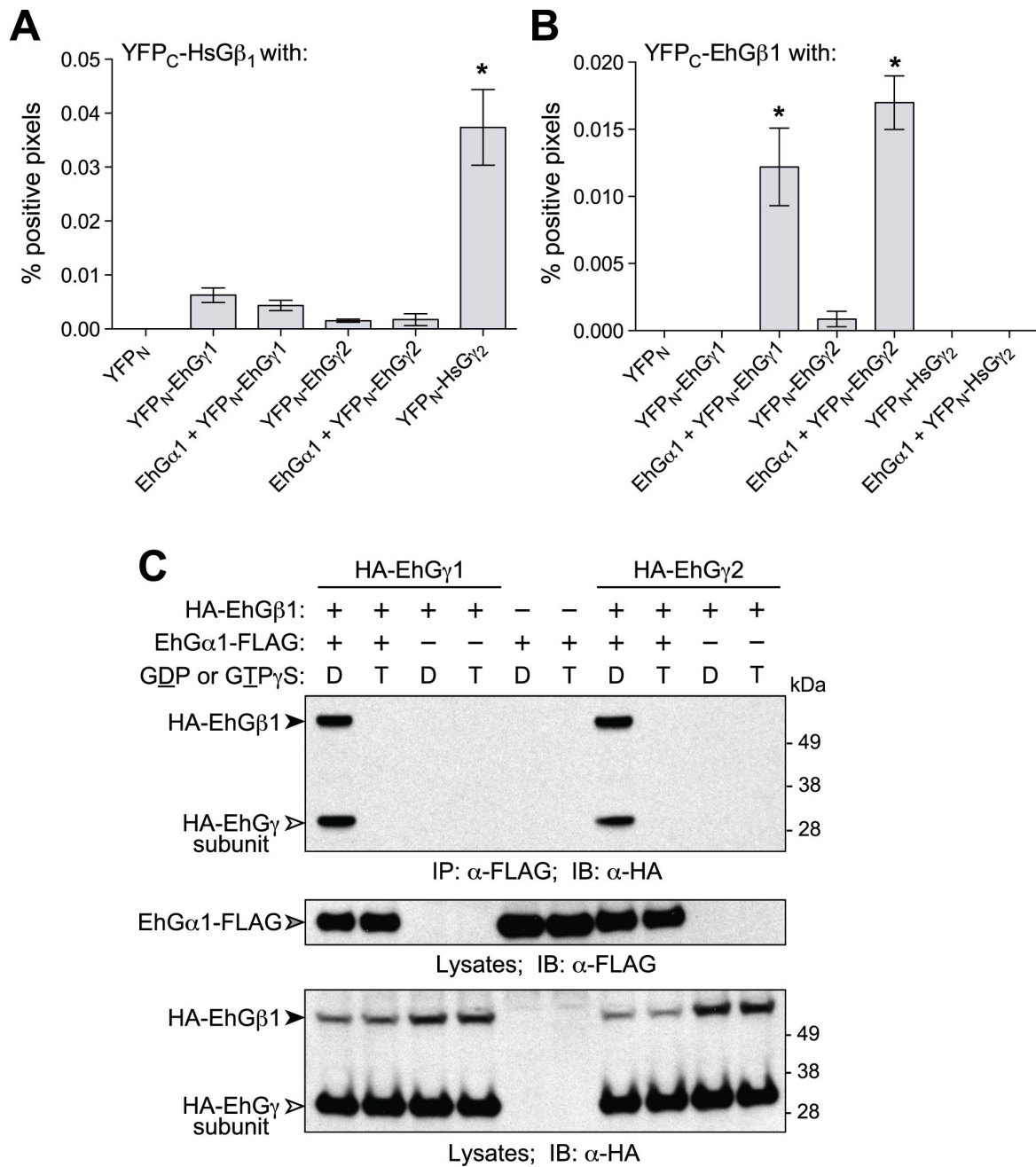
Figure 2.3. *E. histolytica* **G-protein subunits form a heterotrimer in a nucleotide-dependent manner.** Interactions between Gβ and Gγ subunits were detected with split-YFP protein complementation in COS-7 cells. **(A)** Human Gβ1 heterodimerized with human Gγ2, but not with *E. histolytica* Gγ subunits. **(B)** EhGβ1 interacts with EhGγ1 or EhGγ2 when co-expressed with EhGα1. **(C)** G-protein heterotrimer formation in the presence of excess GDP ("D") or the non-hydrolyzable GTP analog, GTPγS ("T"), was examined with co-immunoprecipitation. EhGβ1 and EhGγ1 or EhGγ2 interacted selectively with EhGα1 in its GDP-bound, inactive state. Error bars represent standard error of the mean for three experiments. * represents statistically significant difference from zero, as determined by 95% confidence intervals excluding zero.
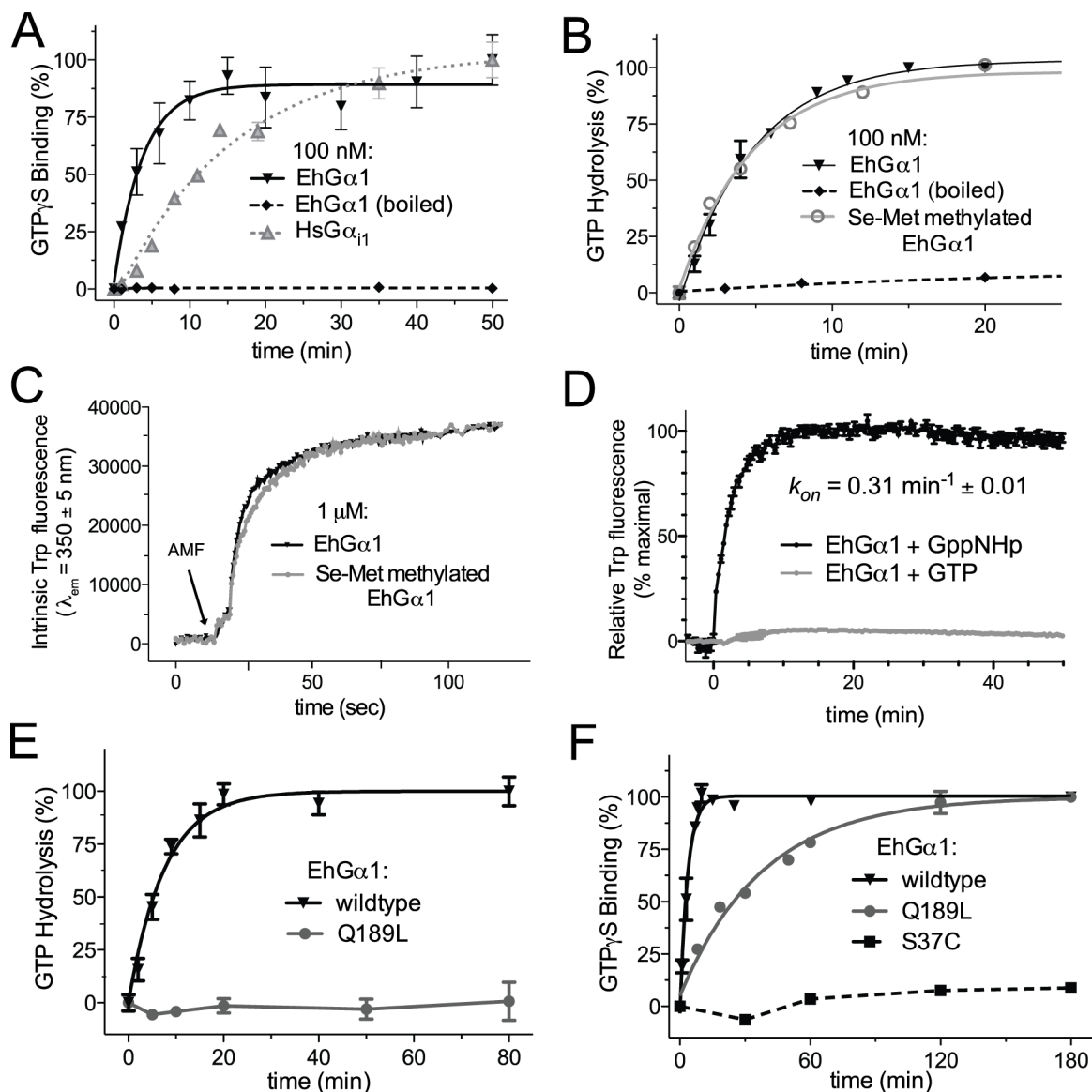
Figure 2.4. EhGα1 cycles between an active, GTP-bound state and an inactive, GDP-bound state. **(A)** EhGα1 bound non-hydrolyzable GTPγS as determined by radionucleotide binding. The observed exchange rate, $k_{obs} = 0.27$ min^{-1} ± 0.06, indicates faster spontaneous GDP release than human Gα$_{i1}$ ($k_{obs} = 0.06$ min^{-1} ± 0.01). **(B)** EhGα1 hydrolyzed GTP[γ-^{32}P] at 0.21 min^{-1} ± 0.02, as determined by single turnover hydrolysis assays. No difference was observed for selenomethionine, lysine-methylated EhGα1 used for crystallization. **(C)** EhGα1 changes conformation upon binding the transition state mimetic aluminum tetrafluoride. Intrinsic EhGα1 fluorescence following excitation at 285 nm increases upon activation, reflecting burial of a conserved tryptophan on switch 2 (Trp-196). **(D)** EhGα1 adopts an active switch conformation upon addition of the non-hydrolyzable GTP analog GppNHp, as reflected by increased intrinsic tryptophan fluorescence. The kinetics of GppNHp-mediated activation are consistent with the kinetics of radiolabeled GTP analog binding (Fig. 2.4A). In contrast, addition of hydrolyzable GTP does not result in EhGα1

activation, indicating that nucleotide exchange, rather than GTP hydrolysis, is the rate-limiting step in the nucleotide cycle of EhGα1. **(E, F)** Two EhGα1 point mutants were profiled for effects on nucleotide cycle. The dominant negative S37C possessed negligible GTP binding. The constitutively active Q189L bound but did not hydrolyze GTP. Error bars in all panels represent standard error of the mean.
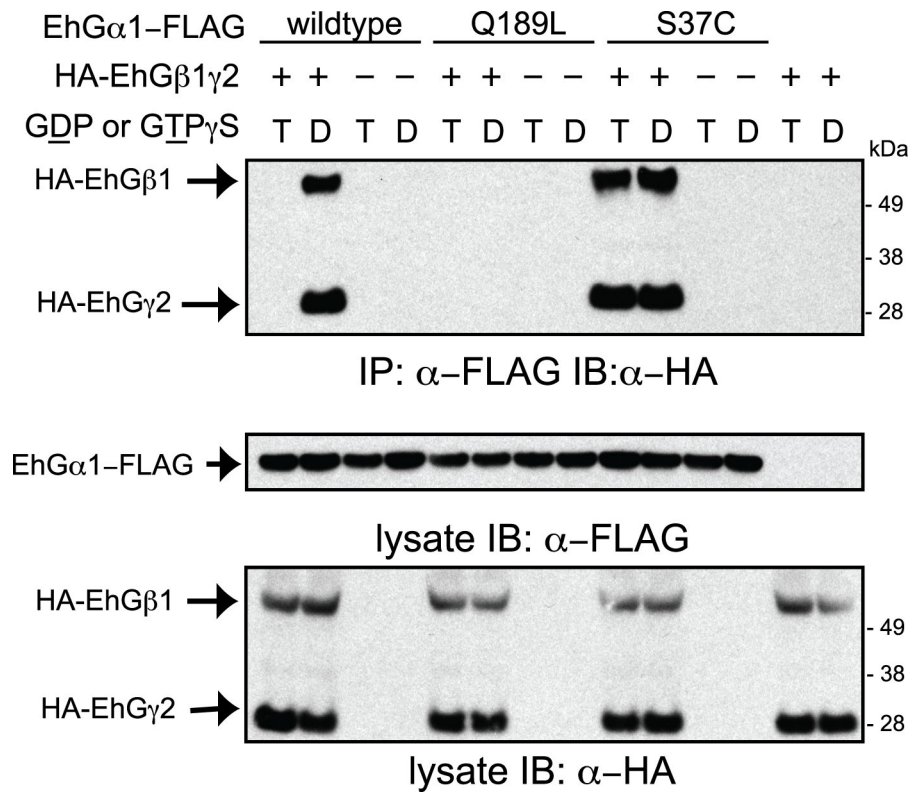
Figure 2.5. The inactive EhGα1(S37C) constitutively binds to EhGβ1γ2, while the constitutively active EhGα1(Q189L) mutant does not. Co-immunoprecipitations of EhGα1 and mutants with EhGβ1 and EhGγ2 were conducted as in Figure 2.3. As predicted, the dominant negative S37C mutant remains bound to EhGβ1γ2, even in excess GTPγS. The constitutively active, GTPase-deficient Q189L mutant does not bind EhGβ1γ2 in either nucleotide state.
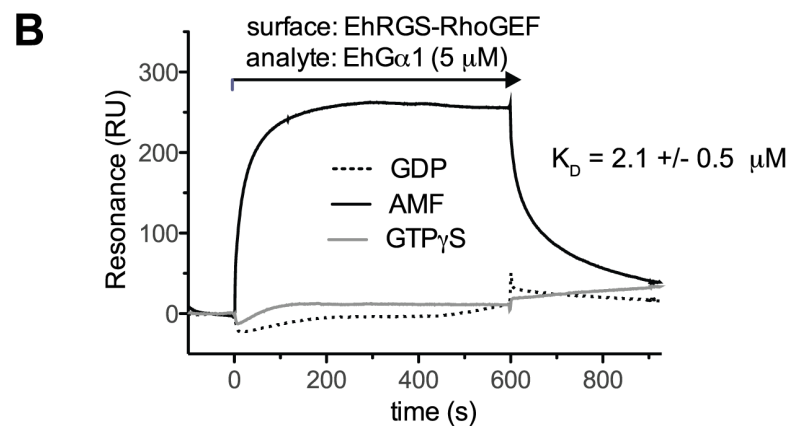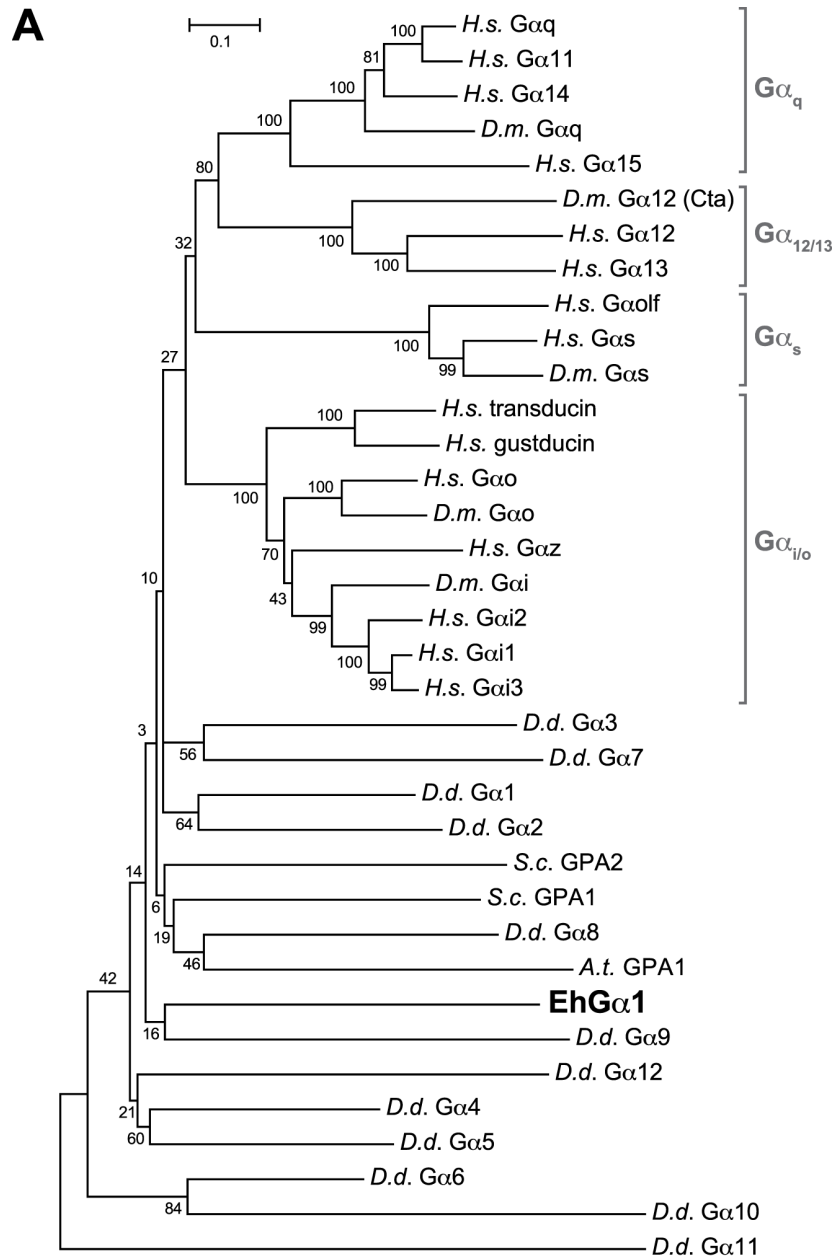
Figure 2.6. Evolutionary relationship of Gα subunits and identification of EhRGS-RhoGEF as a putative effector for activated EhGα1. (A) Gα subunit protein sequences from *E. histolytica*, *D. discoideum* (D.d.), *A. thaliana* (A.t.), *S. cerevisiae* (S.c.), *D. melanogaster* (D.m), and *H. sapiens* (H.s.) were aligned and a bootstrapping consensus phylogram created using MEGA5 [42]. Bootstrap values are indicated at each branch point. EhGα1 is distantly related to metazoan Gα subunits, specifically the adenylyl cyclase stimulatory $G\alpha_s$, adenylyl cyclase inhibitory $G\alpha_{i/o}$, phospholipase Cβ coupled $G\alpha_q$, and RGS-RhoGEF activating $G\alpha_{12/13}$ subfamilies. **(B)** Recombinant EhRGS-RhoGEF protein was immobilized on a surface plasmon resonance chip and EhGα1 protein flowed over in one of two nucleotide states. The EhRGS-RhoGEF biosensor bound EhGα1 selectively in the activated, $GDP \cdot AlF_4^-$-bound state (AMF).
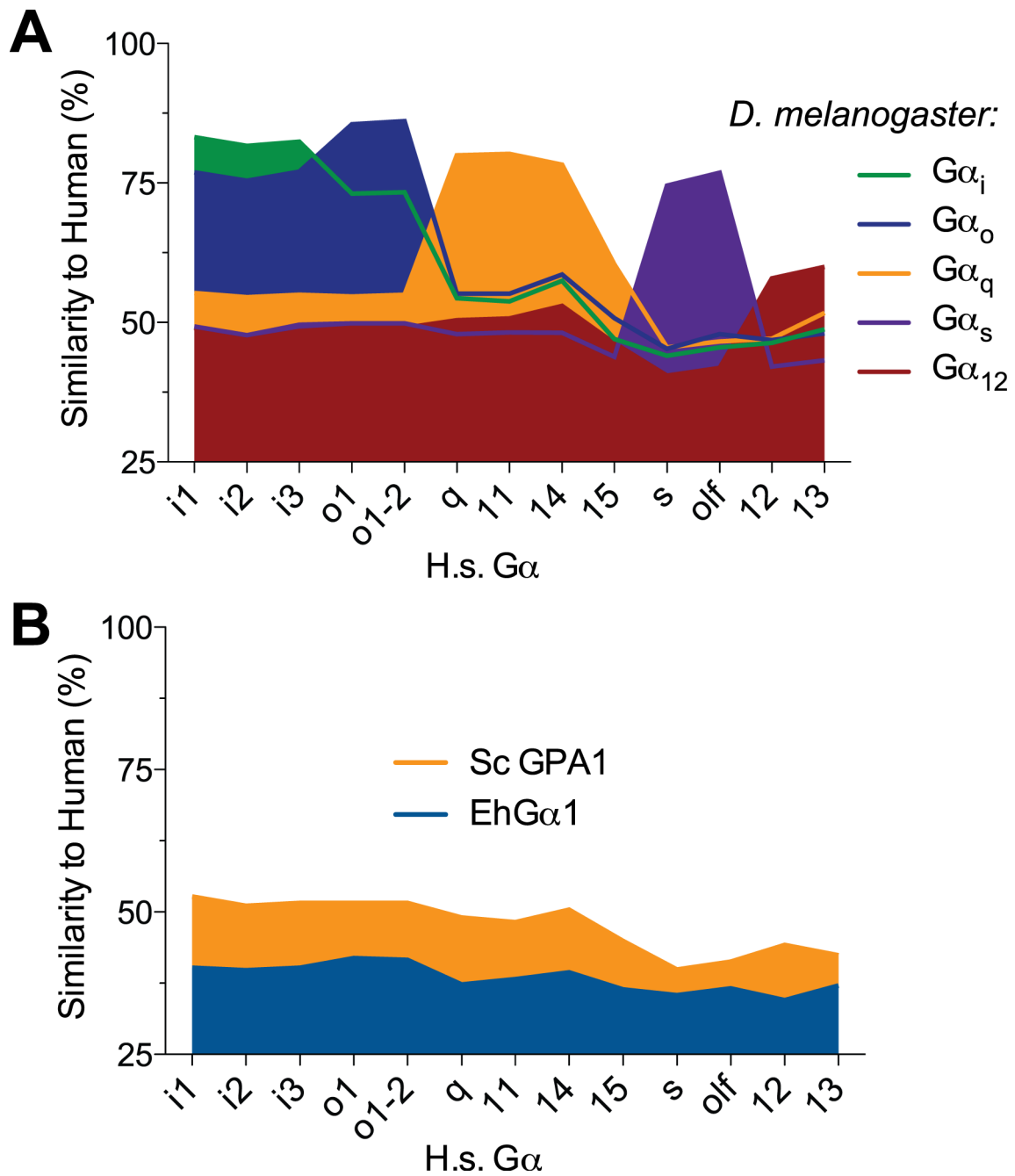
Figure 2.7. Mammalian Gα subfamily homology analyses. Sequence similarity to human Gα subunits was plotted for the Gα subunits from *Drosophila melanogaster* **(A)**, *Saccharomyces cerevisiae* GPA1, and EhGα1 **(B)**. In contrast with *D. melanogaster* subunits, EhGα1 cannot be classified as a member of any particular Gα subfamily.
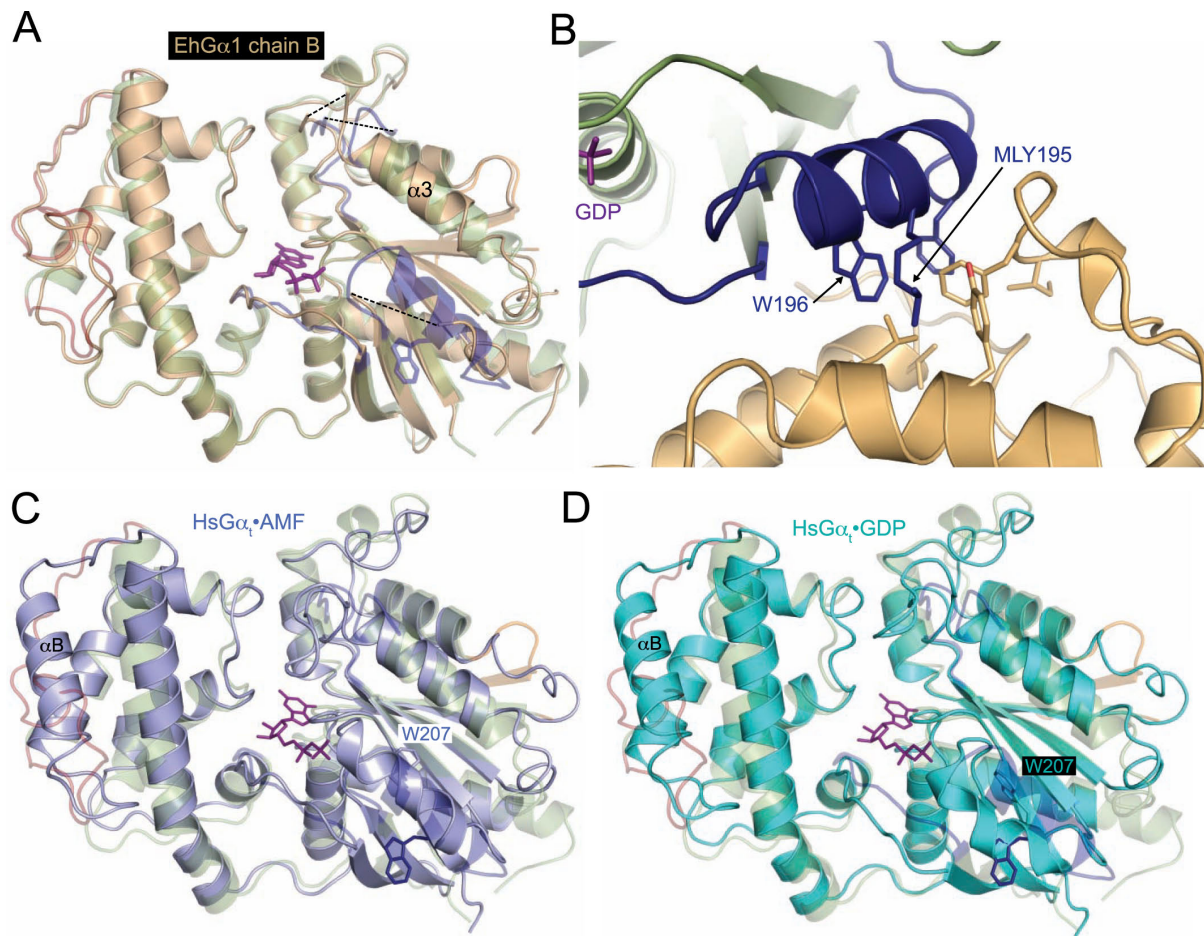
Figure 2.8. Structural comparison of EhGα1 with *Hs* transducin and switch 2 crystal contacts. **(A)** The two EhGα1 molecules in the asymmetric unit are highly similar, although switch 2 of chain B (wheat) is partially disordered. **(B)** Crystal contacts between the ordered switch 2 of chain A (blue) and a neighboring molecule (orange) likely account for the structural differences between the two molecules in the asymmetric unit. The non-polar Trp-196 and N-dimethyl lysine-195 (MLY-195) interface with a hydrophobic patch on a neighboring molecule. Switch 2 may be drawn away from the nucleotide pocket, accounting for the absence of bound AlF$_4^-$ (see discussion below). **(C, D)** The model of EhGα1 is superposed with human transducin in two nucleotide states (slate blue, AMF, PDB 1TAD; teal, GDP, PDB 1TAG). EhGα1 lacks an αB helix seen in transducin and all other Gα subunits and contains a unique α4-β6 insert (orange). Switch 2 of EhGα1 (chain A) adopts a distinct conformation from both the active and inactive forms of transducin, likely due to crystal contacts with a neighboring molecule.

A

All-helical Domain   Ras-like Domain

B

83
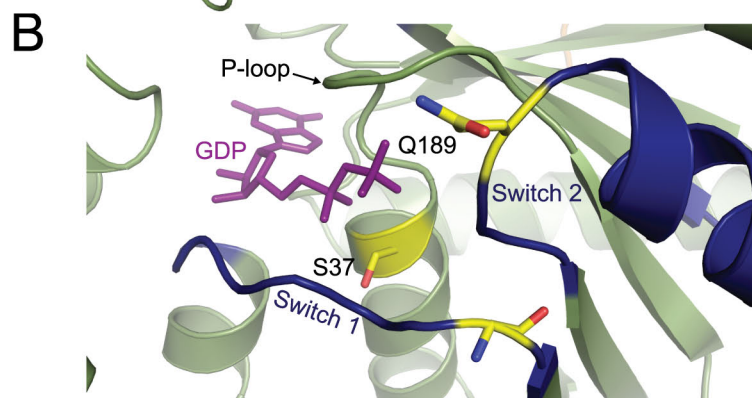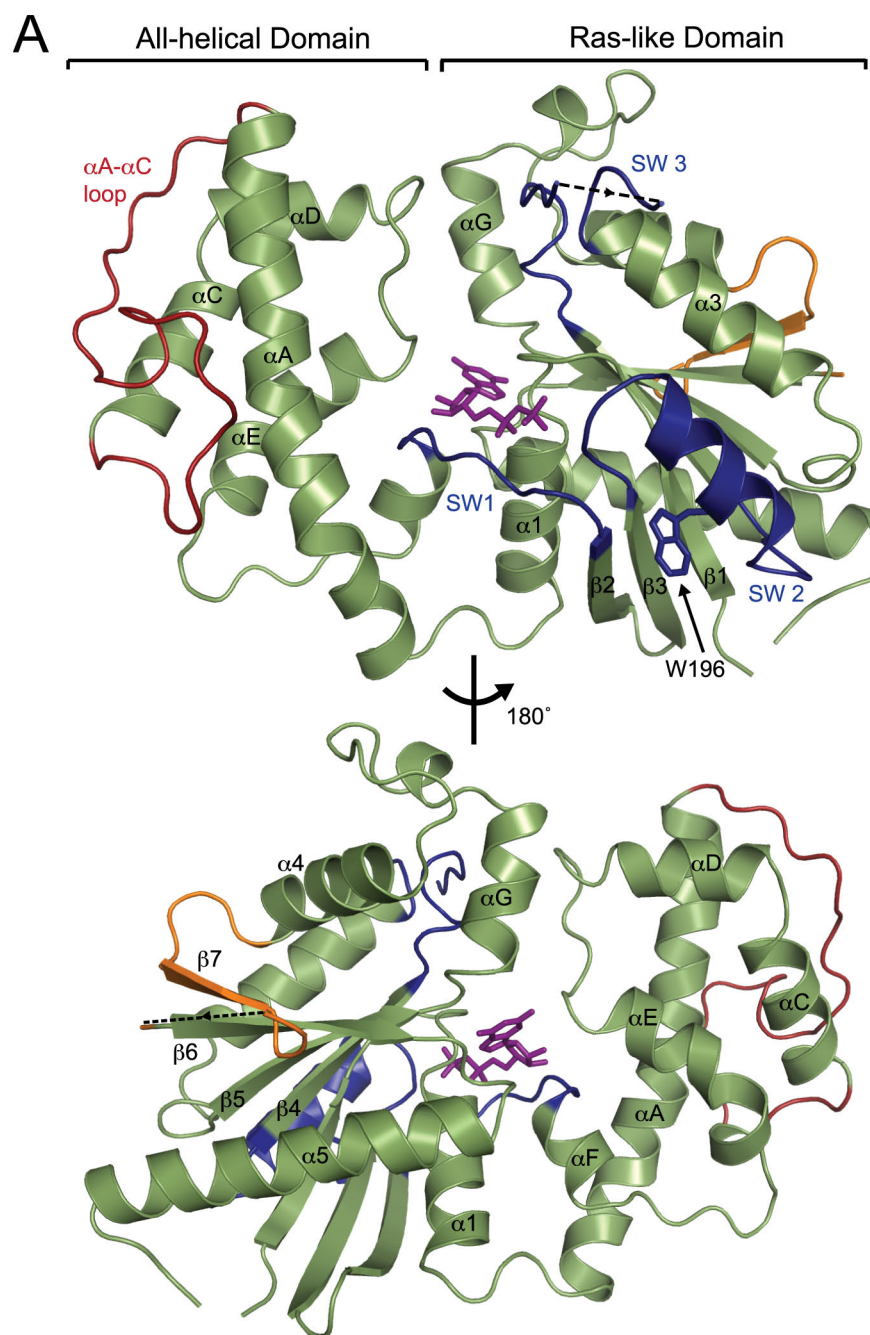
Figure 2.9. Structure of EhGα1 reveals a conserved fold with unique features. The crystal structure of EhGα1 was determined by single anomalous dispersion (SAD) using 2.6 Å resolution data (Table 2.1). **(A)** The EhGα1 Cα backbone is shown in green, bound to GDP in purple sticks. Conserved switch regions (SW 1-3) are dark blue. Trp-196 is solvent-exposed in the inactive state and buried when switch 2 adopts its activated conformation. Unique among Gα subunits, EhGα1 lacks an αB helix in the all-helical domain (red; labeled 'αA-αC loop') but possesses a unique short β-strand insert (β7) and a loop (orange) between the conserved α4 helix and β6 strand. Disordered regions in switch 3 (residues 222 and 223) and the β7-β6 loop (residues 302-310) are indicated by dashed lines. **(B)** Ser-37, conserved among Gα subunits, is an important ligand for Mg^{2+}, a cofactor for GTP binding and hydrolysis. Mutation of Ser-37 to Cys is predicted to produce a dominant negative EhGα1 [41]. Gln-189 is required for orienting the nucleophilic water during GTP hydrolysis; its mutation to Leu is predicted to cripple GTPase activity, yielding a constitutively active EhGα1.
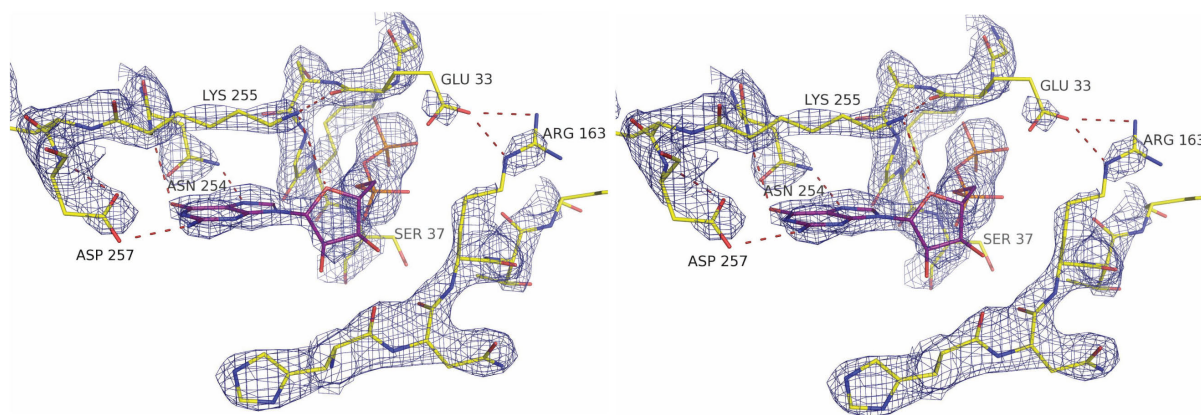
Figure 2.10. Electron density map of the guanine nucleotide binding pocket of EhGα1.
A region of the 2F$_o$-F$_c$ electron density map is shown in stereo view from the structure of
EhGα1 (*yellow sticks*) bound to GDP (*purple sticks*). The nucleotide binding pocket is highly
similar to mammalian Gα subunits, featuring a conserved phosphate binding loop (P-loop;
Glu-33 shown) and an NKxD motif (residues 254-257). Switch one also directly contacts the
nucleotide, and Arg-163 forms polar contacts with the P-loop Glu-33.
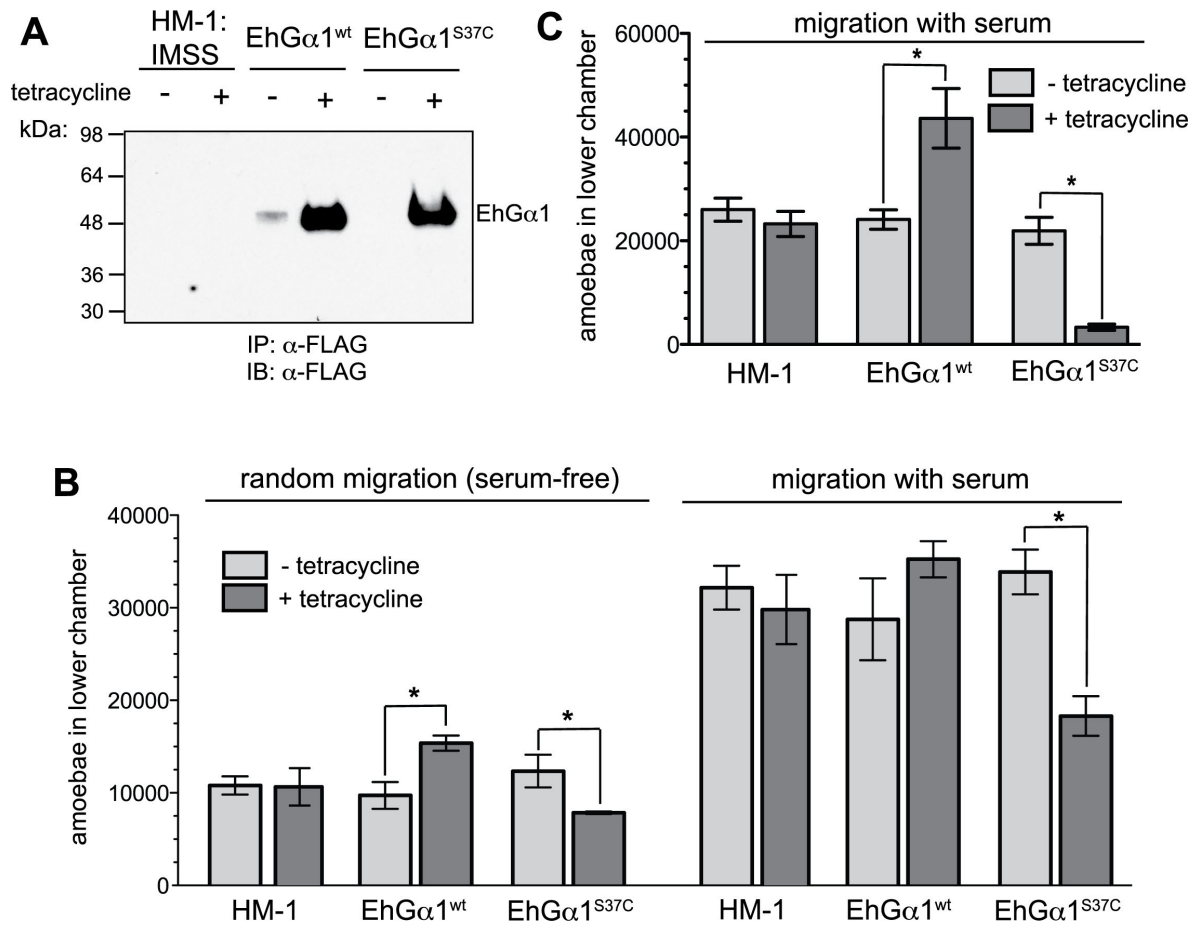
Figure 2.11. Heterotrimeric G-protein signaling increases trophozoite migration across porous membranes and Matrigel layers. (A) Trophozoites were stably transfected to express wildtype $EhG\alpha1$ or dominant negative $EhG\alpha1^{S37C}$ under tetracycline control. **(B)** $EhG\alpha1^{wt}$-expressing trophozoites showed greater migration across a porous membrane in the absence of stimuli (serum-free) while amoebae expressing $EhG\alpha1^{S37C}$ showed lower migration toward both serum-free and serum-containing nutritive media. Migration of HM-1:IMSS trophozoites was not significantly different from the non-induced $EhG\alpha1^{wt}$ and $EhG\alpha1^{S37C}$ strains. Tetracycline treatment was 5 µg/mL over 24 hours. **(C)** Trophozoites expressing $EhG\alpha1^{wt}$ were better able to migrate through a Matrigel layer than uninduced controls. Conversely, $EhG\alpha1^{S37C}$ expression greatly reduced Matrigel transmigration. Parent strain HM-1:IMSS trophozoites were unaffected by tetracycline treatment and were indistinguishable from non-induced $EhG\alpha1^{wt}$ and $EhG\alpha1^{S37C}$. Error bars represent standard error of the mean. * represents statistical significance by an unpaired, two-tailed Student's t-test ($p < 0.05$) for four independent experiments.
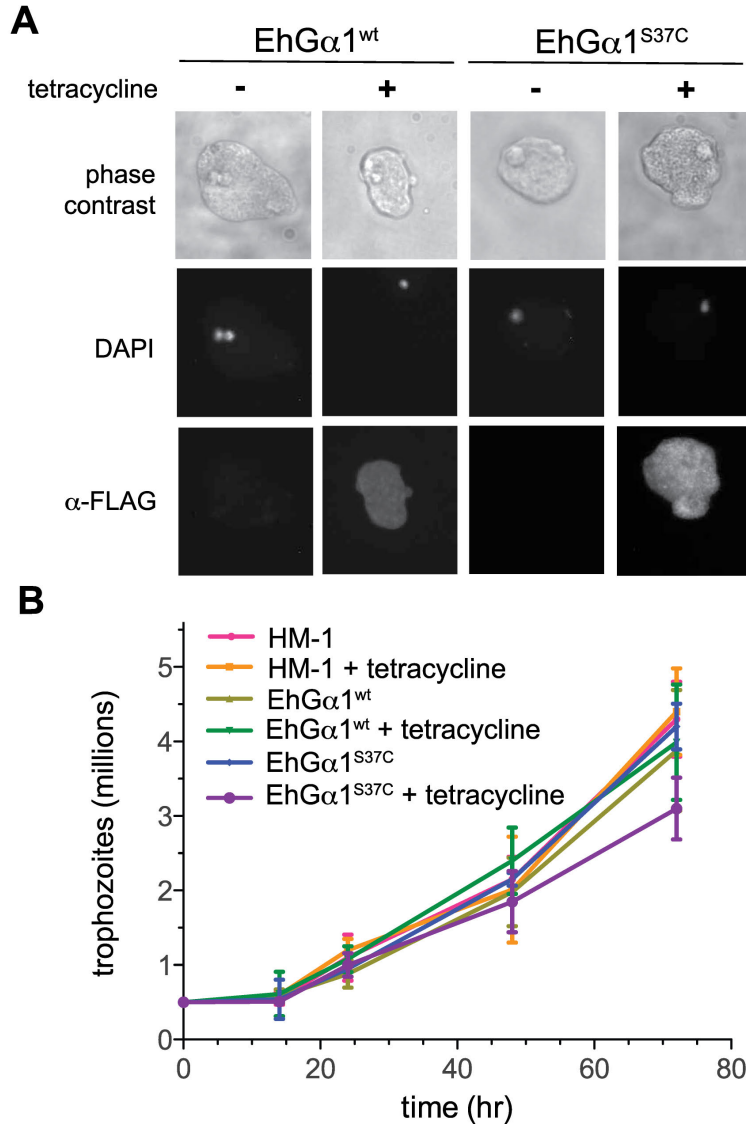
Figure 2.12. Expression of EhGα1^wt^ or EhGα1^S37C^ does not significantly alter trophozoite proliferation. (A) The cellular distribution of overexpressed FLAG-EhGα1^{wt} and FLAG-EhGα1^{S37C} were assessed by immunofluorescence with a Cy3 anti-FLAG conjugate. Both wild type and mutant EhGα1 exhibited similar diffusely cytoplasmic localizations following induced expression by treatment with 5 μg/mL tetracycline for 24 hr. Nuclei were stained with DAPI. **(B)** Trophozoites of the parent HM-1:IMSS, EhGα1^{wt}, and EhGα1^{S37C} strains were seeded in TYI medium with or without 5 μg/mL tetracycline and cell numbers assessed over 3 days. Cell viability was >90% at each measurement, as determined by trypan blue dye exclusion. No significant differences in growth were identified among the strains, although trophozoites induced to express EhGα1^{S37C} trended toward slower growth at day 3. Error bars represent standard error of the mean for three independent experiments.
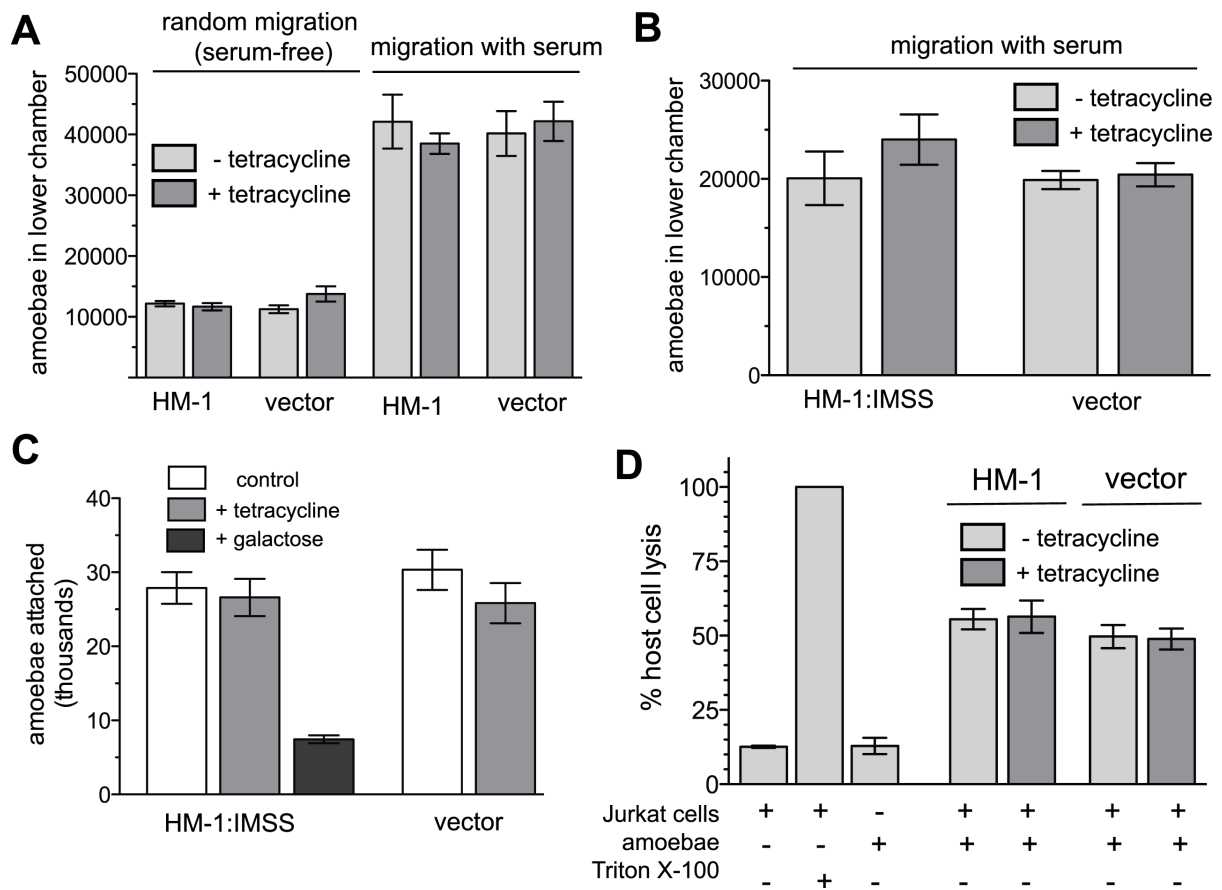
Figure 2.13. *E. histolytica* **transfected with empty vector is not affected by tetracycline treatment.** HM-1:IMSS trophozoites were stably transfected with empty tetracycline-inducible expression vector. **(A)** Transwell migration and **(B)** Matrigel transmigration of parent strain and vector-transfected trophozoites did not differ significantly upon tetracycline treatment of 24 hours prior to the assay. Similarly, transfection with empty vector and tetracycline treatment had no significant effect on host cell attachment **(C)** or host cell killing **(D)**. Error bars represent standard error of the mean for four independent experiments in panels A-C and three independent experiments in panel D. Statistical significance was tested using an unpaired, two-tailed Student's t-test.
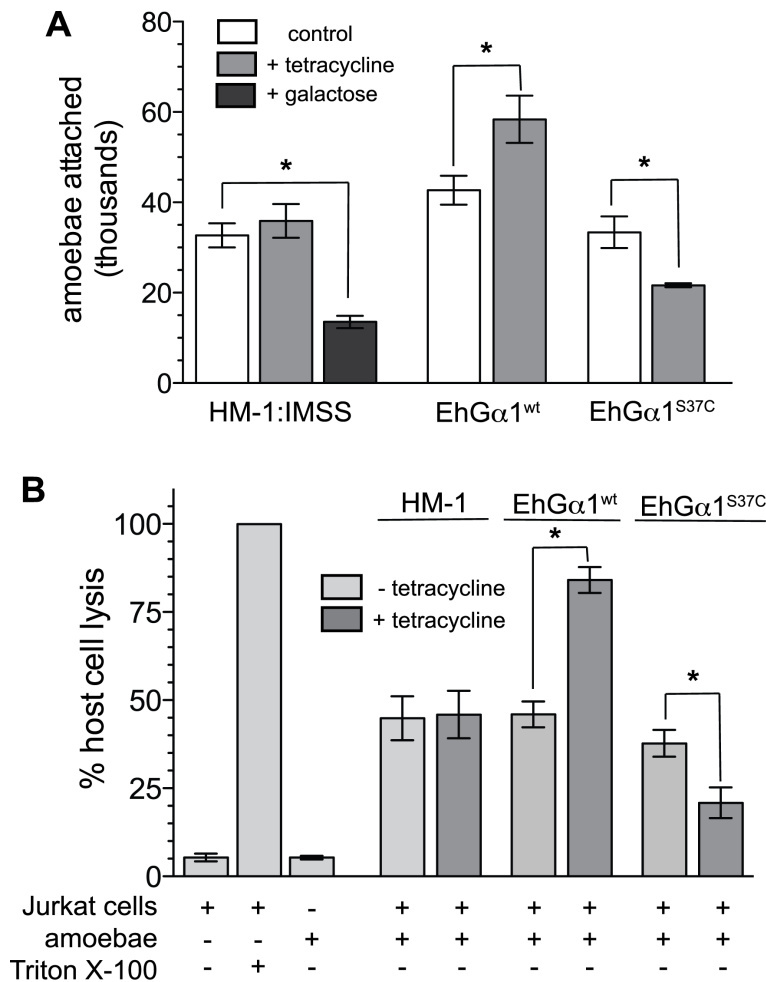
Figure 2.14. Heterotrimeric G-protein signaling positively regulates *E. histolytica* attachment to host cells as well as host cell killing. (A) Trophozoites attach to CHO cell monolayers, primarily through a galactose-inhibitable lectin. Overexpression of EhGα1^{wt} enhanced monolayer attachment, while expression of EhGα1^{S37C} reduced attachment. Parent strain HM-1:IMSS trophozoites were unaffected by tetracycline treatment and were indistinguishable from non-induced EhGα1^{wt} and EhGα1^{S37C}. Attached trophozoites quantities were obtained by multiplying detached cell concentrations by a dilution factor. * indicates a statistically significant difference ($p < 0.05$) between quadruplicate experiments. Error bars represent standard error of the mean. * indicates statistical significance by an unpaired, two-tailed Student's t-test ($p < 0.05$) for four independent experiments. **(B)** Amoebae overexpressing EhGα1^{wt} or EhGα1^{S37C} displayed enhanced or reduced abilities to kill Jurkat (human T-lymphocyte) cells, respectively, as measured by LDH release in a membrane integrity assay. Cell killing by HM-1:IMSS trophozoites was not altered by tetracycline treatment. 0.5% Triton X-100 was added to Jurkat cells to define 100% host cell lysis. Tetracycline treatment was 5 µg/mL over 24 hours. Error bars represent standard error of the mean. * indicates statistical significance by an unpaired, two-tailed Student's t-test ($p < 0.05$) for three independent experiments, with four technical replicates each.

89
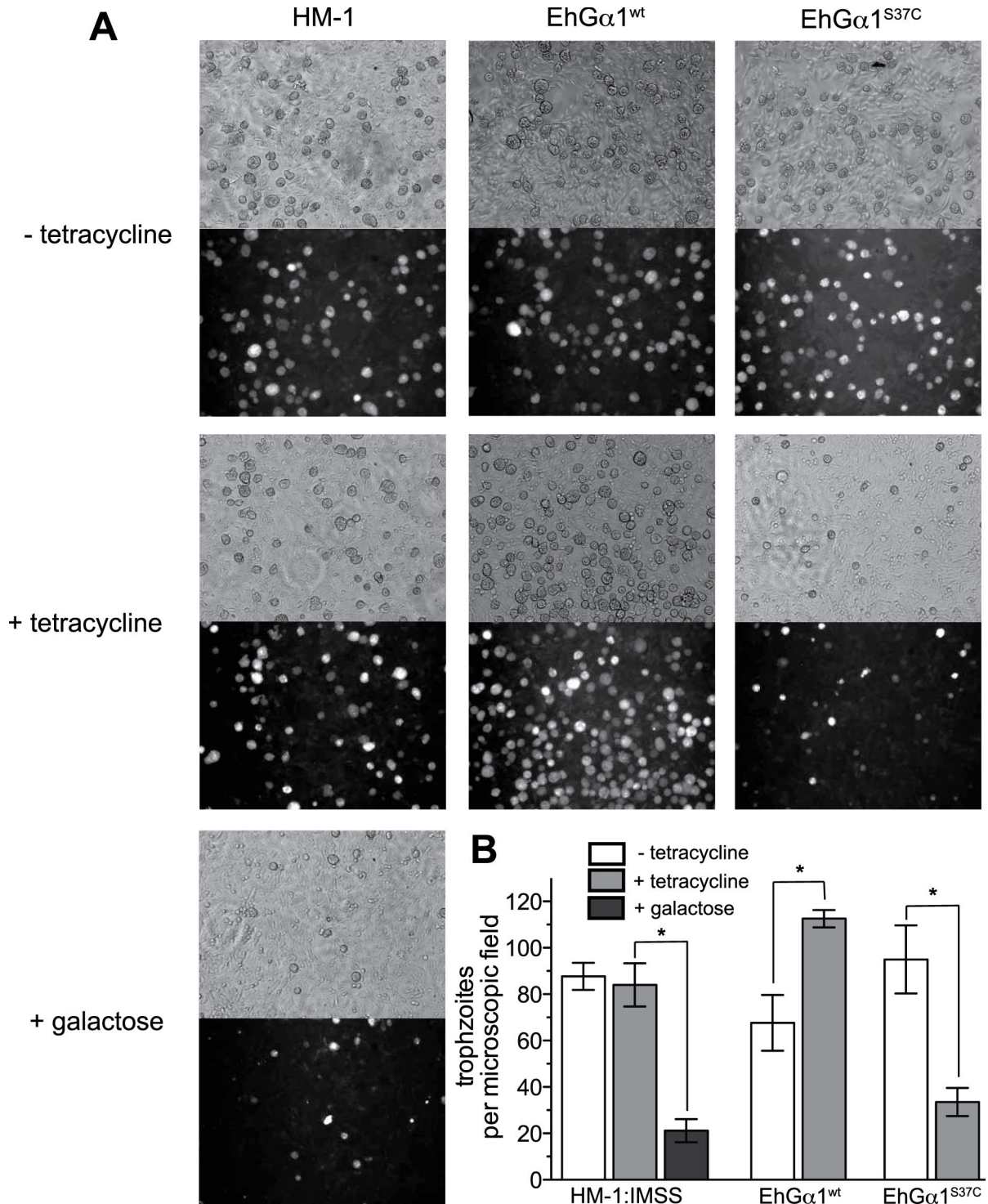
Figure 2.15. Microscopic analysis of perturbed *E. histolytica* attachment to host cells upon overexpression of EhGα1^{wt} or EhGα1^{S37C}. (A) Trophozoites grown in the presence or absence of 5 µg/mL tetracycline were fluorescently labeled with CFDA and allowed to attach to fixed, confluent layers of CHO cells. Phase contrast (upper panels) and epifluorescence (lower panels) images were obtained of attached trophozoites. **(B)** Attachment was quantified by counting trophozoites in three microscopic fields (10X).

Overexpression of EhGα1^{wt} enhanced monolayer attachment, while expression of EhGα1^{S37C} reduced attachment. Parent strain HM-1:IMSS trophozoites were unaffected by tetracycline treatment and were indistinguishable from non-induced EhGα1^{wt} and EhGα1^{S37C}. Error bars represent standard error of the mean. * represents statistical significance by an unpaired, two-tailed Student's t-test ($p < 0.05$) for three independent experiments.
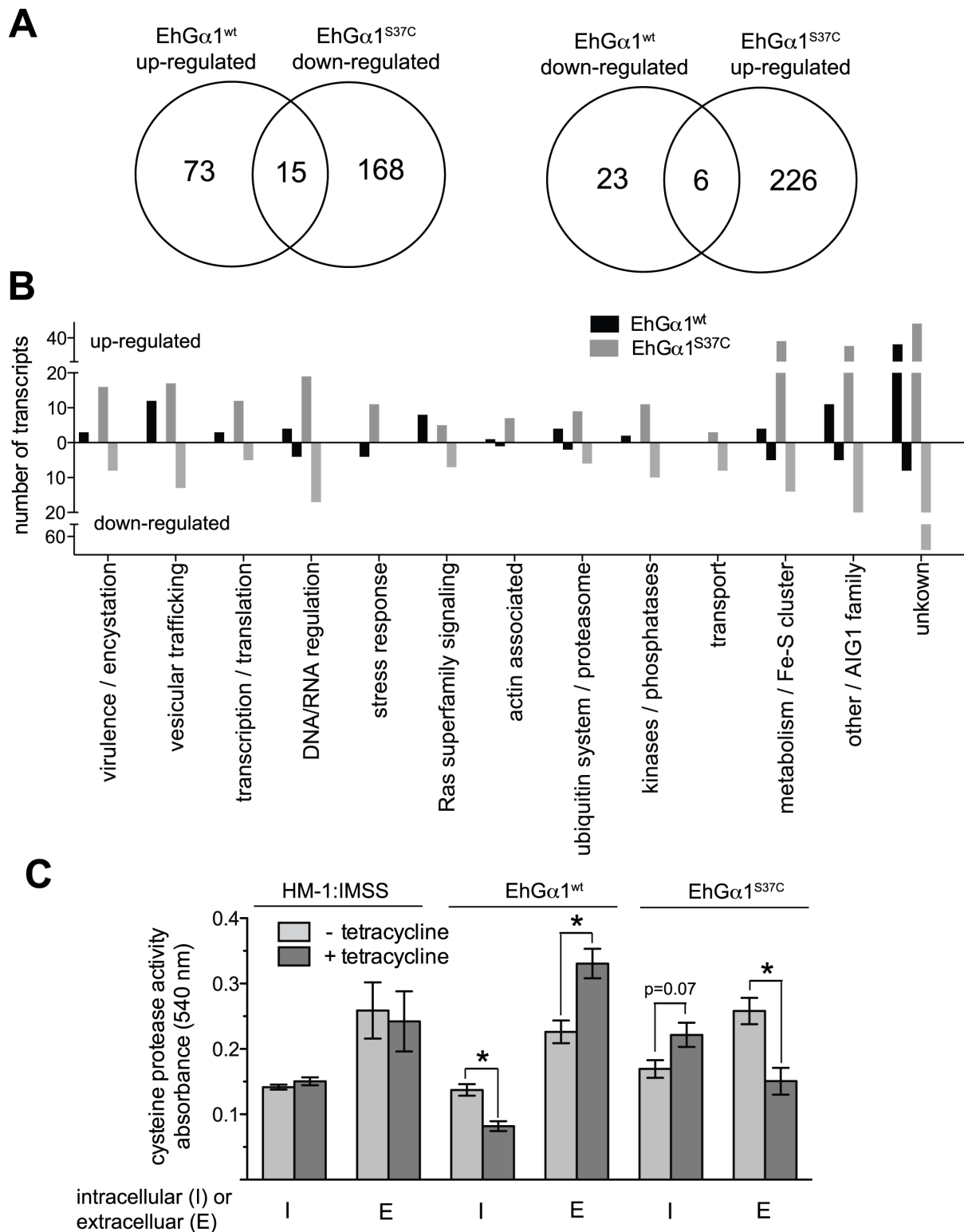
Figure 2.16. Heterotrimeric G-protein signaling alters _E. histolytica_ transcription to modulate cysteine protease secretion. (A) 96 genes or 394 genes were differentially transcribed upon overexpression of EhGα1^wt or EhGα1^S37C, respectively, when compared to

92

uninduced controls as determined by RNA-seq. 21 transcripts were oppositely regulated in trophozoites expressing EhGα1^{wt} *vs* EhGα1^{S37C}. **(B)** Differentially transcribed genes were categorized by putative function based on prior studies, homology to genes of known function, or predicted protein domains of known function. "Virulence/encystation" category includes genes known to modulate *E. histolytica* pathogenesis, such as cysteine proteases [1]. **(C)** Both intracellular and secreted cysteine protease activities were assessed with an azo-collagen assay. EhGα1^{wt} overexpression enhanced cysteine protease secretion, while EhGα1^{S37C} expression resulted in less extracellular (E), despite higher intracellular (I), cysteine protease activity, suggesting that transcriptional responses downstream of heterotrimeric G-protein signaling modulate *E. histolytica* pathogenic processes in part by regulating cysteine protease secretion. Tetracycline treatment in all experiments was 5 μg/mL over 24 hours. * = statistical significance by an unpaired, two-tailed Student's t-test ($p < 0.05$) for four independent experiments.
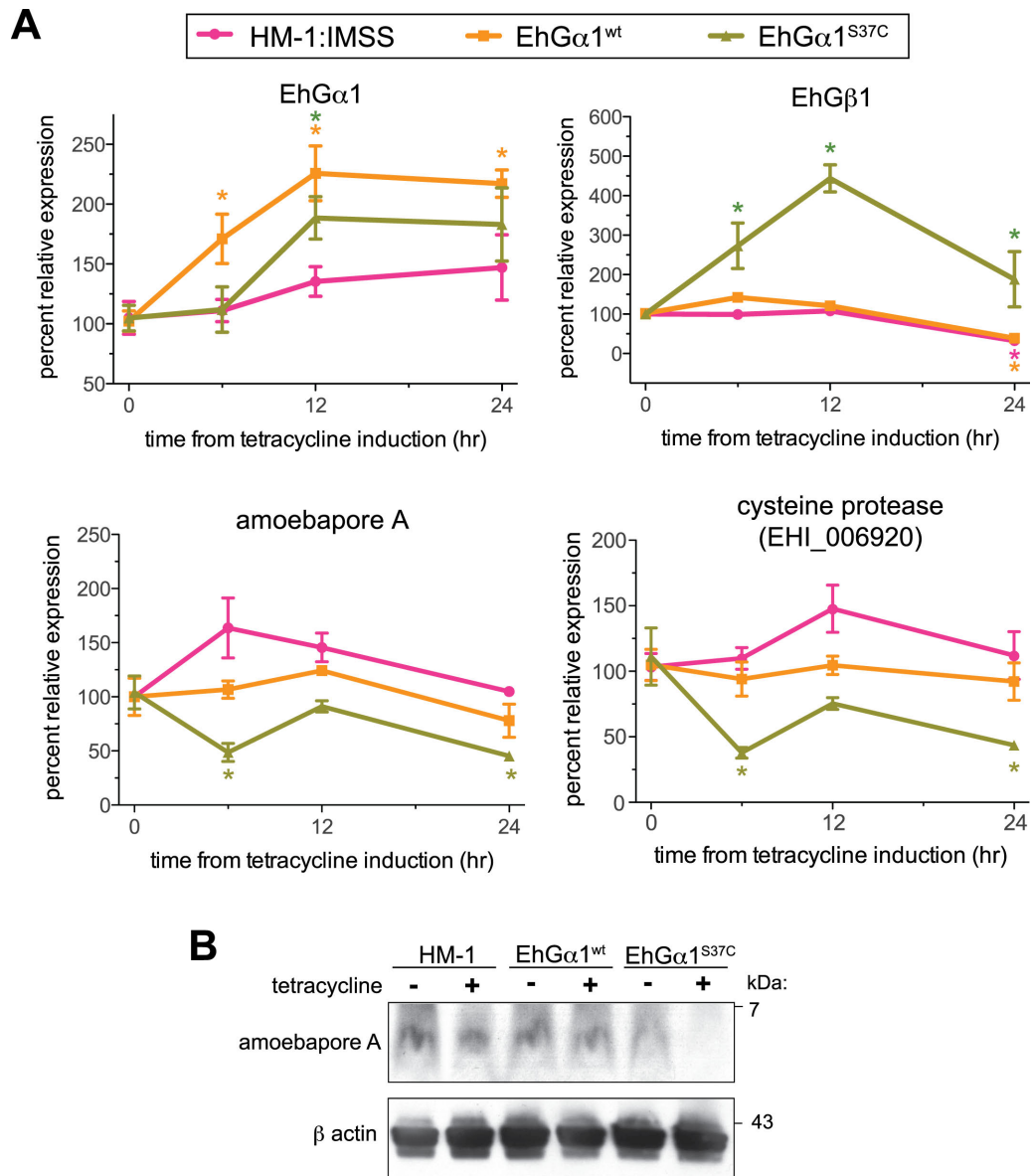
Figure 2.17. RT-PCR analysis of differentially transcribed genes and altered expression of amoebapore A protein. (A) qRT-PCR amplification of RNA isolated from HM1 *E. histolytica* trophozoites confirmed differential transcription of EhGα1, EhGβ1, amoebapore A, and a cysteine protease (EHI_006920) upon tetracycline treatment of the parent HM-1:IMSS, EhGα1^{wt}, or EhGα1^{S37C} strains over 24 hours. * indicates statistically significant difference from time zero (no tetracycline exposure), using an unpaired, two-tailed Student's t-test for two technical duplicates of two independent experiments. EhGα1 expression was significantly up-regulated in the EhGα1^{wt} and EhGα1^{S37C} strains, while EhGβ1 was up-regulated and amoebpore A and cysteine protease (EHI_006920) were down-regulated upon expression of EhGα1^{S37C}. **(B)** Trophozoite lysates were subjected to western blotting with anti-amoebapore A (kind gift of M. Leippe, U. of Kiel, Germany), with actin serving as a loading control. Amoebapore A protein expression is reduced in parallel with its transcriptional downregulation upon overexpression of EhGα1^{S37C}.

94

Table 2.1. Data collection and refinement statistics for lysine-methylated selenomethionine EhGα1.

	EhGα1·GDP
PDB accession code	4FID
Data collection	
Space group	P2$_1$2$_1$2$_1$
Cell dimensions	
a, b, c (Å)	56.3, 56.9, 229.8
α, β, γ (°)	90, 90, 90
	Peak
Wavelength (Å)	0.97954
Resolution (Å)	45.9 - 2.60 (2.62 – 2.60)*
No. unique reflections	43,503
R_{merge} (%)	7.7 (98.2)**
$I / \sigma I$	30.7 (2.6)
Completeness (%)	99.3 (94.0)
Redundancy	13.3 (13.3)
Refinement	
Resolution (Å)	45.9 – 2.6 (2.7 – 2.6)
No. reflections	43503 (4336)
R_{work} / R_{free} (%)	18.9 / 25.8 (25.8 / 36.4)
No. atoms	5013
Protein	4905
Ligand/ion	56
Water	39
B-factors (Å^2)	
Protein	58.1
Ligand/ion	57.4
Water	44.8
R.m.s deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.205

*Values in parentheses are for highest-resolution shell.
**All data were collected from a single crystal.

Table 2.2. Genes differentially transcribed in *E. histolytica* trophozoites expressing EhGα1 or EhGα1^{S37C} with known roles in pathogenesis or putative vesicular trafficking functions.

Gene name	AmoebaDB accession no.	fold change upon EhGα1 expression (p value)	fold change upon EhGα1^{S37C} expression (p value)
Virulence genes			
amoebapore C	EHI_118270	2.6 (0.002)	
myosin II heavy chain	EHI_014010	2.4 (< 0.001)	
vacuolar ATP synthase	EHI_029370	1.7 (0.001)	
EhRGS-RhoGEF	EHI_010670	2.2 (0.001)	
amoebapore A precursor	EHI_159480		-1.6 (0.004)
cysteine protease	EHI_062480		-2.2 (< 0.001)
cysteine protease	EHI_126170		-1.5 (0.021)
cysteine protease	EHI_138460		-1.5 (0.007)
cysteine protease	EHI_006920		-2.5 (< 0.001)
cysteine protease	EHI_045290		1.5 (0.010
cysteine protease	EHI_064430		2.3 (< 0.001)
Gal/GalNac lectin light subunit	EHI_049690		1.6 (0.005)
Gal/GalNac lectin intermediate subunit	EHI_183000		2.6 (< 0.001)
EhGβ	EHI_000240		1.6 (0.004)
Rab family GTPases			
Rab GTPase	EHI_168450	1.9 (0.002)	-1.9 (0.001)
Rab GTPase (RabK5)	EHI_012380	1.5 (0.011)	
Rab GTPase (RabX22B)	EHI_014060	1.6 (0.004)	
hypothetical protein (Rab domain)	EHI_014210	1.8 (< 0.001)	
Rab GTPase (Rabl1)	EHI_177550	2.3 (< 0.001)	
Rab GTPase (Rab7C)	EHI_189990		-161.1 (< 0.001)
Rab GTPase (Rab11A)	EHI_005460		2.3 (< 0.001)
Rab GTPase (Rab11D)	EHI_056100		2.4 (< 0.001)
Rab GTPase (Rab2C)	EHI_067850		2.0 (< 0.001)
Rab11B	EHI_107170		10.3 (< 0.001)
Other secretion/trafficking proteins			
VHS domain containing protein	EHI_087590	2.1 (< 0.001)	-1.7 (0.002)
hypothetical protein (DENN domain)	EHI_014090	1.8 (< 0.001)	-1.6 (< 0.001)
hypothetical protein (DENN domain)	EHI_010320	1.8 (0.006)	
Rab GTPase activating protein	EHI_060420	2.6 (< 0.001)	
ankyrin	EHI_004990	2.1 (< 0.001)	
hypothetical protein (SFT2-like)	EHI_200980	1.9 (0.006)	
MIT domain protein	EHI_093860	1.8 (0.007)	
Rab GTPase activating protein	EHI_170310		-1.5 (0.018)
TBC1 domain family RabGAP 5	EHI_058320		-2.0 (< 0.001)
ankyrin repeat protein	EHI_134800		-1.7 (0.019)
hypothetical protein (M6PR binding)	EHI_161980		-2.1 (< 0.001)
kinesin motor protein	EHI_196390		-1.8 (< 0.001)
cysteine protease binding factor 2	EHI_087660		-2.0 (< 0.001)
cysteine protease binding factor 3	EHI_161650		-1.7 (0.001)
cysteine protease binding factor 4	EHI_012340		-1.6 (0.005)
cysteine protease binding factor 5	EHI_137940		-2.3 (< 0.001)
hypothetical protein (DENN domain)	EHI_135120		2.9 (0.002)
hypothetical protein (TBC/RabGAP)	EHI_189120		2.0 (0.005)
vesicle-fusing ATPase	EHI_004640		2.4 (< 0.001)
adaptor protein (AP) family protein	EHI_023600		1.9 (0.001)

Gene name	AmoebaDB accession no.	fold change upon EhGα1 expression (p value)	fold change upon EhGα1^{S37C} expression (p value)
AP-3 complex subunit delta 1	EHI_158840		2.2 (< 0.001)
AP-3 complex subunit	EHI_164810		2.1 (< 0.001)
dynamin-like protein	EHI_052740		2.6 (< 0.001)
clathrin-adaptor medium chain	EHI_099240		2.5 (< 0.001)
coatamer beta subunit	EHI_088220		2.2 (0.002)
coatamer beta subunit	EHI_173180		2.4 (0.018)
coatamer beta subunit	EHI_173390		2.6 (< 0.001)
Sec1 family protein	EHI_093130		2.2 (< 0.001)
hypothetical protein (VHS domain)	EHI_152920		1.8 (0.003)

2.6   REFERENCES

1.   Oldham, W.M. and H.E. Hamm, *Heterotrimeric G protein activation by G-protein-coupled receptors.* Nat Rev Mol Cell Biol, 2008. **9**(1): p. 60-71.

2.   Aittaleb, M., C.A. Boguth, and J.J. Tesmer, *Structure and function of heterotrimeric G protein-regulated Rho guanine nucleotide exchange factors.* Mol Pharmacol, 2010. **77**(2): p. 111-25.

3.   Kimple, A.J., D.E. Bosch, P.M. Giguere, and D.P. Siderovski, *Regulators of G-protein signaling and their Galpha substrates: promises and challenges in their use as drug discovery targets.* Pharmacol Rev, 2011. **63**(3): p. 728-49.

4.   Chen, Z., L. Guo, S.R. Sprang, and P.C. Sternweis, *Modulation of a GEF switch: autoinhibition of the intrinsic guanine nucleotide exchange activity of p115-RhoGEF.* Protein Sci, 2011. **20**(1): p. 107-17.

5.   Gilchrist, A., *Modulating G-protein-coupled receptors: from traditional pharmacology to allosterics.* Trends Pharmacol Sci, 2007. **28**(8): p. 431-7.

6.   Slessareva, J.E. and H.G. Dohlman, *G protein signaling in yeast: new components, new connections, new compartments.* Science, 2006. **314**(5804): p. 1412-3.

7.   Bosch, D.E., et al., *A P-loop mutation in Galpha subunits prevents transition to the active state: implications for G-protein signaling in fungal pathogenesis.* PLoS Pathog, 2012. **8**(2): p. e1002553.

8.   Xu, X., et al., *Coupling mechanism of a GPCR and a heterotrimeric G protein during chemoattractant gradient sensing in Dictyostelium.* Sci Signal, 2010. **3**(141): p. ra71.

9.   Bakthavatsalam, D., J.M. Choe, N.E. Hanson, and R.H. Gomer, *A Dictyostelium chalone uses G proteins to regulate proliferation.* BMC Biol, 2009. **7**: p. 44.

10.   WHO, *WHO/PAHO/UNESCO report. A consultation with experts on amoebiasis. Mexico City, Mexico 28-29 January, 1997.* Epidemiol Bull, 1997. **18**(1): p. 13-4.

11.   Nozaki, T., S. Kobayashi, T. Takeuchi, and A. Haghighi, *Diversity of clinical isolates of Entamoeba histolytica in Japan.* Arch Med Res, 2006. **37**(2): p. 277-9.

12.   Ramos, F., et al., *High prevalence rate of Entamoeba histolytica asymptomatic infection in a rural Mexican community.* Am J Trop Med Hyg, 2005. **73**(1): p. 87-91.

13.   Petri, W.A., Jr., R. Haque, and B.J. Mann, *The bittersweet interface of parasite and host: lectin-carbohydrate interactions during human invasion by the parasite Entamoeba histolytica.* Annu Rev Microbiol, 2002. **56**: p. 39-64.

14. Bracha, R., Y. Nuchamowitz, M. Leippe, and D. Mirelman, *Antisense inhibition of amoebapore expression in Entamoeba histolytica causes a decrease in amoebic virulence.* Mol Microbiol, 1999. **34**(3): p. 463-72.

15. Leippe, M., S. Ebel, O.L. Schoenberger, R.D. Horstmann, and H.J. Muller-Eberhard, *Pore-forming peptide of pathogenic Entamoeba histolytica.* Proc Natl Acad Sci U S A, 1991. **88**(17): p. 7659-63.

16. Tillack, M., et al., *Increased expression of the major cysteine proteinases by stable episomal transfection underlines the important role of EhCP5 for the pathogenicity of Entamoeba histolytica.* Mol Biochem Parasitol, 2006. **149**(1): p. 58-64.

17. Boettner, D.R., et al., *Entamoeba histolytica phagocytosis of human erythrocytes involves PATMK, a member of the transmembrane kinase family.* PLoS Pathog, 2008. **4**(1): p. e8.

18. Shrimal, S., A. Saha, S. Bhattacharya, and A. Bhattacharya, *Lipids induce expression of serum-responsive transmembrane kinase EhTMKB1-9 in an early branching eukaryote Entamoeba histolytica.* Sci Rep, 2012. **2**: p. 333.

19. Buss, S.N., et al., *Members of the Entamoeba histolytica transmembrane kinase family play non-redundant roles in growth and phagocytosis.* Int J Parasitol, 2010. **40**(7): p. 833-43.

20. Somlata, S. Bhattacharya, and A. Bhattacharya, *A C2 domain protein kinase initiates phagocytosis in the protozoan parasite Entamoeba histolytica.* Nat Commun, 2011. **2**: p. 230.

21. Kumar, S., et al., *Crystal structure and trimer-monomer transition of N-terminal domain of EhCaBP1 from Entamoeba histolytica.* Biophys J, 2010. **98**(12): p. 2933-42.

22. Jain, R., et al., *Calcium-binding protein 1 of Entamoeba histolytica transiently associates with phagocytic cups in a calcium-independent manner.* Cell Microbiol, 2008. **10**(6): p. 1373-89.

23. Sahoo, N., et al., *Calcium binding protein 1 of the protozoan parasite Entamoeba histolytica interacts with actin and is involved in cytoskeleton dynamics.* J Cell Sci, 2004. **117**(Pt 16): p. 3625-34.

24. Guillen, N., P. Boquet, and P. Sansonetti, *The small GTP-binding protein RacG regulates uroid formation in the protozoan parasite Entamoeba histolytica.* J Cell Sci, 1998. **111 (Pt 12)**: p. 1729-39.

25. Aguilar-Rojas, A., et al., *Entamoeba histolytica: inhibition of cellular functions by overexpression of EhGEF1, a novel Rho/Rac guanine nucleotide exchange factor.* Exp Parasitol, 2005. **109**(3): p. 150-62.

26.	Arias-Romero, L.E., et al., *EhGEF3, a Novel Dbl Family Member, Regulates EhRacA Activation During Chemotaxis and Capping in Entamoeba histolytica.* Cell Motility and the Cytoskeleton, 2007. **64**: p. 390-404.

27.	Ghosh, S.K. and J. Samuelson, *Involvement of p21racA, phosphoinositide 3-kinase, and vacuolar ATPase in phagocytosis of bacteria and erythrocytes by Entamoeba histolytica: suggestive evidence for coincidental evolution of amebic invasiveness.* Infect Immun, 1997. **65**(10): p. 4243-9.

28.	Welter, B.H., R.R. Powell, M. Leo, C.M. Smith, and L.A. Temesvari, *A unique Rab GTPase, EhRabA, is involved in motility and polarization of Entamoeba histolytica cells.* Mol Biochem Parasitol, 2005. **140**(2): p. 161-73.

29.	Guzman-Medrano, R., et al., *Entamoeba histolytica: alterations in EhRabB protein in a phagocytosis deficient mutant correlate with the Entamoeba dispar RabB sequence.* Exp Parasitol, 2005. **110**(3): p. 259-64.

30.	Nakada-Tsukui, K., Y. Saito-Nakano, V. Ali, and T. Nozaki, *A retromerlike complex is a novel Rab7 effector that is involved in the transport of the virulence factor cysteine protease in the enteric protozoan parasite Entamoeba histolytica.* Mol Biol Cell, 2005. **16**(11): p. 5294-303.

31.	Mitra, B.N., Y. Saito-Nakano, K. Nakada-Tsukui, D. Sato, and T. Nozaki, *Rab11B small GTPase regulates secretion of cysteine proteases in the enteric protozoan parasite Entamoeba histolytica.* Cell Microbiol, 2007. **9**(9): p. 2112-25.

32.	Anamika, K., A. Bhattacharya, and N. Srinivasan, *Analysis of the protein kinome of Entamoeba histolytica.* Proteins, 2008. **71**(2): p. 995-1006.

33.	Stols, L., et al., *A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site.* Protein Expr Purif, 2002. **25**(1): p. 8-15.

34.	Bosch, D.E., et al., *Structural determinants of affinity enhancement between GoLoco motifs and G-protein alpha subunit mutants.* J Biol Chem, 2010. **286**(5): p. 3351-8.

35.	Johnston, C.A., et al., *Structural determinants underlying the temperature-sensitive nature of a Galpha mutant in asymmetric cell division of Caenorhabditis elegans.* J Biol Chem, 2008. **283**(31): p. 21550-8.

36.	Rayment, I., *Reductive alkylation of lysine residues to alter crystallization properties of proteins.* Methods Enzymol, 1997. **276**: p. 171-9.

37.	Otwinowski, Z.a.W.M., *Processing of X-ray Diffraction Data Collected in Oscillation Mode*, in *Methods in Enzymology*. 1997, Academic Press: New York. p. 307-326.

38.   Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.

39.   Emsley, P., B. Lohkamp, W.G. Scott, and K. Cowtan, *Features and development of Coot.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.

40.   Painter, J. and E.A. Merritt, *Optimal description of a protein structure in terms of multiple groups undergoing TLS motion.* Acta Crystallogr D Biol Crystallogr, 2006. **62**(Pt 4): p. 439-50.

41.   Chen, V.B., et al., *MolProbity: all-atom structure validation for macromolecular crystallography.* Acta Crystallogr D Biol Crystallogr, 2009. **66**(Pt 1): p. 12-21.

42.   Kreutz, B., et al., *A new approach to producing functional G alpha subunits yields the activated and deactivated structures of G alpha(12/13) proteins.* Biochemistry, 2006. **45**(1): p. 167-74.

43.   Mixon, M.B., et al., *Tertiary and quaternary structural changes in Gi alpha 1 induced by GTP hydrolysis.* Science, 1995. **270**(5238): p. 954-60.

44.   Kim, H.S., G. Lee, S.W. John, N. Maeda, and O. Smithies, *Molecular phenotyping for analyzing subtle genetic effects in mice: application to an angiotensinogen gene titration.* Proc Natl Acad Sci U S A, 2002. **99**(7): p. 4602-7.

45.   Willard, F.S., et al., *Regulator of G-protein signaling 14 (RGS14) is a selective H-Ras effector.* PLoS One, 2009. **4**(3): p. e4884.

46.   Notredame, C., D.G. Higgins, and J. Heringa, *T-Coffee: A novel method for fast and accurate multiple sequence alignment.* J Mol Biol, 2000. **302**(1): p. 205-17.

47.   Temple, B.R., C.D. Jones, and A.M. Jones, *Evolution of a signaling nexus constrained by protein interfaces and conformational States.* PLoS Comput Biol, 2010. **6**(10): p. e1000962.

48.   Kimple, A.J., R.E. Muller, D.P. Siderovski, and F.S. Willard, *A capture coupling method for the covalent immobilization of hexahistidine tagged proteins for surface plasmon resonance.* Methods Mol Biol, 2010. **627**: p. 91-100.

49.   Hamann, L., H. Buss, and E. Tannich, *Tetracycline-controlled gene expression in Entamoeba histolytica.* Mol Biochem Parasitol, 1997. **84**(1): p. 83-91.

50.   Olvera, A., et al., *Stable transfection of Entamoeba histolytica trophozoites by lipofection.* Arch Med Res, 1997. **28 Spec No**: p. 49-51.

51.   Gilchrist, C.A., et al., *Targets of the Entamoeba histolytica transcription factor URE3-BP.* PLoS Negl Trop Dis, 2008. **2**(8): p. e282.

52.   Langmead, B., C. Trapnell, M. Pop, and S.L. Salzberg, *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.* Genome Biol, 2009. **10**(3): p. R25.

53.   Trapnell, C., et al., *Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks.* Nat Protoc, 2012. **7**(3): p. 562-78.

54.   Santi-Rocca, J., et al., *Endoplasmic reticulum stress-sensing mechanism is activated in Entamoeba histolytica upon treatment with nitric oxide.* PLoS One, 2012. **7**(2): p. e31777.

55.   Dolabella, S.S., et al., *Amoebic liver abscess production by Entamoeba dispar.* Ann Hepatol, 2012. **11**(1): p. 107-17.

56.   Hirata, K.K., et al., *A phagocytosis mutant of Entamoeba histolytica is less virulent due to deficient proteinase expression and release.* Exp Parasitol, 2007. **115**(2): p. 192-9.

57.   Wedegaertner, P.B., P.T. Wilson, and H.R. Bourne, *Lipid modifications of trimeric G proteins.* J Biol Chem, 1995. **270**(2): p. 503-6.

58.   Kerppola, T.K., *Bimolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells.* Annu Rev Biophys, 2008. **37**: p. 465-87.

59.   Berman, D.M., T.M. Wilkie, and A.G. Gilman, *GAIP and RGS4 are GTPase-activating proteins for the Gi subfamily of G protein alpha subunits.* Cell, 1996. **86**(3): p. 445-52.

60.   Soundararajan, M., et al., *Structural diversity in the RGS domain and its interaction with heterotrimeric G protein alpha-subunits.* Proc Natl Acad Sci U S A, 2008. **105**(17): p. 6457-62.

61.   Johnston, C.A., et al., *GTPase acceleration as the rate-limiting step in Arabidopsis G protein-coupled sugar signaling.* Proc Natl Acad Sci U S A, 2007. **104**(44): p. 17317-22.

62.   Coleman, D.E., et al., *Structures of active conformations of Gi alpha 1 and the mechanism of GTP hydrolysis.* Science, 1994. **265**(5177): p. 1405-12.

63.   Slepak, V.Z., et al., *Random mutagenesis of G protein alpha subunit G(o)alpha. Mutations altering nucleotide binding.* J Biol Chem, 1993. **268**(29): p. 21889-94.

64.   Tamura, K., et al., *MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods.* Mol Biol Evol, 2011. **28**(10): p. 2731-9.

65. Booker, K.S., J. Schwarz, M.B. Garrett, and A.M. Jones, *Glucose attenuation of auxin-mediated bimodality in lateral root formation is partly coupled by the heterotrimeric G protein complex.* PLoS One, 2010. **5**(9).

66. Nilson, S.E. and S.M. Assmann, *The alpha-subunit of the Arabidopsis heterotrimeric G protein, GPA1, is a regulator of transpiration efficiency.* Plant Physiol, 2010. **152**(4): p. 2067-77.

67. Pai, E.F., et al., *Structure of the guanine-nucleotide-binding domain of the Ha-ras oncogene product p21 in the triphosphate conformation.* Nature, 1989. **341**(6239): p. 209-14.

68. Wennerberg, K., K.L. Rossman, and C.J. Der, *The Ras superfamily at a glance.* J Cell Sci, 2005. **118**(Pt 5): p. 843-6.

69. Rasmussen, S.G., et al., *Crystal structure of the beta2 adrenergic receptor-Gs protein complex.* Nature, 2011. **477**(7366): p. 549-55.

70. Ralston, K.S. and W.A. Petri, Jr., *Tissue destruction and invasion by Entamoeba histolytica.* Trends Parasitol, 2011. **27**(6): p. 254-63.

71. Arhets, P., J.C. Olivo, P. Gounon, P. Sansonetti, and N. Guillen, *Virulence and functions of myosin II are inhibited by overexpression of light meromyosin in Entamoeba histolytica.* Mol Biol Cell, 1998. **9**(6): p. 1537-47.

72. Fang, D. and D. Shu, *Entamoeba histolytica liver abscess.* CMAJ, 2010. **182**(16): p. 1758.

73. Haque, R., C.D. Huston, M. Hughes, E. Houpt, and W.A. Petri, Jr., *Amebiasis.* N Engl J Med, 2003. **348**(16): p. 1565-73.

74. Saffer, L.D. and W.A. Petri, Jr., *Role of the galactose lectin of Entamoeba histolytica in adherence-dependent killing of mammalian cells.* Infect Immun, 1991. **59**(12): p. 4681-3.

75. Nakada-Tsukui, K., K. Tsuboi, A. Furukawa, Y. Yamada, and T. Nozaki, *A novel class of cysteine protease receptors that mediate lysosomal transport.* Cell Microbiol, 2012.

76. Gonzalez De la Rosa, C.H., et al., *EhGEF2, a Dbl-RhoGEF from Entamoeba histolytica has atypical biochemical properties and participates in essential cellular processes.* Mol Biochem Parasitol, 2007. **151**(1): p. 70-80.

77. Letunic, I., T. Doerks, and P. Bork, *SMART 7: recent updates to the protein domain annotation resource.* Nucleic Acids Res, 2012. **40**(Database issue): p. D302-5.

78.   Piasecki, B.P., J. Burghoorn, and P. Swoboda, *Regulatory Factor X (RFX)-mediated transcriptional rewiring of ciliary genes in animals.* Proc Natl Acad Sci U S A, 2010. **107**(29): p. 12969-74.

79.   Meza, I., P. Talamas-Rohana, and M.A. Vargas, *The cytoskeleton of Entamoeba histolytica: structure, function, and regulation by signaling pathways.* Arch Med Res, 2006. **37**(2): p. 234-43.

80.   Picazarri, K., J.P. Luna-Arias, E. Carrillo, E. Orozco, and M.A. Rodriguez, *Entamoeba histolytica: identification of EhGPCR-1, a novel putative G protein-coupled receptor that binds to EhRabB.* Exp Parasitol, 2005. **110**(3): p. 253-8.

81.   Pan, C.L., et al., *C. elegans AP-2 and retromer control Wnt signaling by regulating mig-14/Wntless.* Dev Cell, 2008. **14**(1): p. 132-9.

82.   Soid-Raggi, L.G., M.E. Torres-Marquez, and I. Meza, *Entamoeba histolytica: identification of functional Gs and Gi proteins as possible signal transduction elements in the interaction of trophozoites with fibronectin.* Exp Parasitol, 1998. **90**(3): p. 262-9.

CHAPTER 3

STRUCTURAL DETERMINANTS OF RGS-RhoGEF SIGNALING CRITICAL

TO Entamoeba histolytica PATHOGENESIS[1]

3.1   OVERVIEW

G-protein signaling pathways, as key components of physiologic responsiveness and timing, are frequent targets for pharmacologic intervention. Here, we identify an effector for heterotrimeric G-protein α subunit (EhGα1) signaling from *Entamoeba histolytica*, the causative agent of amoebic colitis. EhGα1 interacts with this effector and GTPase-accelerating protein (GAP), EhRGS-RhoGEF, in a nucleotide state-selective fashion. Co-expression of EhRGS-RhoGEF with constitutively active EhGα1 and EhRacC leads to Rac-dependent spreading in *Drosophila* S2 cells. EhRGS-RhoGEF overexpression in *E. histolytica* trophozoites leads to reduced migration toward serum and lower cysteine protease activity, as well as reduced attachment to, and killing of, host cells. A 2.3 Å crystal structure of the full-length EhRGS-RhoGEF reveals a putative inhibitory helix engaging the DH domain Rho-binding surface and the PH domain. Mutational analysis of the EhGα1/EhRGS-RhoGEF interface confirms a canonical RGS domain rather than a RhoGEF-RGS ("rgRGS") domain*,* suggesting a convergent evolution toward heterotrimeric and small G-protein cross-talk.

[1] Bosch, D.E., Kimple, A.J., Manning, A.J., Muller, R.E., Willard, F.S., Machius, M., Rogers, S.L., and Siderovski, D.P. (2012) Structural Determinants of RGS-RhoGEF Signaling Critical to *Entamoeba histolytica* Pathogenesis. *Structure*. 21(1):65-75.

3.2   INTRODUCTION

Heterotrimeric G-protein signaling pathways are frequent targets for pharmacologic manipulation [1]. The Gα subunit in its inactive, GDP-bound conformation engages the obligate G$\beta\gamma$ dimer [2]. A ligand-activated seven-transmembrane <u>G</u>-<u>p</u>rotein <u>c</u>oupled <u>r</u>eceptor (GPCR) promotes nucleotide exchange on the Gα subunit. Upon binding of GTP by the Gα subunit, three switch regions undergo a conformational change, leading to separation from G$\beta\gamma$ and subsequent activation of downstream effectors such as adenylyl cyclases, phospholipase C, and Rho-family guanine nucleotide exchange factors (RhoGEFs) [2, 3]. Signaling is terminated by the GTPase activity of the Gα subunit and reassembly of the Gα·GDP/G$\beta\gamma$ heterotrimer.

‘Regulator of G-protein signaling’ (RGS) proteins accelerate the intrinsic GTP hydrolysis activity of Gα subunits and thereby serve as negative regulators of signaling [4]. Canonical nine-helix RGS domains exhibit highest affinity for Gα in its transition-state mimetic form, stabilizing the switch regions for efficient hydrolysis [5]. Members of the RGS-RhoGEF family of Gα effectors contain N-termini with similarity to RGS domains (called ‘RhoGEF-RGS’ or ‘rgRGS’ domains), in combination with Dbl homology (DH) and pleckstrin homology (PH) domains that together mediate activation of Rho family GTPases [3]. The DH domain engages substrate Rho GTPases, promoting nucleotide release, while the PH domain frequently modulates exchange in various Dbl-family RhoGEF members [6]. In contrast to 9-helix RGS domains, rgRGS domains have a distinct 12-helix fold and engage G$\alpha_{12/13}$ subunits through an effector-like interface involving primarily switch 2 and the α3 helix [3]. An N-terminal extension of the rgRGS domain containing an "IIG" sequence motif contacts the G$\alpha_{12/13}$ switch regions and all-helical domain and, in the case of p115 RhoGEF,

is required for GTPase accelerating protein (GAP) activity toward $G\alpha_{12/13}$ subunits. Although structures of RGS-RhoGEFs with both the rgRGS and DH-PH domain tandems have not yet been elucidated, p115 RhoGEF is thought to be activated by an allosteric "GEF switch" mechanism that involves a conformational change of an N-terminal extension of the DH domain [7]. Recent low-resolution structural studies of full-length p115 RhoGEF suggest an elongated domain architecture and a potential bimodal interaction with $G\alpha13$, namely the effector interface with the rgRGS domain and a potential additional interface with the DH domain [8]. The activation mechanism of another mammalian RGS-RhoGEF, PDZ-RhoGEF, is thought to be complex, involving disruption of an electrostatic RGS-DH linker and DH domain interaction, perturbation of a putative RGS-DH linker "molten globule", and membrane recruitment, as well as a "GEF switch" [9, 10]. Low-resolution SAXS studies of PDZ-RhoGEF also suggested an ensemble of elongated domain architectures [10].

Entamoeba histolytica causes an estimated 50 million infections and 100,000 deaths per year worldwide, with highest incidence in countries with poor barriers between drinking water and sewage [11]. Following ingestion of encysted *E. histolytica*, the trophozoite, or amoeboid form of the parasite attaches to and destroys intestinal epithelial cells, giving rise to amoebic colitis [12]. Although the molecular details of signaling pathways in *E. histolytica* remain understudied, a relatively large family of Rho GTPases and Dbl family RhoGEFs are known to modulate cytoskeletal dynamics, as well as key pathogenic processes such as trophozoite migration, extracellular matrix invasion, and killing and phagocytosis of host cells [13-15].

Recently, we identified a functional heterotrimeric G-protein signaling pathway in *E. histolytica*; perturbation of the $G\alpha$ subunit, $EhG\alpha1$, elucidated positive regulatory roles for

G-protein signaling in pathogenic processes such as trophozoite migration and invasion, host cell attachment, and cell killing [16]. Overexpression of either wild type EhGα1 or a dominant negative mutant in *E. histolytica* trophozoites resulted in altered transcription of numerous genes that have implicated multiple potential mechanisms by which G-protein signaling modulates pathogenesis. From this study, a set of Rho GTPase signaling proteins, including an RGS-RhoGEF, and actin-associated proteins was observed to be differentially transcribed downstream of heterotrimeric G-proteins (Table 3.1). We hypothesized that EhRGS-RhoGEF, like its mammalian homologs, serves as an EhGα1 effector and signals through Rho family GTPases, with important roles in *Entamoeba histolytica* pathogenesis. Here, our results describe EhRGS-RhoGEF as an effector and GTPase accelerating protein (GAP) of EhGα1, with importance for *E. histolytica* motility, host cell attachment, cell killing, and cysteine protease secretion. Activation of EhRGS-RhoGEF by co-expression of constitutively active EhGα1 and EhRacC leads to Rac family GTPase-dependent cell spreading in *Drosophila* S2 cells. Furthermore, we provide a crystal structure of the full-length RGS-RhoGEF in the inactive state, elucidating its molecular architecture and likely distinct evolutionary origin relative to the mammalian RGS-RhoGEFs.

3.3   EXPERIMENTAL PROCEDURES

3.3.1 Cloning and protein purification

The open reading frames (ORFs) of EhGα1, EhRGS-RhoGEF, and EhRacC were amplified from *E. histolytica* genomic DNA (a gift from Dr. W.A. Petri, U. of Virginia) by polymerase chain-reaction (PCR) using Phusion polymerase (New England BioLabs) and Invitrogen primers. Amplicons were subcloned using ligation-independent cloning [17] into a Novagen

pET vector-based prokaryotic expression construct ("pET-His-LIC-C") to form N-terminal tobacco etch virus (TEV) protease-cleavable, hexahistidine-tagged fusions. Mutations were made using QuikChange site-directed mutagenesis (Stratagene). For hexahistidine-tagged EhGα1, EhRGS-RhoGEF, and EhRacC, B834 *E. coli* were grown to an OD_{600nm} of 0.7−0.8 at 37°C before induction with 700 μM, 1 mM, or 500 μM isopropyl-β-D-thiogalactopyranoside (IPTG), respectively, for 14-16 hours at 20°C. Cell pellets were resuspended in N1 buffer (for EhGα1: 50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 10 mM $MgCl_2$, 10 mM NaF, 30 μM $AlCl_3$, 50 μM GDP, 30 mM imidazole, 5% (w/v) glycerol; for EhRGS-RhoGEF: 50 mM HEPES pH 8.0, 150 mM NaCl, 30 mM imidazole, 5% (w/v) glycerol; for EhRacC: 50 mM HEPES pH 7.5, 250 mM NaCl, 5 mM $MgCl_2$, 50 μM GDP or GTP, 30 mM imidazole, and 5% (w/v) glycerol) and lysed at 10,000 kPa using an Emulsiflex (Avestin, Ottowa, Canada). Cleared lysates were applied to nickel-nitrilotriacetic acid resin (GE Healthcare), washed, and eluted with N1 buffer containing 300 mM imidazole. Eluted protein was resolved using a calibrated size exclusion column (GE Healthcare) with S200 buffer for EhGα1 (50 mM Tris-HCl pH 8.0, 250 mM NaCl, 5 mM DTT, 5% (w/v) glycerol, and 50 μM GDP), EhRGS-RhoGEF (50 mM Tris-HCl pH 8.0, 250 mM NaCl, 5 mM DTT, 5% (w/v) glycerol), or EhRacC (50 mM Tris pH 8.0, 200 mM NaCl, 5 mM $MgCl_2$, and 5% (w/v) glycerol). Comparison of global fold between wild type and E39K mutant EhRGS-RhoGEF was performed using circular dichroism as previously described [18]. For crystallization, two residues were removed from both the N- and C-terminus of EhRGS-RhoGEF, and the hexahistidine tag cleaved with TEV protease prior to size exclusion chromatography.

For mammalian HEK293T and *Drosophila* S2 cellular experiments, open reading frames of EhRGS-RhoGEF, EhGα1, and EhRacC were synthesized with codon optimization for mammalian cell expression and subcloned into pcDNA3.1 or pMT/V5-His vectors (Life Technologies). EhRGS-RhoGEF and EhRacC were tagged with N-terminal HA and FLAG epitopes, respectively, and an internal FLAG tag was inserted after His-83 in EhGα1. *D.m.* Rac1(G14V) with an N-terminal myc epitope tag was also subcloned into the pMT/V5-His vector.

3.3.2 Crystallization and structure determination

Crystals of the full-length EhRGS-RhoGEF (a.a. 1-519) yielded diffraction data not suitable for either molecular replacement or anomalous dispersion. However, by removing two residues on both the N- and C-termini of EhRGS-RhoGEF we obtained another crystal form with improved diffraction quality, ultimately allowing structure determination by SAD. Crystallization was achieved by hanging drop vapor diffusion at 18°C. EhRGS-RhoGEF at 15 mg/mL in crystallization buffer (50 mM HEPES pH 7.5, 100 mM NaCl, and 1 mM DTT) was mixed 1:1 and equilibrated against crystallization solution containing 16% (w/v) PEG 3350 and 100 mM sodium malonate pH 5.0. Hexagonal plate crystals grew to 400 x 150 x 20 μm over 5 days. EhRGS-RhoGEF crystals displayed the symmetry of space group C2 (a = 86.1 Å, b = 46.3 Å, c = 142.6 Å, α = γ = 90°, β = 104.2°), with one monomer in the asymmetric unit. Prior to data collection, crystals were cryoprotected in crystallization solution supplemented with 25% (v/v) glycerol.

Anomalous diffraction data were obtained at 0.97954 Å wavelength (selenium absorption peak) and 100K temperature at the GM/CA-CAT ID-D beamline (APS, Argonne

National Labs) and processed using HKL2000 [19]. A highly redundant data set was obtained combining partial data sets from five points along the EhRGS-RhoGEF plate crystal. Heavy atom searching identified 8 of 8 possible sites, and refinement yielded an estimated Bayes correlation coefficient of 48.2 to 2.5 Å resolution. After density modification, the estimated Bayes correlation coefficient increased to 66.0. ~75% of the model was constructed automatically, and the remaining portion was built manually. The current model (Table 3.2) contains one EhRGS-RhoGEF monomer.

Refinement was carried out against peak anomalous data with Bijvoet pairs kept separate using phenix.refine [20] interspersed with manual model revisions using the program Coot [21] and consisted of conjugate-gradient minimization and calculation of individual atomic displacement and translation/libration/screw (TLS) parameters [22]. Residues that could not be identified in the electron density were: 1, 139-140, 153-156, and 452-454. The model exhibits excellent geometry as determined by MolProbity [23]. A Ramachandran analysis identified 97.6% favored, 2.4% allowed, and 0% disallowed residues. Coordinates and structure factors are deposited in the RCSB Protein Data Bank (id 4GOU).

3.3.3 Single turnover nucleotide hydrolysis

GTP hydrolysis by single turnover assays was quantified as previously described [24]. For GTPase acceleration assays, increasing concentrations of purified EhRGS-RhoGEF were added along with the hydrolysis-initiating magnesium.

3.3.4 Surface plasmon resonance

Optical detection of protein binding was conducted as described previously [25]. Briefly, His$_6$-tagged wild type or mutant EhRGS-RhoGEF was immobilized on an NTA chip surface by capture coupling and increasing concentrations of wildtype EhGα1 and mutants were flowed over at 10 μL/s in various nucleotide states. In complementary experiments, GST-EhGα1 was immobilized on an anti-GST chip surface, as described [26], and increasing concentrations of EhRGS-RhoGEF and mutants flowed over.

3.3.5 NTA affinity co-preciptiation

HEK293T cells were transfected with DNA encoding an HA epitope-tagged EhRGS-RhoGEF RGS domain (a.a. 1-152), codon optimized for mammalian cell expression. Cells were plated in 6-well dishes and transfected with 1.5 μg of DNA and Lipofectamine 2000 (Invitrogen). Cells were lysed in buffer containing 20 mM HEPES pH 7.5, 100 mM NaCl, 5 mM MgCl$_2$, 1 mM EDTA, 1% NP-40 alternative (Calbiochem), 10% deoxycholate, and 5 mM imidazole. The lysis buffer was supplemented with 10 μM GTP; 10 μM GDP; or 10 μM GDP, 10 mM NaF, and 30 μM AlCl$_3$. Cellular lysates were cleared by high-speed centrifugation at 4°C for 20 minutes. Purified His$_6$-EhGα1 (50 μg) was added to the cleared cell lysate and rocked at 4°C for 30 minutes prior to addition of NTA-agarose and continued rocking for 4 hours. NTA-agarose was pelleted and washed three times with lysis buffer, and bound proteins resolved by SDS-PAGE, transferred to nitrocellulose, and detected by α-HA immuno-blotting or Ponceau S staining.

3.3.6 Trophozoite stable transfection

EhRGS-RhoGEF was subcloned with an N-terminal FLAG epitope tag into an *E. histolytica* expression vector, described previously [27]. Axenic cultures were transfected by lipofection as previously described [28]. Briefly, amoebae at ~5 x 10^6/mL were suspended in medium 199 (Sigma) supplemented with 5.7 mM cysteine, 1 mM ascorbic acid, 25 mM HEPES (pH 6.9), 15 μg of DNA, and 30 μL of Superfect (Qiagen). After 3 hours at 37 °C, trophozoites were transferred to TYI-S-33 medium overnight and selected for stable transfection with 10 μg/mL hygromycin over 3 weeks.

3.3.7 Chemotactic migration

Trophozoite migration assays were performed essentially as described previously [29]. Briefly, amoebae harvested in log growth phase were suspended in serum free TYI growth medium, and 50,000 cells loaded in the upper chamber of a Transwell migration chamber (Costar, 8 μm pore size). The lower chamber contained growth medium with or without 15% adult bovine serum. Transwell plates were incubated at 37°C for 2 hr under anaerobic conditions (GasPak EZ, BD Biosciences).

3.3.8 Host cell attachment

Attachment of *E. histolytica* trophozoites to epithelial monolayers was assessed as previously described [30]. Chinese hamster ovary (CHO) cells were grown to confluency in 24-well plates, washed, and fixed in 4% paraformaldehyde for 30 minutes. Trophozoites (3 x 10^5) were added to the fixed monolayers in medium 199 supplemented with 5.7 mM cysteine, 1 mM ascorbic acid, and 25 mM HEPES (pH 6.9). After incubation at 37°C for 30 minutes,

each well was washed gently two times with warm PBS to remove unattached trophozoites. Monolayer-attached trophozoites were quantified by counting with an inverted microscope. Each experiment was performed in quadruplicate and statistical significance determined by Student's t-test.

3.3.9 Cell killing

Killing of mammalian cells (Jurkat) was assessed using the CytoTox-ONE membrane integrity assay (Promega). In 96-well plates, 5×10^5 Jurkat cells and/or 2.5×10^5 trophozoites were incubated at 37°C in 200 μL of medium 199 (Sigma) supplemented with 5.7 mM cysteine, 0.5% BSA, and 25 mM HEPES pH 6.8. After 2 hr, 50 μL of medium from each well was incubated with Cytotox reagent and a colorimetric measure of extracellular lactate dehydrogenase activity was obtained after 10 min. Each experiment was performed in quadruplicate and statistical significance determined by Student's t-test.

3.3.10 Cysteine protease activity

Secreted cysteine protease activity in amoebic lysates was assayed essentially as described previously [31]. 10^6 trophzoites were incubated at 37 °C for 3 hr in 500 μL PBS supplemented with 20 mM cysteine, 0.15 mM $CaCl_2$, and 0.5 mM $MgCl_2$, conditions known to sustain *E. histolytica* growth and allow cysteine protease secretion [32]. Following centrifugation, the cell-free conditioned medium was assayed for cysteine protease activity. 2 mg of azo dye-impregnated collagen (Sigma) with conditioned medium in 500 μL of protease activation buffer (100 mM Tris pH 7.0 and 10 mM $CaCl_2$) were incubated at 37°C for 18 hr, then terminated with 500 μL of 10% TCA. Samples were centrifuged to exclude

intact collagen fibers, and supernatants collected for absorbance reading at 540 nm. In parallel experiments, the inhibitor p-hydroxy-mercuribenzoic acid (PHMB) was included at 1 mM to assess the fraction of specific cysteine protease activity. Residual protease activity (after PHMB treatment) was subtracted to determine total cysteine protease activity.

3.3.11 S2 cell culture and spreading assay

S2 cells were obtained from the Drosophila Genome Resource Center (Bloomington, IL), and cultivated as described previously [33]. S2 cells were maintained in SF900 SFM (Invitrogen, Carsbad, CA) and transfected with 2 μg total DNA using the Amaxa nucleofector system (Lonza, Basel, Switzerland). Expression of transfected constructs was induced with 35 μM CuSO$_4$. Double stranded RNAs were produced using Promega (Madison, WI) Ribomax T7 kit according to manufacturer instructions. S2 cells at 50-90% confluency in 6-well plates were treated every other day for 7 days with 7.5 μg/ml of dsRNA. On day 5 of RNAi treatment, cells were transfected as above and then induced on day 6. Cells were resuspended and plated on poly-L-lysine (Sigma-Aldrich, St. Louis, MO) coated coverslips and allowed to spread for 1 hour. For quantifying numbers of cells with spreading, each condition was repeated at least three times and ≥100 cells were counted per experiment.

Double stranded RNAs were produced using Promega (Madison, WI) Ribomax T7 kit according to manufacturer instructions. Primers used for dsRNA synthesis are as follows and are all preceded by the T7 sequence (5'-TAATACGACTCACTATAGG-3'). Control-fwd: 5'-TAAATTGTAAGCGTTAATATTTTG-3' and Control-rev: 5'-AATTCGATATCAAGCTTATCGAT-3' to amplify a region from the pBluescript plasmid; Rho-fwd: 5'-GTAAAACTTGCCTTCTGATTGTCT-3', Rho-rev: 5'-

115

ATCTGGTCTTCTTCCTCTTTTTGA-3';   Rac1/2-fwd: 5'-CTGATCAGCTACACGACCAAT-3', Rac1/2-rev: 5'-CGAGCACTCCAGATACTTGAC-3'; mtl-fwd: 5'-ATGTCAACCGGAAGGCCCATAAAG-3', mtl-rev: 5'-TTACATTATTAAACACTTTCG-3'.

3.3.12 Immunofluorescence microscopy

S2 cells were fixed with 4% formaldehyde (EM Sciences, Gibbstown, NJ) in HL3 buffer (70 mM NaCl; 5 mM KCl; 1.5 mM $CaCl_2$-$2H_2O$; 20 mM $MgCl_2$-$6H_2O$; 10 mM $NaHCO_3$; 5 mM trehalose; 115 mM sucrose; 5 mM HEPES; pH to 7.2), and permeabilized with PBST (PBS + 0.1% Triton X-100). Cells were blocked with 5% normal goat serum (Sigma-Aldrich, St. Louis, MO) in PBST and stained with primary antibodies diluted into blocking solution at 1:500. Following washing, cells were incubated with secondary antibodies (1:1000 dilution blocking solution) and Alexa 488-phalloidin (1:100 dilution; Invitrogen). After washing, the cells were mounted in fluorescence mounting medium (Dakocytomation, Glostrup, Denmark). Antibodies are as follows: mouse anti-myc (Developmental Studies Hybridoma Bank, Iowa City, Iowa), mouse anti-FLAG and rabbit anti-HA (Sigma-Aldrich), RhodamineX-goat anti-rabbit, Cy5-goat anti-mouse, and DyLight 549–goat anti-mouse (Jackson Immunoresearch, West Grove, PA). Images were acquired using a CoolSnap HQ CCD camera (Roper Scientific, Ottobrunn, Germany) on a Nikon Eclipse Ti inverted microscope driven by Nikon Elements software (Tokyo, Japan).

3.4 RESULTS

3.4.1 EhGα1 engages an RGS-RhoGEF effector and GTPase accelerating protein

The *E. histolytica* genome encodes a single classical Gα subunit effector, an RGS domain-containing RhoGEF (GenBank XP_653063; named EhRGS-RhoGEF) with distant similarity to the RGS-RhoGEF effectors of mammalian Gα$_{12/13}$ subunits (Fig. 3.1). Transcription of EhRGS-RhoGEF was seen to be up-regulated upon overexpression of EhGα1 in *E. histolytica* trophozoites, suggesting a functional link to heterotrimeric G-protein signaling (Table 3.1). EhRGS-RhoGEF was purified from *E. coli* (Fig. 3.2) and found to bind directly to EhGα1 selectively in its GDP·AlF$_4^-$ (AMF) nucleotide state as measured by surface plasmon resonance (Fig. 3.3 A-D). Reciprocal immobilization experiments each indicated an ~5 μM EhGα1·AMF/EhRGS-RhoGEF dissociation constant (K$_D$). EhRGS-RhoGEF also interacted with the constitutively active, GTPase deficient EhGα1(Q189L) mutant in its GTP-bound form, although with ~20-fold lower affinity than for EhGα1·AMF (K$_D$ ≈ 110 μM) (Fig. 3.3B). The EhGα1(Q189L)·GTP/EhRGS-RhoGEF binding affinity could not be precisely determined by equilibrium binding analysis due to concentration limitations of our assay. The preference of EhRGS-RhoGEF for the AMF nucleotide state over the GTP-bound state of EhGα1 is consistent with a similar order-of-magnitude difference in p115 RhoGEF affinity for Gα$_{i1/13}$·AMF compared to Gα$_{i1/13}$·GTPγS [34]. To determine whether the interaction occurs through the RGS domain of EhRGS-RhoGEF, a conserved glutamate at the predicted Gα subunit-binding surface was mutated to lysine (E39K; Fig. 3.1). Despite proper global folding of the E39K mutant (Fig. 3.2B), it exhibited drastically reduced affinity for EhGα1·AMF (Fig. 3.3D). The isolated RGS domain of EhRGS-RhoGEF could not be produced from *E. coli*. However, RGS domain expressed in

HEK 293T cells was seen to specifically co-precipitate with purified EhGα1·AMF and EhGα1(Q189L)·GTP, but not EhGα1·GDP (Fig. 3.2C), suggesting that the RGS domain alone is sufficient to bind EhGα1.

The relatively high affinity of wild type EhRGS-RhoGEF for the hydrolysis transition state-mimetic (AMF) form of EhGα1 suggested that the RGS domain of EhRGS-RhoGEF may serve as a GTPase accelerating protein (GAP) for EhGα1. Interestingly, EhRGS-RhoGEF lacks the N-terminal extension containing the IIG motif that is required for the GAP activity of p115 RhoGEF (Fig. 3.1) [3]. However, single turnover GTP hydrolysis assays demonstrated that wild type EhRGS-RhoGEF, but not the EhRGS-RhoGEF(E39K) mutant, accelerates the intrinsic GTP hydrolysis activity of EhGα1 in a concentration-dependent fashion (Fig. 3.3E). Mutation of the conserved Asn-83 in the EhRGS-RhoGEF RGS domain, predicted to contact and orient EhGα1 switch residues for efficient catalysis of GTP hydrolysis [5], also eliminated GAP activity (Fig. 3.3F) and dramatically reduced EhRGS-RhoGEF/EhGα1 binding affinity (Fig. 3.3C).

3.4.2 EhGα1 and EhRacC activate EhRGS-RhoGEF to promote Rho-dependent cell spreading

To determine the ability of EhRGS-RhoGEF to modulate Rho-dependent cellular processes and its potential regulation by EhGα1, we utilized *Drosophila* S2 cells that undergo a dramatic morphological transition with distinctive reorganization of actin structures when Rho family GTPases are activated by various stimuli [35]. For example, overexpression of a GTPase-deficient and constitutively active *D.m.* Rac1(G14V) lead to cell spreading in ~80% of S2 cells on a poly-lysine coated surface, as compared to ~20% of cells

expressing RFP (Figure 3.4A,B). Expression of constitutively active EhGα1(Q189L) and EhRGS-RhoGEF was not sufficient to significantly increase cell spreading. However, additional expression of constitutively active EhRacC(Q65L) lead to spreading, suggesting that GTP-bound EhGα1 and EhRacC are necessary to activate EhRGS-RhoGEF. Importantly, EhRacC(Q65L) alone (not shown) or in combination with EhGα1 (Figure 3.4) did not trigger S2 cell spreading, indicating its inability to productively engage the endogenous *D.m.* Rho GTPase signaling machinery independently of EhRGS-RhoGEF. Wild type EhGα1 did not activate EhRGS-RhoGEF, and the EhRGS-RhoGEF(E39K) mutation prevented an increase cell spreading (Fig. 3.4A,B). Thus, co-expression of EhRGS-RhoGEF with constitutively active EhRacC and interaction with constitutively active EhGα1 at the RGS domain are required for enhanced cell spreading.

The observed cell spreading phenotype strongly suggested that EhRGS-RhoGEF was signaling through endogenous *Drosophila* Rho family GTPases. To investigate the dependence of the observed phenotype on endogenous Rho family GTPases, we knocked down expression of *D.m.* Rac1/2 (both isoforms targeted), Rho, and mtl by RNA interference, as demonstrated previously in S2 cells [33]. Specific knock down of *D.m.* Rac GTPases abolished the cell spreading effect of EhRGS-RhoGEF, EhGα1(Q189L), and EhRacC(Q65L), suggesting that either or both *D.m.* Rac isoforms may serve as substrates for overexpressed EhRGS-RhoGEF (Figure 3.4A).

To test whether activated EhRacC directly engages EhRGS-RhoGEF, surface plasmon resonance was utilized. EhRGS-RhoGEF selectively bound EhRacC(Q65L)·GTP, to the exclusion of EhRacC·GDP or nucleotide-free EhRacC (Figure 3.4D). This pattern of nucleotide state selectivity was consistent with a Rho GTPase and effector-like interaction,

rather than a RhoGEF and substrate Rho GTPase relationship. Although the observed recombinant EhRacC(Q65L)/EhRGS-RhoGEF affinity was relatively low ($K_D \approx 34 \pm 9$ μM), the cell-spreading experiments suggest that a productive interaction occurs in a cellular context. The interaction of EhRGS-RhoGEF with an active species of Rho family GTPase was reminiscent of the structurally elucidated interaction between human RhoA·GTPγS and a hydrophobic patch on the PH domain of PDZ-RhoGEF [36]. An analysis of the PDZ-RhoGEF PH domain residues involved in activated RhoA binding (PDB id 3KZ1) revealed poor conservation with the corresponding PH domain residues of EhRGS-RhoGEF (19% identity, 25% similarity); thus, we do not predict that EhRacC·GTP binds EhRGS-RhoGEF in a similar fashion. However, a direct interaction does occur between activated EhRacC and EhRGS-RhoGEF, potentially explaining the required co-expression of EhRacC(Q65L), together with EhGα1(Q189L) and EhRGS-RhoGEF to enhance cell spreading (Figure 3.4A,B).

3.4.3 EhRGS-RhoGEF modulates pathogenic processes of *E. histolytica* trophozoites

We next investigated the role(s) of RGS-RhoGEF signaling in *E. histolytica* trophozoites by engineering the virulent HM-1:IMSS strain to stably overexpress wild type EhRGS-RhoGEF (Fig. 3.5). We focused on measuring the effect of EhRGS-RhoGEF overexpression on trophozoite chemotactic migration, host cell attachment, and cell killing, given the known dependence of these vital pathogenic processes on actin cytoskeletal dynamics, as well as Rho GTPases and Dbl family RhoGEFs [13, 15]. Trophozoites ectopically overexpressing EhRGS-RhoGEF killed host cells less efficiently than the parent strain, as indicated by a membrane integrity assay (Fig. 3.5A). A number of cytotoxic

proteins are involved in host cell killing, including membrane-perforating amoebapores and cysteine proteases [12]. Reduced secretion of active cysteine proteases, as measured by an azo-collagen assay (Fig. 3.5B), may account in part for the impaired cell killing of the EhRGS-RhoGEF-expressing trophzoite strain. Direct attachment of *E. histolytica* trophozoites to host epithelial cells, primarily through a galactose-inhibitable lectin [37], is also required for tissue destruction. Amoebae overexpressing EhRGS-RhoGEF showed reduced attachment to Chinese Hamster Ovary (CHO) cell monolayers compared to the parent strain (Fig. 3.5D). *E. histolytica* trophozoites are also highly motile, a feature that is dependent on a dynamic actin cytoskeleton regulated by Rho family GTPase signaling [15]. Transwell migration experiments indicated that overexpression of EhRGS-RhoGEF also decreases trophozoite chemotactic migration toward serum, but not random migration (Figure 3.5C), suggesting that interfering with the EhGα1/EhRGS-RhoGEF signaling axis modulates the *E. histolytica* migratory response to serum factors.

3.4.4 A crystal structure of EhRGS-RhoGEF

While isolated domains from mammalian RGS-RhoGEFs have been structurally characterized [3], a high-resolution structure of an RGS domain together with a DH-PH domain tandem has not been elucidated to date. We obtained a crystal structure of a nearly full-length EhRGS-RhoGEF protein (lacking only two residues from each terminus) to 2.3 Å by single wavelength anomalous dispersion (SAD) (Table 3.2). The structural model exhibits an N-terminal, canonical nine-helix RGS domain, an oblong DH domain, and a C-terminal PH domain (Figs. 3.6A). The RGS domain interacts with the DH domain surface opposite from the PH domain and the putative Rho GTPase binding site (Fig. 3.6). The linker between

the RGS and DH domains wraps around the oblong helical bundle of the DH domain (Fig. 3.6A), forming an additional helix (termed the 'inhibitory helix').

The RGS domain of EhRGS-RhoGEF is involved in multiple crystal contacts (Fig. 3.7); specifically, the putative Gα-binding surfaces of neighboring EhRGS-RhoGEF molecules in the crystal lattice contact one another. Although crystal contacts may modestly affect the disposition of the RGS domain, the similarity of its conformation to mammalian RGS domains in both crystallographic and NMR studies (Fig. 3.8B) [38], and significant burial of hydrophobic surface area (~850 Å^2) at the RGS-DH domain interface suggest that the crystal structure architecture accurately reflects that of EhRGS-RhoGEF in solution.

The PH domain exhibits a conserved overall fold despite weak sequence similarity (2.9 Å r.m.s.d. compared to 324 equivalent residues of the Dbs PH domain with only 51% sequence similarity) (Fig. 3.9A). An analysis of protein sequence motifs and comparison of the EhRGS-RhoGEF PH domain structure to other PH domains in complex with phospholipid head groups [39] revealed poor conservation of a potential phospholipid-binding site on EhRGS-RhoGEF. Thus, we do not hypothesize that the EhRGS-RhoGEF PH domain directly associates with phospholipids.

The EhRGS-RhoGEF DH domain is most similar to that of Intersectin (Dali server Z-score 19.4; PDB id 1KI1). Superposition of Intersectin/Cdc42 [40] and EhRGS-RhoGEF highlights a number of DH domain structural differences (Fig. 3.6B). The α6 helix of EhRGS-RhoGEF, which is the longest of the Intersectin DH domain, is disrupted by a loop, giving rise to two helices at ~90° relative orientations (termed α6a and α6b). The PH domain adopts a very different orientation relative to the DH domain in EhRGS-RhoGEF as compared to Intersectin (Fig. 3.6B). The PH domain of RGS-RhoGEF directly obstructs the

122

putative Rho binding site, similar to a number of mammalian RhoGEFs, *e.g.* Vav and Sos [41]. The DH and PH domains of EhRGS-RhoGEF share a substantial interface (~1200 Å^2 buried surface area) that occurs predominantly through hydrophobic interactions between the α7 helix of the PH domain and the α3d, α4, α5, and α6b helices of the DH domain (Figure 3.9B). Particularly, the hydrophobic side chains of Phe-393 and Met-397 project into an approximately triangular concavity formed by helices α3d, α4, and α5 (Figure 3.9B). The nature of the DH/PH domain interface suggests that the structural relationship between the two domains observed in the crystal structure likely also exists in solution. However, additional contacts of each domain with the inhibitory helix may also be necessary to maintain the observed DH and PH domain relationship (Figure 3.10), and alternative conformations are also possible.

3.4.5 The inhibitory helix coordinates occlusion of the Rho GTPase binding site

In the inactive state of EhRGS-RhoGEF, the inhibitory helix, the α6b helix of the DH domain, and the PH domain all obstruct the presumptive Rho GTPase interaction surface of the DH domain (Fig. 3.6), as predicted based on comparison with the Intersectin/Cdc42 structure [40]. In fact, the entire RGS-DH domain linker inhibitory helix lies within the space occupied by the Rho GTPase substrate in numerous, well-conserved Dbl family GEF/Rho interactions [6]. The inhibitory helix interacts with both the DH and PH domains through a series of hydrophobic and polar interactions (Fig. 3.10, 3.11). The hydrophobic residues Leu-164, Ile-167, Ile-168, and Trp-175 interface with a hydrophobic patch at the DH α6B helix/PH domain interface, consisting primarily of the hydrophobic portion of Lys-386, Leu-387, Ile-406, and Ile-450 (Fig. 3.10, 3.11). Surrounding the hydrophobic patch are a number

of apparent polar and ionic interactions, including those between Lys-161 of the inhibitory helix and Glu-511 of the PH domain as well as Lys-172 and Asp-383 of the inhibitory and DH domain α6B helices, respectively. The inhibitory helix residues Lys-166 and Ile-170 also form limited contacts with a DH domain loop from a neighboring molecule in the crystal lattice (Figure 3.7C), but these contacts likely do not contribute to the observed main chain conformation. Notably, the RGS-DH linker containing the putative inhibitory helix is much shorter in EhRGS-RhoGEF (26 residues) than the corresponding linker in its mammalian homologs, with p115 RhoGEF possessing the next shortest linker at 164 residues (Fig. 3.1C). Thus, it is likely that this region exhibits different structural features and potentially performs different functions in mammalian RGS-RhoGEFs.

3.4.6 Convergent evolution of the EhGα1/EhRGS-RhoGEF interface

The RGS/DH domain interface consists of a central hydrophobic region with peripheral hydrogen bond and ionic interactions (Fig. 3.8A, 3.11C). The residues corresponding to this domain interface are not highly conserved among mammalian RGS-RhoGEFs, such as p115 RhoGEF (Fig. 3.1). This observation, together with a previous small-angle X-ray scattering (SAXS) analysis of the elongated p115 RhoGEF [8], suggests that the structural relationships among the EhRGS-RhoGEF domains differ from those of mammalian homologs.

The RGS domain of EhRGS-RhoGEF closely resembles the nine-helix bundle found in canonical RGS domains, such as RGS4 (Fig. 3.8B). This canonical RGS domain fold is distinct from the 12-helix rgRGS domains of mammalian RGS-RhoGEFs, such as p115 (Fig. 3.8C) [3]. EhRGS-RhoGEF is unique in possessing a canonical nine-helix RGS domain,

suggesting that the RGS and DH-PH domain combination within *E. histolytica* may have arisen through an independent evolutionary mechanism.

In addition to possessing a distinctive RGS domain fold, the mammalian RGS-RhoGEFs engage $G\alpha_{12/13}$ subunits through an effector-like interface, primarily utilizing switch 2 and the $\alpha3$ helix on $G\alpha$, although the N-terminal extension of the rgRGS domain required for GAP activity also contacts the three switches and the all-helical domain (Fig. 3.12B) [3]. In contrast, canonical nine-helix RGS domains primarily interface with $G\alpha$ switches 1 and 2 (Fig. 3.12A) [38]; hence, the $G\alpha$ subunit switch 1 Gly-to-Ser "RGS-insensitivity" mutation selectively disrupts canonical RGS domain interactions but not $G\alpha$/rgRGS domain interactions [42, 43]. To test whether the EhRGS-RhoGEF RGS domain interfaces with EhGα1 in a canonical fashion, we generated the EhGα1(G168S) mutant. EhRGS-RhoGEF exhibited drastically lower affinity for EhGα1(G168S) than wildtype EhGα1, as measured by SPR, and was unable to affect the intrinsic GTPase rate of EhGα1(G168S) (Fig. 3.3B, G). These experiments suggest that the EhGα1/EhRGS-RhoGEF interface most likely resembles a canonical RGS/$G\alpha$ interaction, providing further evidence for an independent evolutionary mechanism giving rise to a $G\alpha$/RGS-RhoGEF signaling axis in *E. histolytica*.

3.5   DISCUSSION

The *E. histolytica* $G\alpha$ subunit is divergent in its amino acid sequence as compared to mammalian $G\alpha$ subunits and, in particular, does not belong to the $G\alpha_{12/13}$ subfamily that couples to mammalian RGS-RhoGEFs (Fig. 3.13). However, EhGα1 does engage the RGS domain of EhRGS-RhoGEF in a nucleotide state-dependent fashion, resulting in accelerated

GTP hydrolysis. A search of publicly available sequenced genomes identified the RGS and DH-PH tandem domain combination exclusively in metazoan species (*e.g. C. elegans* and *D. melanogaster*) with the only non-metazoan exception being the *Entamoeba* species. Resistance of EhGα1 to conventional Gα subfamily classification, the RGS4-like nine-helix RGS domain fold of the EhRGS-RhoGEF N-terminus, and the canonical nature of the EhGα1/EhRGS-RhoGEF interface, as evidenced by the EhGα1(G168S), EhRGS-RhoGEF(N83A), and EhRGS-RhoGEF(E39K) mutants, all suggest an evolutionary origin independent of the Gα$_{12/13}$/RGS-RhoGEF signaling axis present in mammals.

Mammalian RGS-RhoGEFs are thought to achieve full activation through integration of multiple signals including, but not limited to, interactions with Gα$_{12/13}$. For instance, Gα$_{12}$-mediated stimulation of leukemia-associated RhoGEF requires tyrosine phosphorylation by Tec [44]. Consistent with this theme, EhRGS-RhoGEF requires co-expression, not only with constitutively active EhGα1, but also with constitutively active EhRacC, to achieve apparent activation as evidenced by S2 cell spreading. Little is currently known about EhRacC signaling in *E. histolytica*, although it is evidently a substrate for EhGEF2 *in vitro* [45]. EhRacC was seen to bind EhRGS-RhoGEF directly, exclusively in the GTP-bound conformation, suggesting that EhRGS-RhoGEF may serve as an EhRacC effector. Activated human RhoA GTPase has been demonstrated to bind the PH domain of PDZ-RhoGEF in an analogous fashion [36]. RhoA also serves as a substrate for PDZ-RhoGEF-mediated exchange, suggesting a possible mode of feedback regulation in mammals. However, there is currently no evidence that EhRacC is a substrate for EhRGS-RhoGEF; in fact, EDTA-treated, nucleotide-free EhRacC did not bind appreciably to EhRGS-RhoGEF (Fig. 3.4D). However, the full-length isolated EhRGS-RhoGEF used in these experiments is expected to

have an obstructed Rho substrate-binding site, and activation by EhGα1·GTP, EhRacC·GTP, and/or other factors may be required to allow efficient substrate binding. Although the putative exchange factor activity of EhRGS-RhoGEF was not directly measureable, selective knockdown of endogenous *D.m.* Rac1/2 in S2 cells impaired the cell spreading triggered by co-expression of EhRGS-RhoGEF, EhGα1(Q/L), and EhRacC(Q/L), suggesting that *Drosophila* Rac may serve as a substrate for EhRGS-RhoGEF in a cellular context. However, the cell spreading experiments provide limited insight into the precise signaling mechanics. For instance, additional cellular factors may contribute to EhRGS-RhoGEF activation, and we cannot rule out the possibility that overexpressed *E. histolytica* signaling components promote cell spreading through other endogenous signaling pathways.

In isolation, EhRGS-RhoGEF appears to adopt an autoinhibited conformation, with direct obstruction of the presumptive Rho substrate-binding surface by a putative inhibitory helix and its DH and PH domain interactions. We hypothesize that binding of EhGα1·GTP and EhRacC·GTP to EhRGS-RhoGEF, possibly together with other cellular context factors such as membrane localization or post-translational modifications, may lead to a structural rearrangement of the putative inhibitory helix and the PH domain, allowing for substrate Rho GTPase binding and nucleotide exchange. The predicted mode of EhRGS-RhoGEF autoinhibition, as derived from the crystal structure, is comparable to that of mammalian PDZ-RhoGEF, seen in solution studies [9]. However, in the case of PDZ-RhoGEF, an acidic stretch of its rgRGS-DH domain linker interacts with a DH domain surface basic cluster, distinct from the inhibitory helix interface seen here in the crystal structure of EhRGS-RhoGEF. The rgRGS-DH linker of p115 RhoGEF also apparently inhibits the GEF activity of its DH-PH domain tandem, although both SAXS analyses and crystallographic studies of

the DH-PH domains with short segments of the linker intact suggest a different linker disposition than that seen in EhRGS-RhoGEF [7]. The RGS-DH linker in EhRGS-RhoGEF is >100 residues shorter than those of mammalian homologs (Fig. 3.1C), further suggesting that this region does not have a conserved structure across species.

Endogenous EhRGS-RhoGEF likely represents a functional signaling link between heterotrimeric G-proteins and Rho family GTPases in *E. histolytica*. Indeed, Rho GTPases and other Dbl family RhoGEFs in *E. histolytica* are also known to regulate multiple processes important for pathogenesis such as actin reorganization during chemotaxis, surface receptor capping, cell killing, phagocytosis, and tissue destruction [13]. A surprisingly large family of Rho GTPases (>20 members) is apparently simultaneously expressed in the single-celled parasite [14]. Further studies are needed to determine which Rho family members are activated by EhRGS-RhoGEF and what downstream signaling pathways are utilized.

Overexpression of EhRGS-RhoGEF resulted in reduced *E. histolytica* trophozoite chemotactic migration, attachment to and killing of host cells, and secretion of cysteine proteases. By each of these measures, the EhRGS-RhoGEF trophozoite strain phenocopies a strain overexpressing a dominant negative EhGα1 point mutant and exhibits an opposing trend to trophozoites overexpressing wild type EhGα1 [16], consistent with ectopically overexpressed EhRGS-RhoGEF serving to accelerate GTP hydrolysis on EhGα1 and thus inhibit its signaling. Given the amenability of heterotrimeric G-protein signaling to pharmacological manipulation [1], this pathway provides a promising drug target for the treatment of amoebic colitis.

A

| | | | | | | α1 | | |
| EhRGS-RhoGEF | 1 | MAATE | - | NKGMRTRLGCLSHKSDSCSD | 5 |

A (sequence alignment)

EhRGS-RhoGEF   1 MAATE -   5
HsRGS8   1 MAALLMP - - RR - NKGMRTRLGCLSHKSDSCSD   29
HsRGS4   1 MCKGLAGLPASCLRSAKDMKHRLGFLLQKSDSCEHNSSHNKKDKVVICQ - - - - - - - - - - - RVS   52
HsRGS2   1 MQSAMFL - - -AVQHD - CRPMDKSAGSGHKSEEKREKM   33
p115 RGS-RhoGEF 1 MEDFA -   5

                                                                                    α1   20
EhRGS-RhoGEF   6 - TAILEGWP - TLQEVLEDS   22
HsRGS8   30 FTAILPDKPNRA - - - - - - - - - - - - - LKRLSTEEA - - TR - - - - - - - - - - WADSFDVLLSHK   64
HsRGS4   53 QEEVKKWAES - LENLISHE   70
HsRGS2   34 KRTLLKDWKTRLSYFLQNSSTPGKPKTGKKSKQQAFIKPSPEEAQ - - - - - - - - - LWSEAFDELLASK   91
p115 RGS-RhoGEF 6 - RGAASPGP - - - - - - - - - - - - - - SRPGLVPVSIIGAEDEDFENELTNSEEQNSQFQSLEQVKRRP   56

                        α2   α3   α4
EhRGS-RhoGEF   23 FMKRLLRCYLSDERSEENLDFLESVGLYESQFDK - - LTPKVRLEALNFIKDQFLDRNSERQVNLSYQ   87
HsRGS8   65 YGVAAFRAFLKTEFSEENLEFWLACEE - - - - FKKTRS - TAKLVSKAHRIFEEFVDVQAPREVNIDFQ 126
HsRGS4   71 CGLAAFKAFLKSEYSEENLEFWISCEE - - - - YKKIKS - PSKLSPKAKKIYNEFISVQATKEVNLDSC 132
HsRGS2   92 YGLAAFRAFLKSEFCEENIEFWLACED - - - - FKKTKS - PQKLSSKARKIYTDFIEKEAPKEINIDFQ 153
p115 RGS-RhoGEF 57 AHLMALLQHVALQFEPGPLLCCLHA - - - DMLGS - - LGPKEAKKAFLDFYHSFLEKTAVLRVPVPPN 117

                           α5
EhRGS-RhoGEF   88 IQQSILKKLSE - - - - -VTSNA - - PKDV   107
HsRGS8   127 TREATRKNLQE - - - - - PSLT - - - - - C   142
HsRGS4   133 TREETSRNMLE - - - - - PTIT - - - - - C   148
HsRGS2   154 TKTLIAQNIQE - - - - -ATSG - - - - - C   169
RGS-RhoGEFp115 118 VAFELDRTRADLISEDVQRRFVQEVVQSQQVAVGRQLEDFRSKRLMGMTPWEQELAQLEAWVGRDRA 184

                           α7   α8   α9
EhRGS-RhoGEF 108 - - - FNEAKKATEYLLYTEQYTYFINK - - - - - - - LNANTIGT - GKKDVYSLYLNQ - - - - FPKTNPQPL 159
HsRGS8   143 - - - FDQAQGKVHSLMEKDSYPRFLRSKMYL - - - - - - - - - - - - - - - - - DL - - LSQSQRRLS 180
HsRGS4   149 - - - FDEAQKKIFNLMEKDSYRRFLKSRFYLDLVNPSSCGA - EKQKGAK - SSAD - - - - CASLVPQCA 205
HsRGS2   170 - - - FTTAQKRVYSLMENNSYPRFLESEFYQ - - - - - - - - - - - - - - - - - - - DLCKKPQITTEPHAT 211
p115 RGS-RhoGEF 185 SYEARERHVAERLMHLEEMQHTIST - - - - - DEEKSAAVVNAIGLYMRHLGV - - - - RTKSGDKKS 240

B

                                           α1   α2a   α2b
EhRGS-RhoGEF 186 - - - - - - - - - - - SIKSLVESLIQDECNYVGVLTSLSE - - - - - FSEIMTKKQILG - - PDVLKELFDHIPVLIQHHQKFISS 246
Intersectin 1229 - - -DMLTPTERKRQGYIHELIVTEENYVNDLQLVTE - - - - IFQKPLMESELLT - - EKEVAMIFVNWKELIMCNIKLLKA 1298
Dbs   622 - GEEEESLAILRRHVMSELLDTERAYVEELLCVLEGYAAEMDNPLMAHLLSTGLHNKKDVLCLFGNMEEIYHFHNRIFLR 698
p115 RGS-RhoGEF 478 DTLHSLPKSQVKRQEVISELLVTEAAHVRMLRVLHD - - - -LFFQPMAECLFFP - - LEELQNIFPSLDELIEVHSLFLDR 550
Tiam1   1032 - -MRQLSDADKLRKVICELLETERTYVKDLNCLMER - - - - YLKPLQKETFLT - - QDELDVLFGNLTEMVEFQVEFLKT 1101

                   α3a   α3b   α3c   α3d   α4
EhRGS-RhoGEF 247 LQEAKADEKVGEKLNSGLHFLVLYRYYLR - - - - - - -HVPKNIAKLCSIGMTDEIEVGRELYPLVIEEFDKQQKMT - - K 316
Intersectin 1299 LRVRKKMSGEKMPVKMIGDILSAQLP - - - - - - - - - - HMQPYIRFCSRQLNGAALIQQKTDEAPDFKEFVKRLAMDPRCK 1367
Dbs   699 ELENYTDCP - - - - -ELVGRCFLERMED - - - - - - - - FQIYEKYCQNKPRSESLWRQCS - DCPFFQECQRKLD - - - - - H 756
p115 RGS-RhoGEF 551 LMKRRQESG - - YLIEEIGDVLLARFDGA - - - - - EGSWFQKISSRFCSRQSFALEQLKAKQRKDPRFCAFVQEAESRPRCR 623
Tiam1   1102 LEDGVRLVPDLEKLEKVDQFKKVLFSLGGSFLYYADRFKLYSAFCASHTKVPKVLVKAK - TDTAFKAFLDAQNPKQQ - H 1178

                                   α5   α6a   α6b
EhRGS-RhoGEF 317 KMSVLQMLVYPYFRVRTYQAYVDDFIKITKKDSQEVKELEVVHSQLAIFQELINTYSDINKIERISDALKLLFPFSFTS 395
Intersectin 1368 GMPLSSFILKPMQRVTRYPLLIKNILENTPENHMQPDHSHLKHALEKAEELCSQVNEGVREKENSDRLEWIQAHVQCEGLS 1446
Dbs   757 KLSLDSYLLKPVQRITKYQLLLKEMLKYS - RNCEGAEDLQEALSSILGILKAVNDSMH - - - -LIAITGYDG - -NLGDLG 828
p115 RGS-RhoGEF 624 RLQLKDMIPTEMQRLTKYPLLLQSIGQNT - EEPTEREKVELAAECCREILHHVNQAVRDMEDLLRLKDYQRRLDLSHLR 701
Tiam1   1179 SSTLESYLIKPIQRILKYPLLLRELFALTDAESEEHYHLDVAIKTMNKVASHINEMQKIHEEFGAVFDQLIAEQTGEKK 1257

                   α7   β1   β2   β3
EhRGS-RhoGEF 396 IMPLFEGKNGICGIASLDRFDKTDINQLSMSLNSRKKLTLIILYRGVVVTDVPVIRKKGNVSNSIDKSFYS 466
Intersectin 1447 - - -EQLVFNSVTNCLGPRKFLHSGKLYKAKSNK - - - -ELYGFLFNDFLLLTQITKP - LGSSGTDKVFSPKSN - - - - - -LQY 1513
Dbs   829 - - - -KLLMQGSFSVWTDHKRGHTKVKELARFKP - -MQRHLFLHEKAVLFCKKREE - - - -NGEGYEKAPSYS - - - - -YKQ 892
p115 RGS-RhoGEF 702 QSSDPMLSEFKNLDITKKLVHEGPLTWRVTDKDKAVEVHVLLLDDLLLLLQRQDERLLLKSHSRTLTPTPD - - - -GKT 775
Tiam1   1258 - -EVADLSMGDLLLHHTTVIWLNPPASLGKWKKE - PELAAFVFKTAVVLVYKDGSKQKKKLVGSHRLSIYEDWDPFRFR 1332

                   β4   β5   β6   β7   αC
EhRGS-RhoGEF 467 - - - - - -FTLIGDIRDFGTEDSTETIYIDVPEI - - - - -KKRIWFGCESTEEFKSCVEALRTILSN 519
Intersectin 1514 KMYKTPIFLNEVLVKLPTDPSGDEPIFHISHI - - - -DRVYTLRAESINERTAWVQKIKAASELYIETEKKKRE - - - - 1582
Dbs   893 SLN - - -MAAVGITENVKGDAKKFEIWYNAR - - - - -EEVYIVQAPTPEIKAAWVNEIRKVLTSQLQACREASQHRAL 960
p115 RGS-RhoGEF 776 MLRPVLRLTSAMTREVATDHKAFYVLFTWDQE - - - -AQIYELVAQTVSERKNWCALITETAGS - - - - - - - - - - - - - 834
Tiam1   1333 HMIPTEALQVRALASADAEANAVCEIVHVKSESEGRPERVFHLCCSSPESRKDFLKAVHSILRDKHRRQL - - - - - - - - 1402

C

	RGS-DH linker length (a.a.)
EhRGS-RhoGEF	26
p115 RhoGEF	164
PDZ RhoGEF	243
LARG	229

129

Figure 3.1. Multiple sequence alignment of the EhRGS-RhoGEF RGS and DH-PH domains. Sequence alignments were conducted with T-Coffee (www.tcoffee.org) for the RGS domain **(A)** and the DH-PH domains **(B)** of EhRGS-RhoGEF. The RGS domain most closely resembles canonical RGS domains rather than the RGS-like domain of p115 RhoGEF. Glu-39 and Asn-83, residues highly conserved among canonical RGS proteins, were mutated in EhRGS-RhoGEF and eliminated both binding to EhGα1 and acceleration of GTP hydrolysis (Fig. 3.3). A DALI search indicated that the EhRGS-RhoGEF DH domain most closely resembles Intersectin. **(C)** The linker joining the RGS and DH domains of EhRGS-RhoGEF is substantially shorter than that of mammalian homologs.
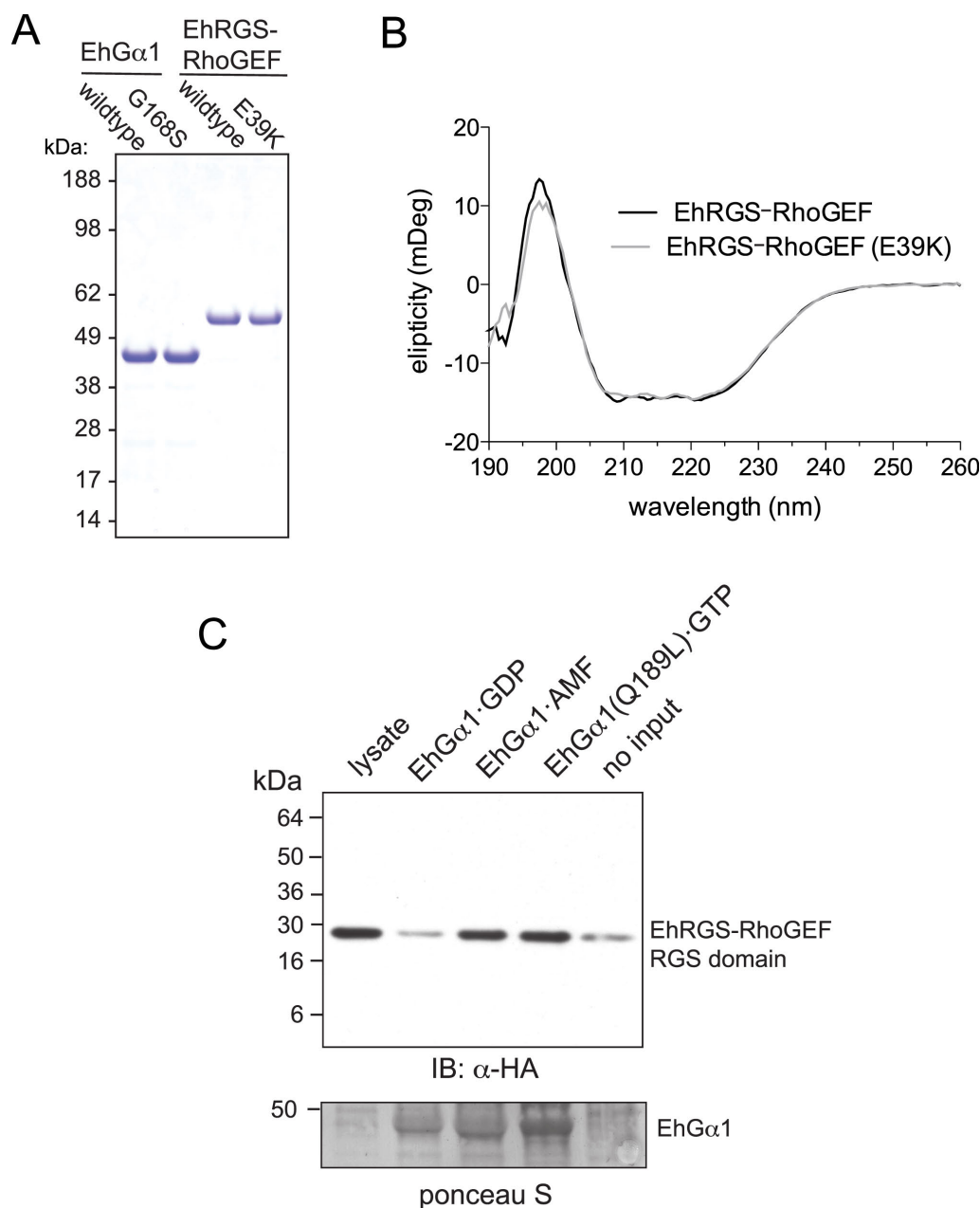
Figure 3.2. Circular dichroism (CD) of EhRGS-RhoGEF(E39K) and nucleotide-dependent interaction of EhGα1 with the isolated RGS domain. (A) EhGα1, EhRGS-RhoGEF, and point mutants used in biochemical experiments were purified from E. coli by affinity chromatography and gel filtration. (B) Mutation of a predicted surface residue on EhRGS-RhoGEF (Glu-39 to Lys) does not disrupt global protein folding, as measured by CD. (C) HA epitope-tagged RGS domain (a.a. 1-152) was expressed in HEK-293T cells and co-precipitated with His₆-tagged EhGα1 on NTA resin. The EhRGS-RhoGEF RGS domain selectively co-precipitated with EhGα1·AMF and EhGα1(Q189L)·GTP, compared to EhGα1·GDP or NTA resin alone. Ponceau S staining indicated approximately equal loading of EhGα1 in each experiment.

131
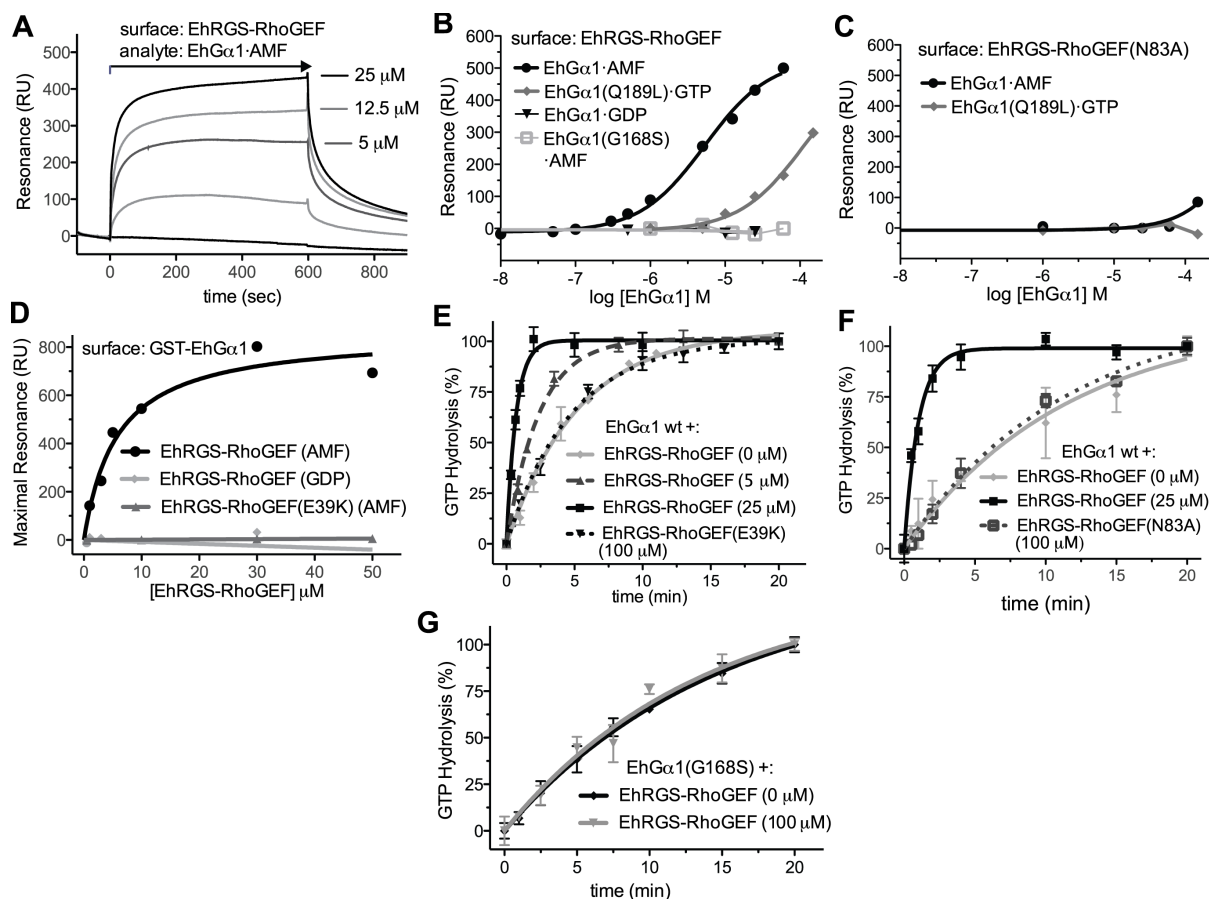
Figure 3.3. EhRGS-RhoGEF is an EhGα1 effector that accelerates its GTP hydrolysis.
(A-D) Either EhGα1 or EhRGS-RhoGEF was immobilized on a surface plasmon resonance chip, and the complementary protein flowed over at increasing concentrations. EhRGS-RhoGEF bound EhGα1 selectively in the GDP·AlF$_4^-$-bound state (AMF); interaction affinity was independent of immobilized species (K_D = 5.2 ± 0.8 μM, panel B; 5.7 ± 1.6 μM, panel D). The GTPase-deficient mutant mutant EhGα1(Q189L) also interacted with EhRGS-RhoGEF, but with lower affinity (K_D ≈ 110 μM). Mutation of the conserved EhRGS-RhoGEF Asn-83, predicted to orient EhGα1 residues for rapid GTP hydrolysis, the predicted EhGα1-binding surface charge reversal EhRGS-RhoGEF(E39K), and the switch 1 mutant EhGα1(G168S) all drastically reduced binding affinity. Sensorgrams and equilibrium binding curves are representative of three experiments. **(E, F)** The GTPase rate of EhGα1 was accelerated by EhRGS-RhoGEF in a dose-dependent fashion (k_{obs} = 0.20 ± 0.02 min^{-1} for EhGα1 alone and 1.45 ± 0.13 min^{-1} upon addition of 25 μM EhRGS-RhoGEF). EhRGS-RhoGEF(E39K) and EhRGS-RhoGEF(N83A) had no effect on the hydrolysis rate. **(G)** GTP hydrolysis rates for the "RGS-insensitivity" mutant EhGα1(G168S) alone or in the presence of a high concentration of EhRGS-RhoGEF were indistinguishable. Error bars represent standard error of the mean for duplicate reactions. Each single turnover hydrolysis experiment was independently repeated at least twice.
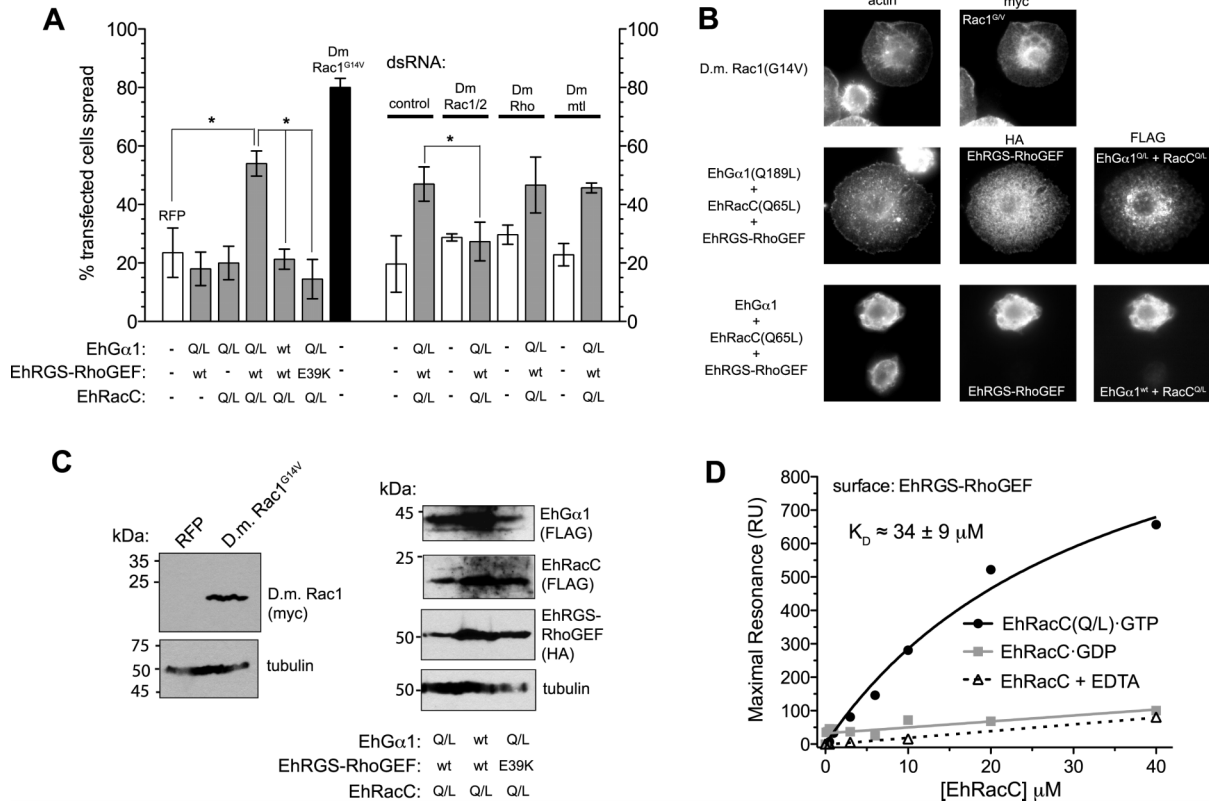
Figure 3.4. EhRGS-RhoGEF activation by constitutively active EhGα1 and EhRacC leads to Rac-dependent S2 cell spreading. (A-C) Rho family GTPase activation in *D. melanogaster* S2 cells leads to spreading on a poly-lysine coated surface [33]. Coexpression of EhRGS-RhoGEF with GTPase-deficient and constitutively active EhGα1(Q189L) was insufficient to effect cell spreading. However expression of the constitutively active EhRacC(Q65L), together with EhGα1(Q189L) and EhRGS-RhoGEF significantly enhanced cell spreading, while EhRacC(Q65L) alone or in combination with EhGα1 had no effect. To determine which *D. melanogaster* Rho family GTPases were necessary for cell spreading, and thus potential substrates for overexpressed EhRGS-RhoGEF, double-stranded RNAi was employed as described previously [35]. RNAi-mediated knockdown of *D.m.* Rac isoforms, but not Rho or mtl, prevented significant enhancement of cell spreading by coexpression of EhRGS-RhoGEF, EhRacC(Q65L), and EhGα1(Q189L). Error bars represent standard deviation for three independent experiments, and * indicates statistically significant difference ($p < 0.05$) by Student's t-test. Example micrographs are shown in panel B and western blots confirming expression of all *E. histolytica* proteins and mutants are shown in panel C. **(D)** Recombinant, activated EhRacC(Q65L)·GTP was seen to directly bind EhRGS-RhoGEF by surface plasmon resonance, while EhRacC·GDP and nucleotide-free EhRacC exhibited negligible binding up to 40 μM concentration.

133
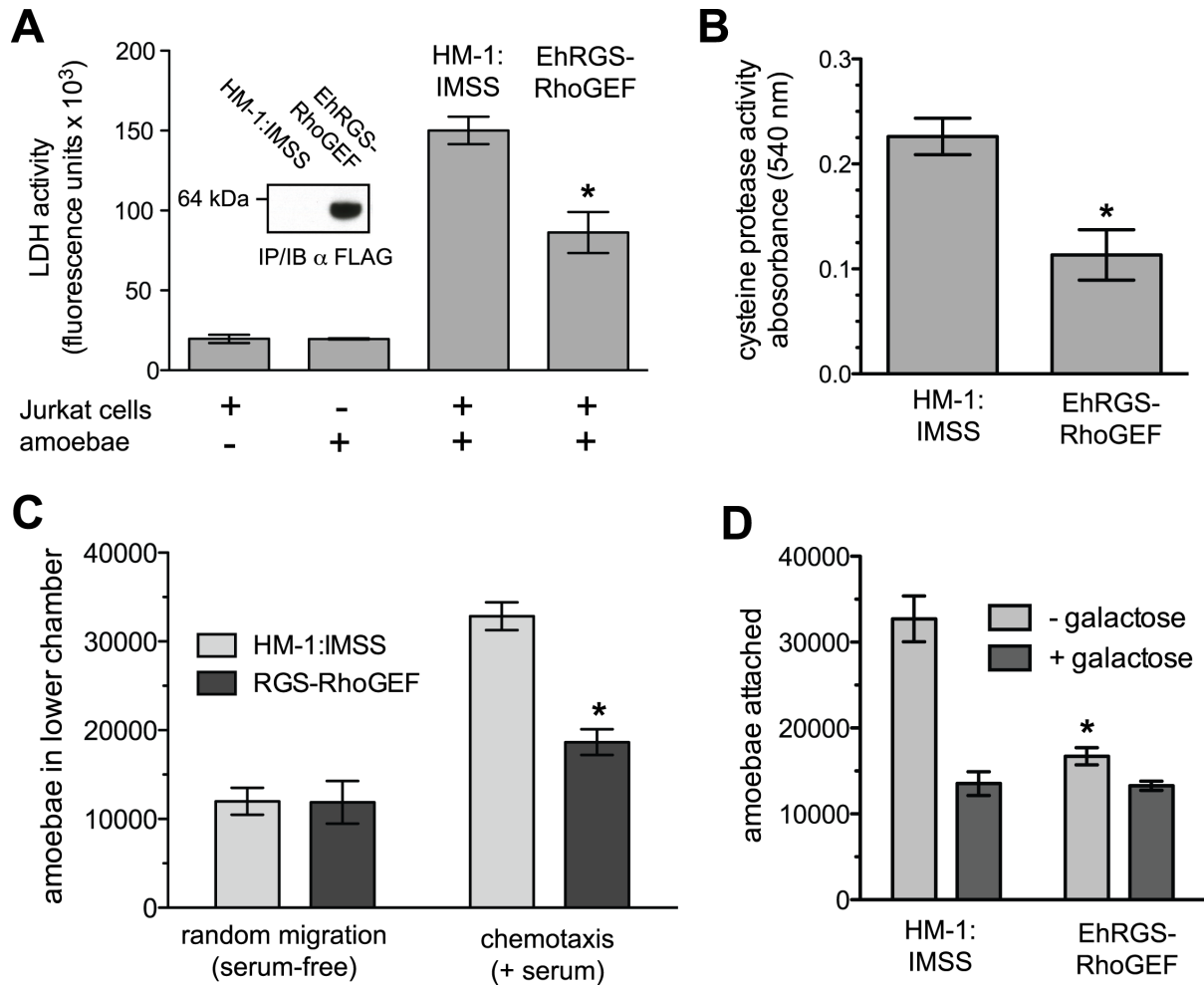
Figure 3.5. EhRGS-RhoGEF expression inhibits host cell attachment and killing, cysteine protease secretion, and chemotactic migration by *E. histolytica* trophozoites. (A) Amoebae were stably transfected to overexpress EhRGS-RhoGEF (*inset)*. Expression of EhRGS-RhoGEF reduced the ability of *E. histolytica* to kill Jurkat (human lymphocyte-derived) cells compared to the HM-1:IMSS virulent parent strain, as determined by a membrane integrity assay. **(B)** Trophozoites expressing EhRGS-RhoGEF exhibited reduced cysteine protease secretion, a process necessary for host cell killing and extracellular matrix invasion. **(C)** Overexpression of RGS-RhoGEF reduced trophzoite chemotactic migration across a porous membrane toward serum-containing nutritive media, but had no measureable effect on random migration. **(D)** Trophozoites attach to CHO cell monolayers, primarily through a galactose-inhibitable lectin. Overexpression of EhRGS-RhoGEF reduced lectin-dependent monolayer attachment. All trophozoite experiments were conducted in quadruplicate. * indicates a statistically significant difference by Student's t-test ($p < 0.05$).

A

RGS

inhibitory helix

PH

DH

N

C

B

ITSN/Cdc42

C

N

α6b

N

α6a

C
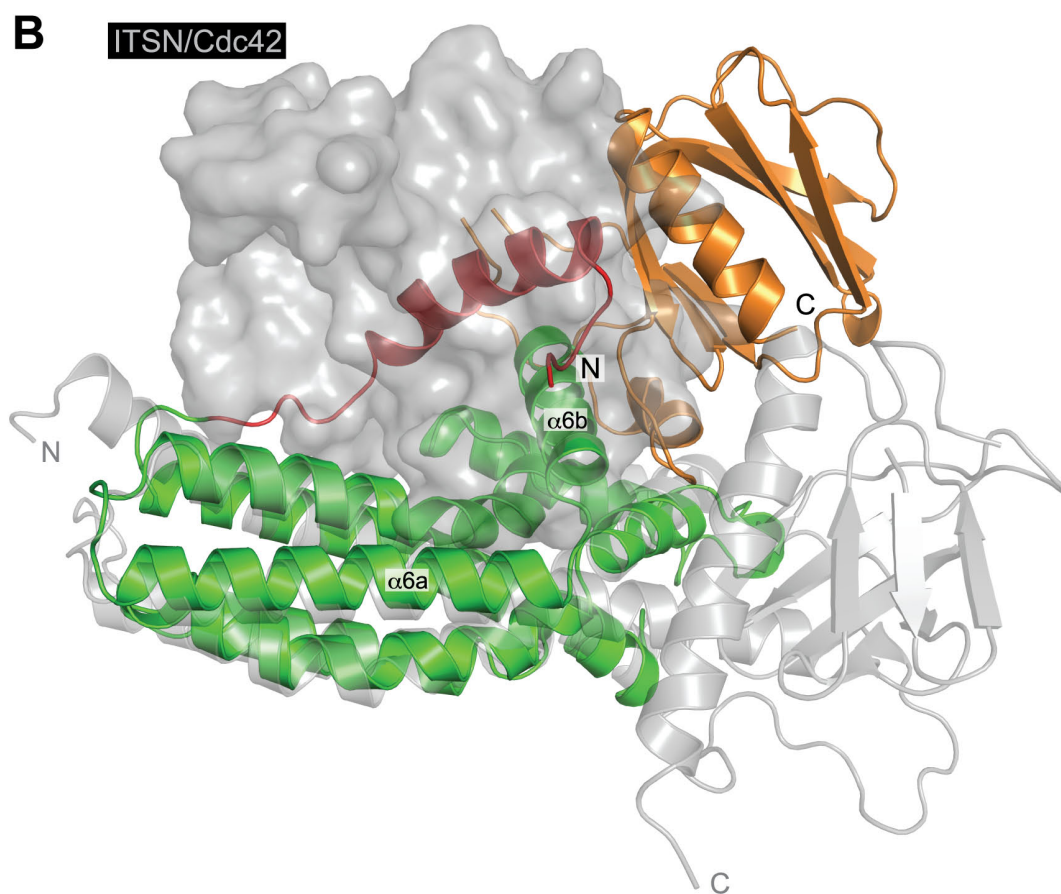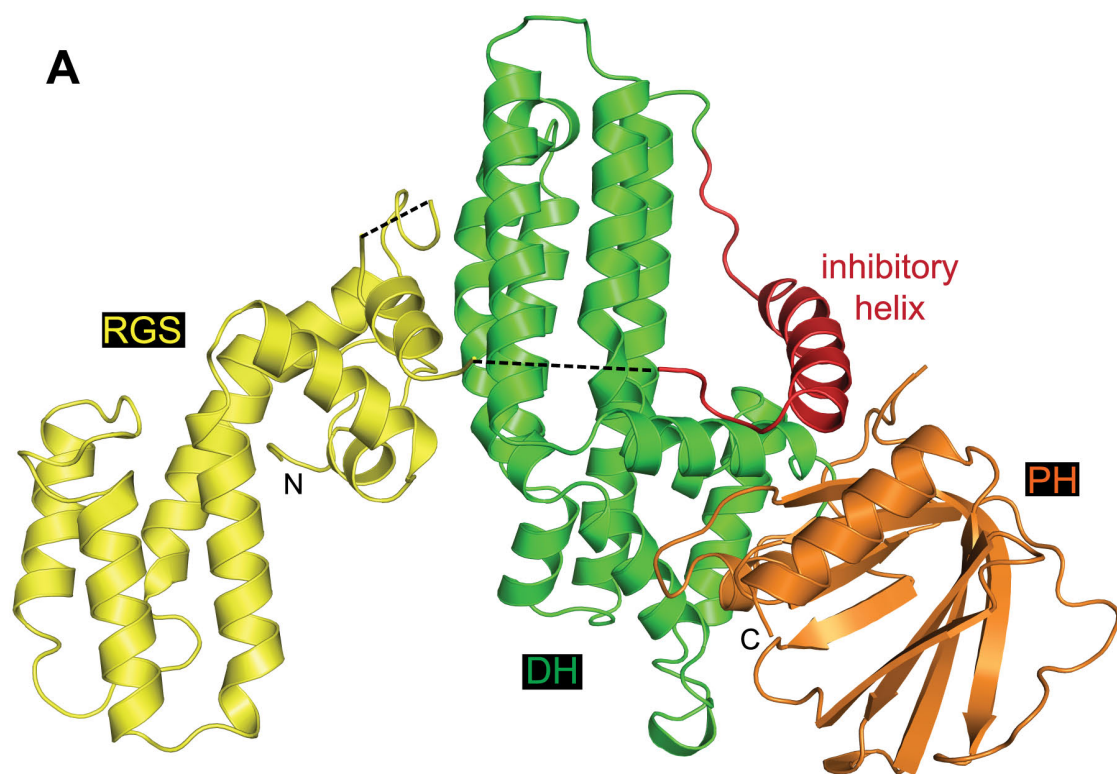
Figure 3.6. The structure of EhRGS-RhoGEF reveals inter-relationship between RGS and DH/PH domains. (A) The RGS domain (*yellow*) adopts a canonical 9-helix fold and interacts with the DH domain (*green*) opposite from the predicted Rho binding site. The linker between the RGS and DH domains wraps ~180° around the DH domain and contains a 15-residue α helix (termed the 'inhibitory helix'; *red*) that engages both the conserved PH domain fold (*orange*) and the C-terminal portion of the DH domain. **(B)** The inhibitory helix, DH, and PH domains are superimposed with the structure of Intersectin/Cdc42 (*gray*; PDB id 1KI1). The conserved site of Rho GTPase interaction, illustrated by a surface rendering of Cdc42, is obstructed in the case of EhRGS-RhoGEF. The inhibitory helix lies entirely within the space corresponding to Cdc42. In addition, the long α6 helix is continuous in Intersectin and other RhoGEFs, but is segmented into two helices related by an ~90° angle in EhRGS-RhoGEF. The α6b helix both interacts with the putative inhibitory helix and contributes to obstruction of the Rho binding site. The EhRGS-RhoGEF PH domain is also predicted to interfere with Rho binding in this inactive conformation.
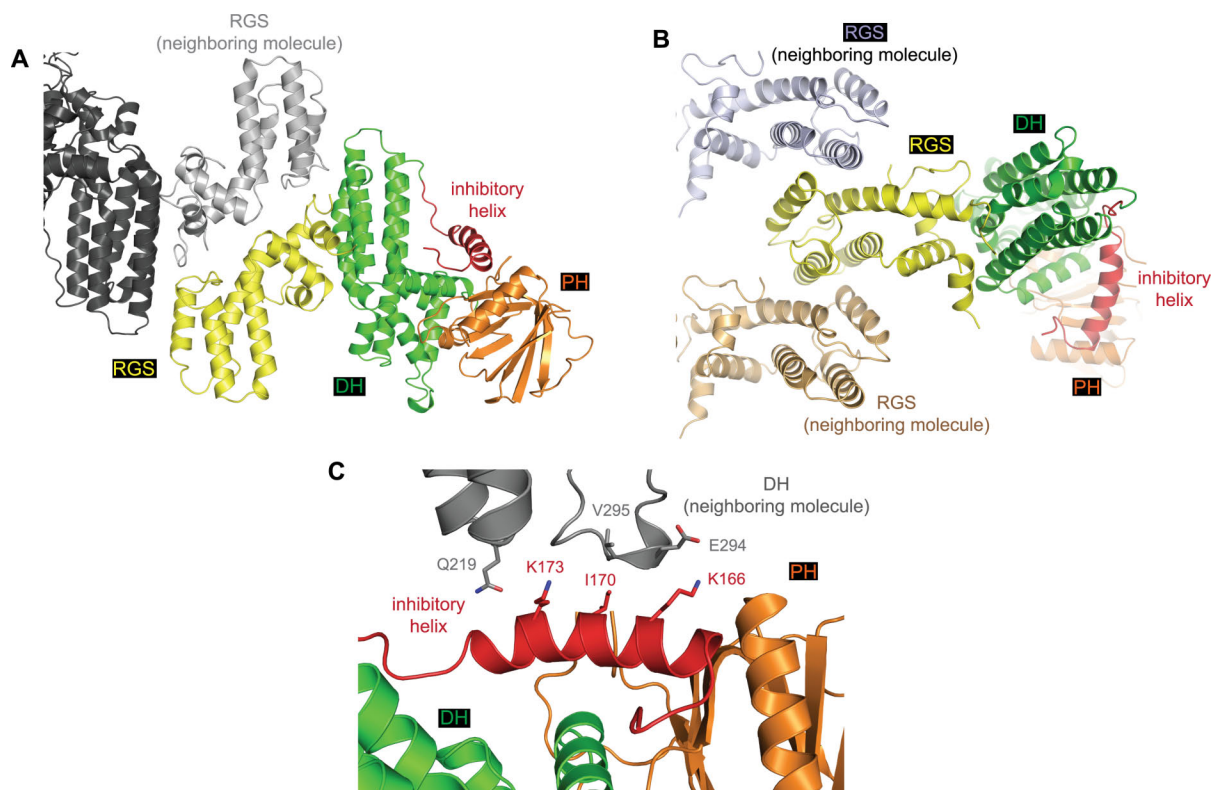
Figure 3.7. Crystal contacts of the EhRGS-RhoGEF RGS domain and inhibitory helix. The RGS domain (*yellow*) of EhRGS-RhoGEF contacts symmetry-related molecules in the crystal lattice. **(A)** The putative EhGα1-binding surface of the RGS domain contacts the corresponding face of a neighboring RGS domain (*gray*). **(B)** Two other RGS domain surfaces interface with additional neighboring RGS domains (*blue* and *brown*). Although crystal contacts may modestly affect the disposition of the RGS domain, the similarity of its conformation to mammalian RGS domains (Fig. 3.8B) and significant burial of hydrophobic surface area (~850 Å^2) at the RGS-DH domain interface (Fig. 3.8A) suggest that the crystal structure architecture accurately reflects that of EhRGS-RhoGEF in solution. **(C)** The inhibitory helix (*red*) forms limited contacts with the neighboring DH domain (*gray*) in the crystal lattice. The side chains of Ile-170 and Lys-166 make van der Waals contacts with the main chain of the neighboring DH domain. Although no side chains depicted in the model are within hydrogen bonding distance, two side chain pairs could potentially form polar contacts, namely Lys-173/Gln-219 and Lys-166/Glu-294.
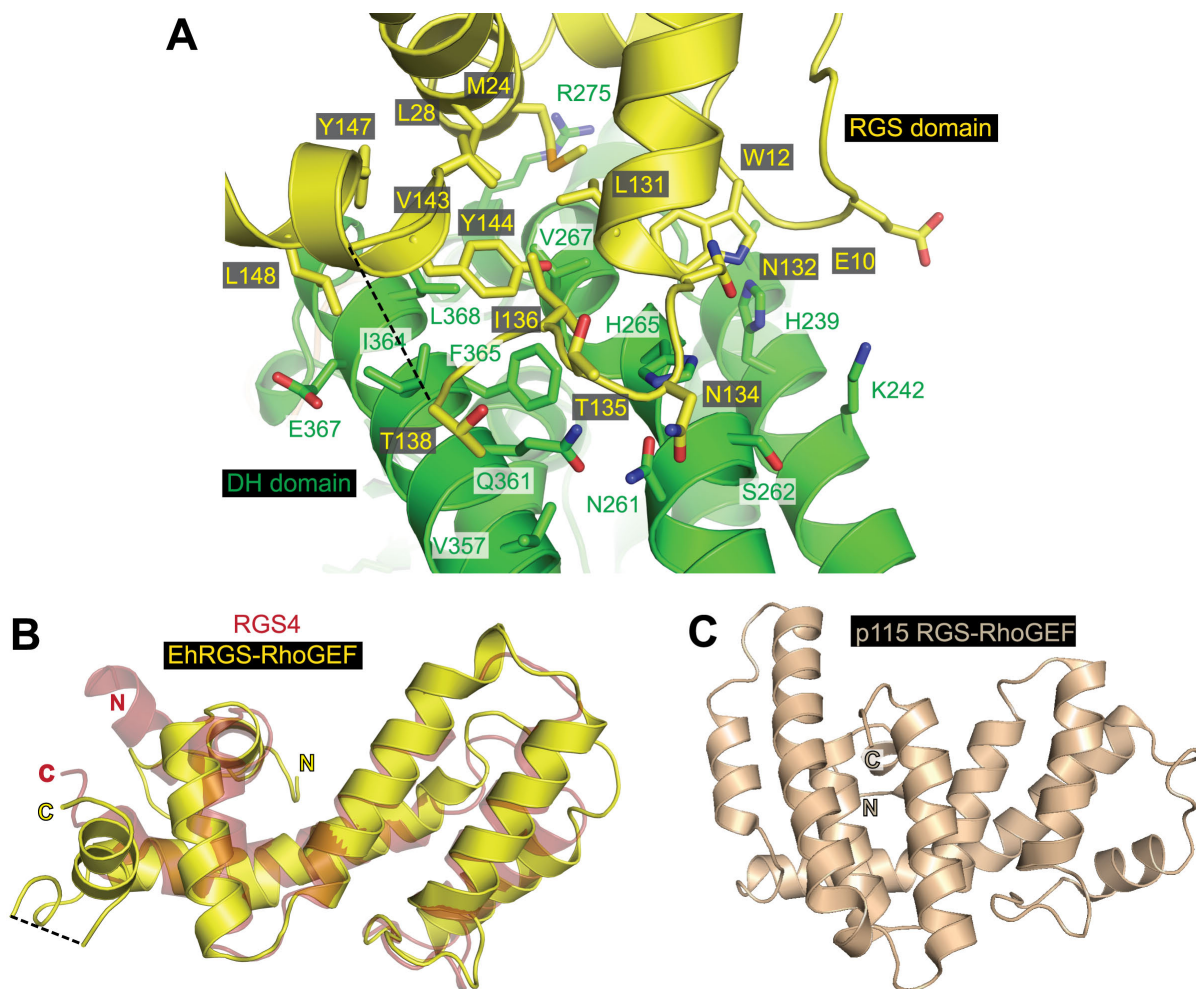
Figure 3.8. The EhRGS-RhoGEF RGS domain adopts a canonical fold and interacts with the DH domain. (A) Residues participating in the RGS (*yellow*) and DH (*green*) domain interface in EhRGS-RhoGEF are shown in sticks. A central hydrophobic region is surrounded by polar and ionic side chain interactions. **(B)** The EhRGS-RhoGEF RGS domain adopts a 9-helical bundle fold very similar to canonical RGS domains, typified by RGS4 (red; PDB id 1AGR). **(C)** In contrast, the mammalian RGS-RhoGEFs possess an RGS-like domain with 12 helices, as seen in p115 RGS-RhoGEF (PDB id 1IAP). Dotted lines indicate loops that could not be accurately modeled.
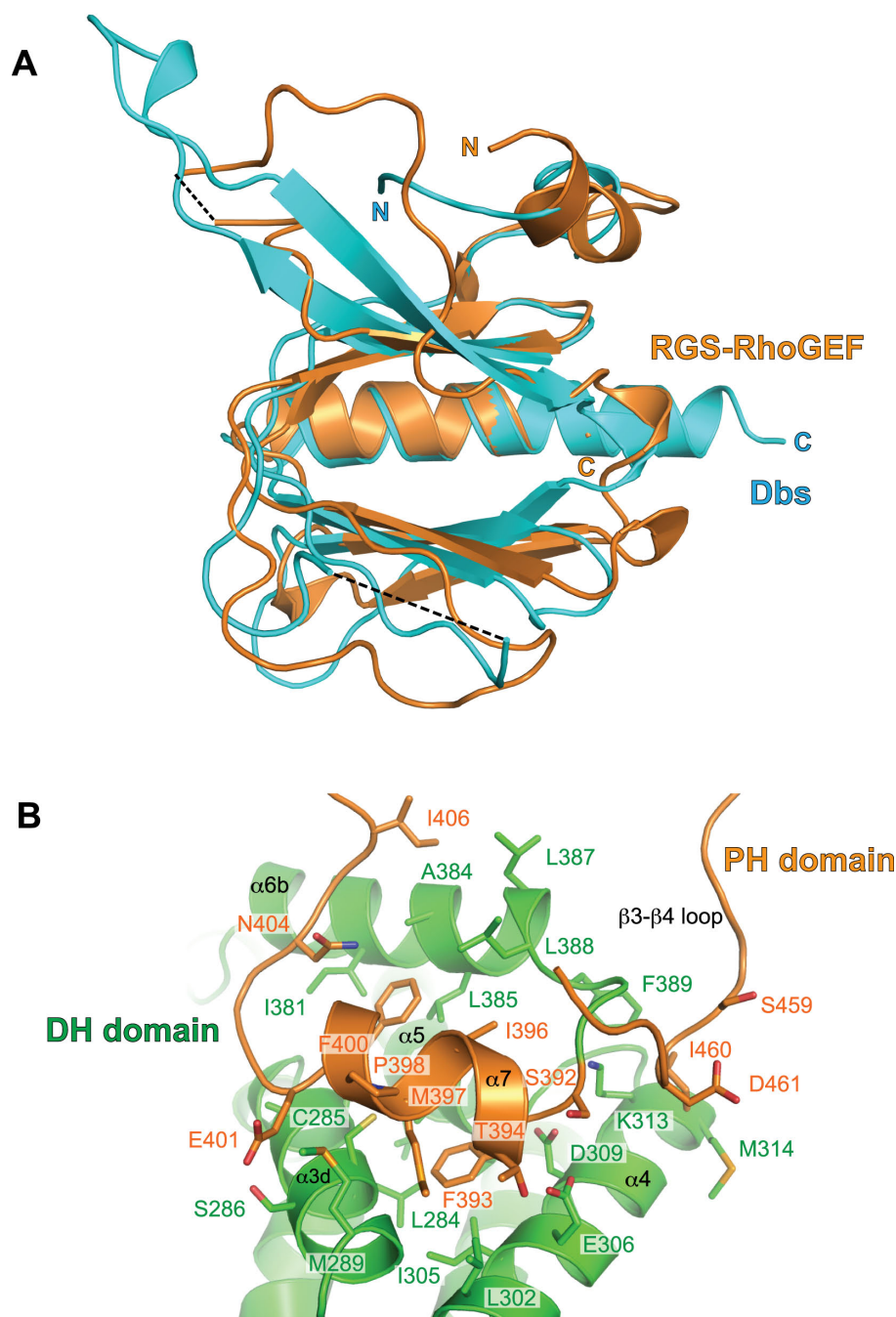
Figure 3.9. The EhRGS-RhoGEF PH domain contacts the DH domain. (A) Despite only 51% sequence similarity to the Dbs RhoGEF (*cyan*; PDB id 1RJ2), the EhRGS-RhoGEF PH domain (*orange*) is structurally similar with a 2.9 Å Cα r.m.s.d. The β-sandwich fold is conserved with variation in the lengths and positioning of individual β-strands. EhRGS-RhoGEF also lacks the C-terminal helical extension seen in Dbs. **(B)** The α7 helix of EhRGS-RhoGEF, together with the β4-β4 loop of the PH domain (*orange*) contact the DH domain (*green*) and buy ~1200 Å^2 of surface area. The side chains of Phe-393 and Met-397 project into an approximately triangular concavity formed by helices α3d, α4, and α5 on the DH domain.
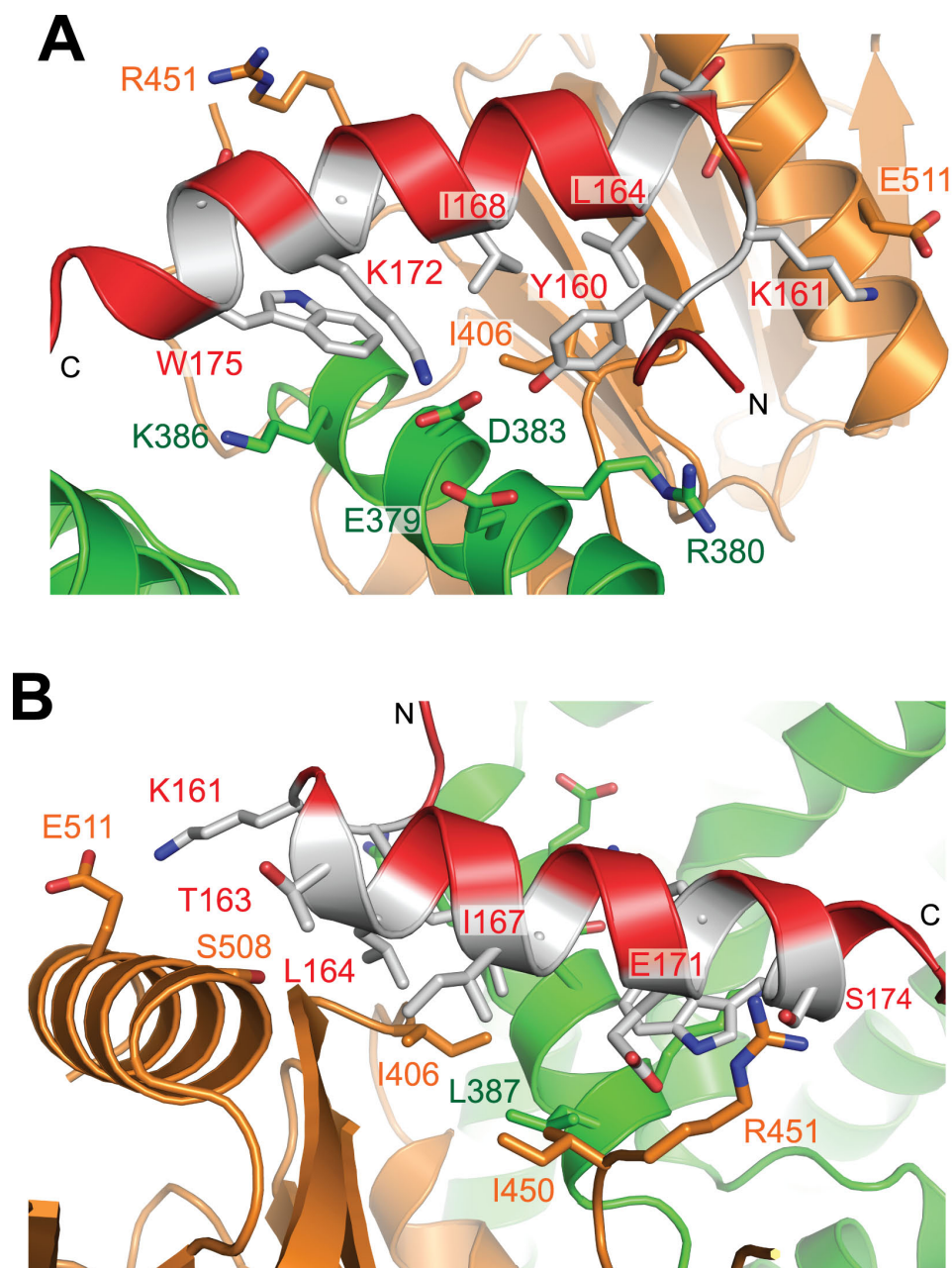
Figure 3.10. The EhRGS-RhoGEF inhibitory helix engages both the DH and PH domains. (A,B) The EhRGS-RhoGEF DH and PH domains share an interface with the hydrophobic inhibitory helix residues Leu-164, Ile-167, Ile-168, Trp-175, and the aromatic ring of Tyr-160 (*gray sticks*). Hydrogen-bond interactions and peripheral ionic interactions, *e.g.* Lys-172·Asp-383 and Lys-161·Glu-511, also contribute to this interface. N and C indicate the N- and C-terminal ends of the inhibitory helix, respectively.
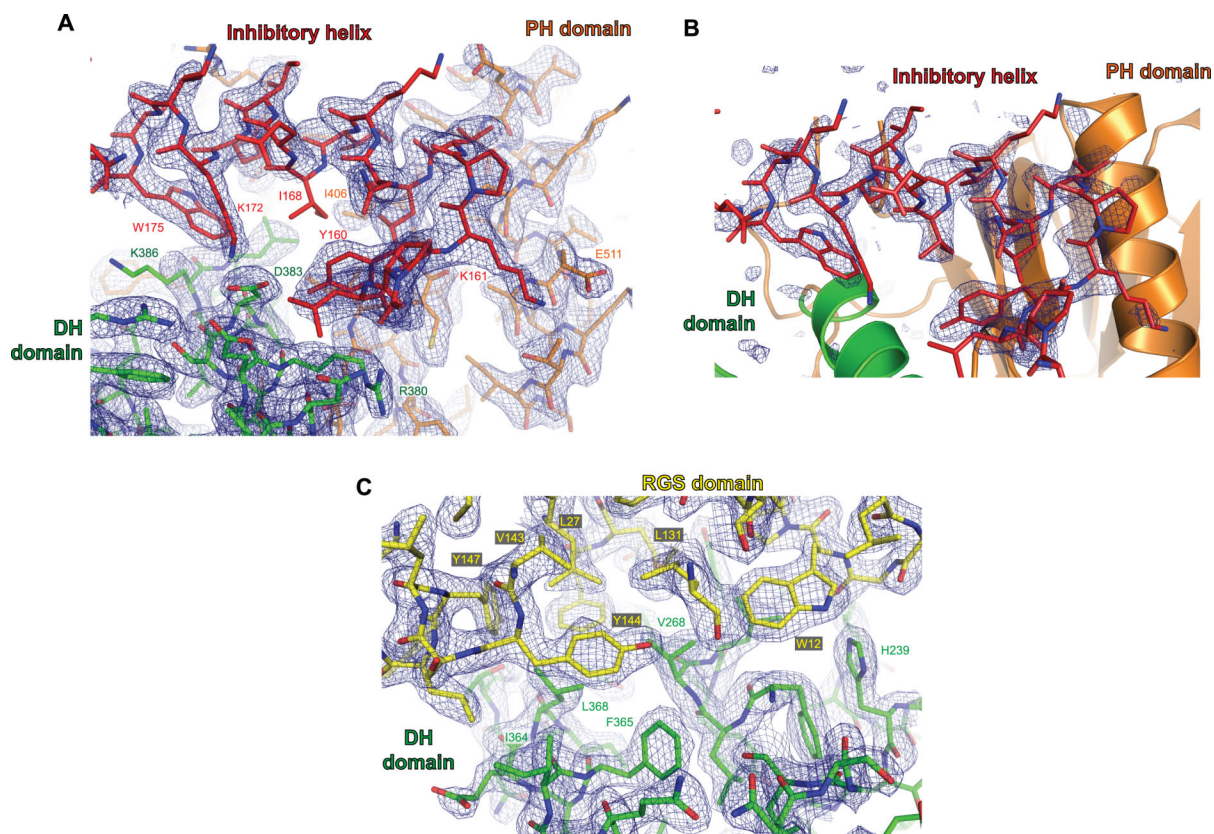
Figure 3.11. Electron density maps of the inhibitory helix and RGS/DH domain interface regions. (A) A $2F_o$-F_c map was rendered using Pymol and contoured to σ = 2.5. The inhibitory helix of the RGS-DH linker (*red*) contacts both the DH (*green*) and PH (*orange*) domains, as shown in Figure 3.10. **(B)** A simulated annealing omit map was generated with a model excluding the RGS-DH linker, shown in red sticks, and contoured to σ = 3.0. The average *B*-factor for the inhibitory helix (a.a. 158-178) was 47.2 Å^2, as compared to 32.5 Å^2 for the entire molecule. **(C)** The RGS and DH domains also share a substantial interface. A $2F_o$-F_c map was contoured to σ = 2.5.
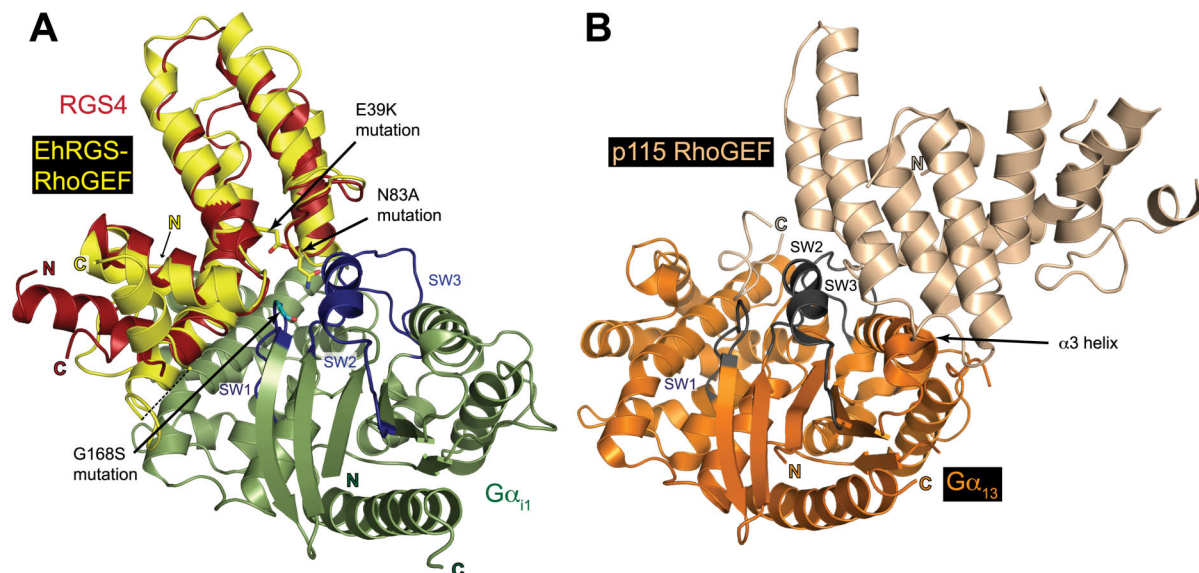
Figure 3.12. Evolutionary analysis of the EhGα1/EhRGS-RhoGEF signaling pathway. Canonical RGS domains, illustrated by RGS4 (PDB id 1AGR), and rgRGS domains of mammalian RGS-RhoGEFs, represented by that of p115 (PDB id 3AB3), exhibit distinct folds. **(A)** The EhRGS-RhoGEF RGS domain structure (*yellow*) closely resembles RGS4 (*red*), suggesting a canonical Gα/RGS domain interaction as exhibited by the RGS4/Gα$_{i1}$ complex. Canonical RGS domains engage primarily switches 1 and 2, while rgRGS domains interact with the effector interface of Gα$_{12/13}$ (*orange*) family members, primarily through switch 2 and the α3 helix, although the N-terminal extension required for GAP activity also contacts the three switch regions and the all-helical domain **(B)**. The EhGα1(G168S), EhRGS-RhoGEF(E39K), and EhRGS-RhoGEF(N83A) mutations can distinguish between the two modes of binding by selectively disrupting the canonical RGS domain binding site. Divergence of EhGα1 sequence from known mammalian subfamilies, together with the canonical nine-helix RGS domain of EhRGS-RhoGEF and its mode of EhGα1 interaction, suggest that the EhGα1/RGS-RhoGEF signaling axis arose by an evolutionary mechanism distinct from and functionally convergent with that of the mammalian Gα$_{12/13}$/RGS-RhoGEF axis.
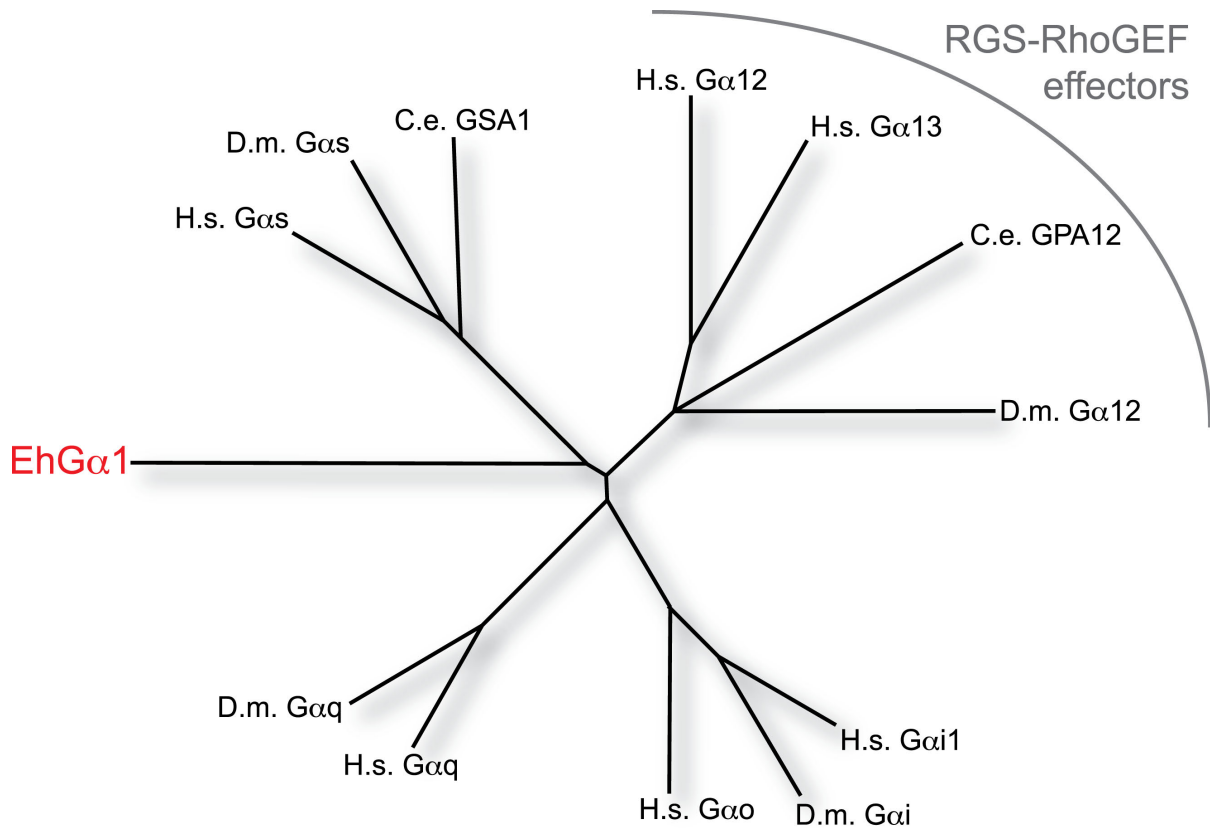
Figure 3.13. EhGα1 diverges from the Gα$_{12/13}$ subfamily despite sharing a function in binding an RGS-RhoGEF protein. An unrooted dendrogram was constructed from a multiple sequence alignment of Gα subunits from *Homo sapiens* (H.s.), *Drosophila melanogaster* (D.m.), and *Caenorhabditis elegans* (C.e.). Other Gα$_{12/13}$ subunits known to engage RGS-RhoGEF effectors cluster together, while EhGα1 exhibits low sequence similarity to each mammalian Gα subfamily.

Table 3.1. Rho family GTPase signaling and actin-associated genes differentially transcribed in *E. histolytica trophozoites* expressing EhGα1 or the dominant negative EhGα1^{S37C}.

Gene name	AmoebaDB accession no.	fold change upon EhGa1 expression (p value)	fold change upon EhGa1^{S37C} expression (p value)
Rho family signaling			
EhRGS-RhoGEF	EHI_010670	2.2 (0.001)	
hypothetical (Rho GTPase activating)	EHI_050670	2.2 (< 0.001)	
RhoGAP domain protein	EHI_098030	2.0 (< 0.001)	
RhoGAP domain protein	EHI_045100		-2.1 (< 0.001)
RhoGAP domain protein	EHI_055340		-1.7 (< 0.001)
RhoGAP domain protein	EHI_182930		1.9 (0.011)
Rho guanine nucleotide exchange factor	EHI_062800		-2.2 (0.013)
Rho guanine nucleotide exchange factor	EHI_100140		1.9 (0.003)
Rho guanine nucleotide exchange factor	EHI_131540		2.0 (< 0.001)
Rho family GTPase	EHI_029020		1.3 (0.001)
Rho family GTPase	EHI_129750		2.3 (< 0.001)
actin-associated proteins			
hypothetical (profilin/allergen family)	EHI_199670	2.0 (0.007)	
actin binding protein, filamin-like	EHI_094030	-2.0 (0.004)	
actin	EHI_008780		2.1 (< 0.001)
actobindin	EHI_039020		1.9 (< 0.001)
calponin-like actin binding domain	EHI_068450		2.6 (< 0.001)
vinculin/alpha-catenin family	EHI_083620		3.2 (< 0.001)
actinin-like	EHI_164430		2.1 (< 0.001)
actinin-like	EHI_164440		2.3 (< 0.001)
actin binding, cofilin/tropomyosin family	EHI_186840		6.0 (< 0.001)

Table 3.2. Data collection and refinement statistics for selenomethionine EhRGS-RhoGEF.

	EhRGS-RhoGEF
PDB accession code	4GOU
Data collection	
Space group	C2
Cell dimensions	
a, b, c (Å)	86.1, 46.3, 142.6
α, β, γ (°)	90, 104.2, 90
	Peak
Wavelength (Å)	0.97954
Resolution (Å)	43.0 - 2.30 (2.32 - 2.30)*
No. unique reflections	46,832
R_{merge} (%)	8.9 (58.4)**
I / sI	18.5 (2.0)
Completeness (%)	98.8 (86.4)
Redundancy	4.3 (2.5)
Wilson B-factor (Å^2)	25.6
Refinement	
Resolution (Å)	43.0 - 2.30 (2.35 - 2.30)
No. reflections	46,587 (2639)
R_{work} / R_{free} (%)	18.2 / 23.6 (26.5 / 32.0)
No. atoms	4357
Protein	4363
Ligand/ion	0
Water	216
B-factors (Å^2)	
Protein	32.5
Ligand/ion	-
Water	32.5
R.m.s deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.080

*Values in parentheses are for highest-resolution shell.
**All data were collected from a single crystal.

3.6 REFERENCES

1. Gilchrist, A., *Modulating G-protein-coupled receptors: from traditional pharmacology to allosterics.* Trends Pharmacol Sci, 2007. **28**(8): p. 431-7.

2. Oldham, W.M. and H.E. Hamm, *Heterotrimeric G protein activation by G-protein-coupled receptors.* Nat Rev Mol Cell Biol, 2008. **9**(1): p. 60-71.

3. Aittaleb, M., C.A. Boguth, and J.J. Tesmer, *Structure and function of heterotrimeric G protein-regulated Rho guanine nucleotide exchange factors.* Mol Pharmacol, 2010. **77**(2): p. 111-25.

4. Kimple, A.J., D.E. Bosch, P.M. Giguere, and D.P. Siderovski, *Regulators of G-protein signaling and their Galpha substrates: promises and challenges in their use as drug discovery targets.* Pharmacol Rev, 2011. **63**(3): p. 728-49.

5. Tesmer, J.J., D.M. Berman, A.G. Gilman, and S.R. Sprang, *Structure of RGS4 bound to AlF4--activated G(i alpha1): stabilization of the transition state for GTP hydrolysis.* Cell, 1997. **89**(2): p. 251-61.

6. Rossman, K.L. and J. Sondek, *Larger than Dbl: new structural insights into RhoA activation.* Trends Biochem Sci, 2005. **30**(4): p. 163-5.

7. Chen, Z., L. Guo, S.R. Sprang, and P.C. Sternweis, *Modulation of a GEF switch: autoinhibition of the intrinsic guanine nucleotide exchange activity of p115-RhoGEF.* Protein Sci, 2011. **20**(1): p. 107-17.

8. Chen, Z., et al., *Activation of p115-RhoGEF requires direct association of Galpha13 and the Dbl-homology domain.* J Biol Chem, 2012.

9. Zheng, M., et al., *On the mechanism of autoinhibition of the RhoA-specific nucleotide exchange factor PDZRhoGEF.* BMC Struct Biol, 2009. **9**: p. 36.

10. Bielnicki, J.A., et al., *Insights into the molecular activation mechanism of the RhoA-specific guanine nucleotide exchange factor, PDZRhoGEF.* J Biol Chem, 2011. **286**(40): p. 35163-75.

11. WHO, *WHO/PAHO/UNESCO report. A consultation with experts on amoebiasis. Mexico City, Mexico 28-29 January, 1997.* Epidemiol Bull, 1997. **18**(1): p. 13-4.

12. Ralston, K.S. and W.A. Petri, Jr., *Tissue destruction and invasion by Entamoeba histolytica.* Trends Parasitol, 2011. **27**(6): p. 254-63.

13. Guillen, N., *Role of signalling and cytoskeletal rearrangements in the pathogenesis of Entamoeba histolytica.* Trends Microbiol, 1996. **4**(5): p. 191-7.

14.   Bosch, D.E., E.S. Wittchen, C. Qiu, K. Burridge, and D.P. Siderovski, *Unique structural and nucleotide exchange features of the Rho1 GTPase of Entamoeba histolytica.* J Biol Chem, 2011. **286**(45): p. 39236-46.

15.   Meza, I., P. Talamas-Rohana, and M.A. Vargas, *The cytoskeleton of Entamoeba histolytica: structure, function, and regulation by signaling pathways.* Arch Med Res, 2006. **37**(2): p. 234-43.

16.   Bosch, D.E., et al., *Heterotrimeric G-protein Signaling Is Critical to Pathogenic Processes in Entamoeba histolytica.* PLoS Pathog, 2012. **8**(11): p. e1003040.

17.   Stols, L., et al., *A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site.* Protein Expr Purif, 2002. **25**(1): p. 8-15.

18.   Johnston, C.A., et al., *Structural determinants underlying the temperature-sensitive nature of a Galpha mutant in asymmetric cell division of Caenorhabditis elegans.* J Biol Chem, 2008. **283**(31): p. 21550-8.

19.   Otwinowski, Z.a.W.M., *Processing of X-ray Diffraction Data Collected in Oscillation Mode*, in *Methods in Enzymology*. 1997, Academic Press: New York. p. 307-326.

20.   Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.

21.   Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2126-32.

22.   Painter, J. and E.A. Merritt, *Optimal description of a protein structure in terms of multiple groups undergoing TLS motion.* Acta Crystallogr D Biol Crystallogr, 2006. **62**(Pt 4): p. 439-50.

23.   Chen, V.B., et al., *MolProbity: all-atom structure validation for macromolecular crystallography.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 1): p. 12-21.

24.   Bosch, D.E., et al., *Structural determinants of affinity enhancement between GoLoco motifs and G-protein alpha subunit mutants.* J Biol Chem, 2010. **286**(5): p. 3351-8.

25.   Kimple, A.J., R.E. Muller, D.P. Siderovski, and F.S. Willard, *A capture coupling method for the covalent immobilization of hexahistidine tagged proteins for surface plasmon resonance.* Methods Mol Biol, 2010. **627**: p. 91-100.

26.   Hutsell, S.Q., R.J. Kimple, D.P. Siderovski, F.S. Willard, and A.J. Kimple, *High-affinity immobilization of proteins using biotin- and GST-based coupling strategies.* Methods Mol Biol, 2010. **627**: p. 75-90.

27.   Hamann, L., H. Buss, and E. Tannich, *Tetracycline-controlled gene expression in Entamoeba histolytica.* Mol Biochem Parasitol, 1997. **84**(1): p. 83-91.

28.   Olvera, A., et al., *Stable transfection of Entamoeba histolytica trophozoites by lipofection.* Arch Med Res, 1997. **28 Spec No**: p. 49-51.

29.   Gilchrist, C.A., et al., *Targets of the Entamoeba histolytica transcription factor URE3-BP.* PLoS Negl Trop Dis, 2008. **2**(8): p. e282.

30.   Shrimal, S., S. Bhattacharya, and A. Bhattacharya, *Serum-dependent selective expression of EhTMKB1-9, a member of Entamoeba histolytica B1 family of transmembrane kinases.* PLoS Pathog, 2010. **6**(6): p. e1000929.

31.   Dolabella, S.S., et al., *Amoebic liver abscess production by Entamoeba dispar.* Ann Hepatol, 2012. **11**(1): p. 107-17.

32.   Hirata, K.K., et al., *A phagocytosis mutant of Entamoeba histolytica is less virulent due to deficient proteinase expression and release.* Exp Parasitol, 2007. **115**(2): p. 192-9.

33.   Rogers, S.L. and G.C. Rogers, *Culture of Drosophila S2 cells and their use for RNAi-mediated loss-of-function studies and immunofluorescence microscopy.* Nat Protoc, 2008. **3**(4): p. 606-11.

34.   Chen, Z., W.D. Singer, P.C. Sternweis, and S.R. Sprang, *Structure of the p115RhoGEF rgRGS domain-Galpha13/i1 chimera complex suggests convergent evolution of a GTPase activator.* Nat Struct Mol Biol, 2005. **12**(2): p. 191-7.

35.   Rogers, S.L., U. Wiedemann, N. Stuurman, and R.D. Vale, *Molecular requirements for actin-based lamella formation in Drosophila S2 cells.* J Cell Biol, 2003. **162**(6): p. 1079-88.

36.   Chen, Z., et al., *Activated RhoA binds to the pleckstrin homology (PH) domain of PDZ-RhoGEF, a potential site for autoregulation.* J Biol Chem, 2010. **285**(27): p. 21070-81.

37.   Petri, W.A., Jr., R. Haque, and B.J. Mann, *The bittersweet interface of parasite and host: lectin-carbohydrate interactions during human invasion by the parasite Entamoeba histolytica.* Annu Rev Microbiol, 2002. **56**: p. 39-64.

38.   Soundararajan, M., et al., *Structural diversity in the RGS domain and its interaction with heterotrimeric G protein alpha-subunits.* Proc Natl Acad Sci U S A, 2008. **105**(17): p. 6457-62.

39.   Ferguson, K.M., et al., *Structural basis for discrimination of 3-phosphoinositides by pleckstrin homology domains.* Mol Cell, 2000. **6**(2): p. 373-84.

40.   Snyder, J.T., et al., *Structural basis for the selective activation of Rho GTPases by Dbl exchange factors.* Nat Struct Biol, 2002. **9**(6): p. 468-75.

41.   Das, B., et al., *Control of intramolecular interactions between the pleckstrin homology and Dbl homology domains of Vav and Sos1 regulates Rac binding.* J Biol Chem, 2000. **275**(20): p. 15074-81.

42.   Lan, K.L., et al., *A point mutation in Galphao and Galphai1 blocks interaction with regulator of G protein signaling proteins.* J Biol Chem, 1998. **273**(21): p. 12794-7.

43.   Meigs, T.E., et al., *Selective uncoupling of G alpha 12 from Rho-mediated signaling.* J Biol Chem, 2005. **280**(18): p. 18049-55.

44.   Suzuki, N., S. Nakamura, H. Mano, and T. Kozasa, *Galpha 12 activates Rho GTPase through tyrosine-phosphorylated leukemia-associated RhoGEF.* Proc Natl Acad Sci U S A, 2003. **100**(2): p. 733-8.

45.   Gonzalez De la Rosa, C.H., et al., *EhGEF2, a Dbl-RhoGEF from Entamoeba histolytica has atypical biochemical properties and participates in essential cellular processes.* Mol Biochem Parasitol, 2007. **151**(1): p. 70-80.

CHAPTER 4

UNIQUE STRUCTURAL AND NUCLEOTIDE EXCHANGE FEATURES OF THE

Rho1 GTPase OF Entamoeba histolytica[1]

4.1   OVERVIEW

The single-celled human parasite *Entamoeba histolytica* possesses a dynamic actin
cytoskeleton vital for its intestinal and systemic pathogenicity. The *E. histolytica* genome
encodes several Rho family GTPases known to regulate cytoskeletal dynamics. EhRho1, the
first family member identified, was reported to be insensitive to the Rho GTPase-specific *C.
botulinum* C3 exoenzyme, raising the possibility that it may be a mis-classified Ras family
member. Here, we report the crystal structures of EhRho1 in both active and inactive states.
EhRho1 is activated by a conserved switch mechanism, but diverges from mammalian Rho
GTPases in lacking a signature Rho insert helix. EhRho1 engages a homolog of mDia,
EhFormin1, suggesting a role in mediating serum-stimulated actin reorganization and
microtubule formation during mitosis. EhRho1, but not a constitutively active mutant,
interacts with a newly-identified EhRhoGDI in a prenylation-dependent manner.
Furthermore, constitutively active EhRho1 induces actin stress fiber formation in mammalian
fibroblasts, thereby identifying it as a functional Rho family GTPase. EhRho1 exhibits a fast
rate of nucleotide exchange relative to mammalian Rho GTPases due to a distinctive switch

---

one isoleucine residue reminiscent of the constitutively active F28L mutation in human Cdc42 which, for the latter protein, is sufficient for cellular transformation. Non-conserved, nucleotide-interacting residues within EhRho1, revealed by the crystal structure models, were observed to contribute a moderating influence on fast spontaneous nucleotide exchange. Collectively, these observations indicate that EhRho1 is a *bona fide* member of the Rho GTPase family, albeit with unique structural and functional aspects compared to mammalian Rho GTPases.

4.2   INTRODUCTION

The parasite *Entamoeba histolytica* is the causative agent of amoebiasis in humans, responsible for an estimated 50 million infections and 100,000 deaths per year worldwide [1]. Spread primarily by contaminated drinking water in its encysted form, *E. histolytica* infection is endemic among poor populations of developing countries, although outbreaks and infection among travelers occur frequently in the United States [2]. The water-borne pathogen can attach and invade intestinal mucosa to cause amoebic colitis, and can also enter the blood stream, leading to systemic amoebiasis characterized by liver, lung, and brain abscesses [3]. *E. histolytica* trophozoites are highly motile and undergo complex, dynamic cytoskeletal rearrangements [4]. Indeed, the cytoskeletal dynamics of *E. histolytica* are vital for many of its pathogenic processes, including chemotaxis and invasion, adhesion to intestinal epithelia, phagocytosis, and host cell killing [5].

Rho family GTPases are small guanine nucleotide binding proteins of the Ras small G-protein superfamily [6] that produce multiple effects in cells when activated. The most prominent and immediate effect of Rho activation is actin cytoskeletal reorganization [7].

Rho GTPases are molecular switches, with an active GTP-bound state and an inactive GDP-bound conformation. Conformational shifts in response to nucleotide exchange are dominated by two conserved switch regions which contribute to nucleotide state-selective engagement of multiple Rho-interacting proteins [8]. Switch 1 contains a highly conserved phenylalanine that forms aromatic interactions with the nucleotide guanine ring [9]. Rho GTPases are regulated by the actions of guanine nucleotide exchange factors (GEFs), GTPase activating proteins (GAPs) and guanine nucleotide dissociation inhibitors (GDIs) [10]. RhoGDIs preferentially bind inactive, GDP-bound Rho GTPases and are critical for shuttling them between cellular membranes [11]. Among the many effectors of activated human Rho GTPases is mDia, a formin protein that produces actin filaments by initiating nucleation and polymerization [12].

The genome of *E. histolytica* encodes a family of small G-proteins with high sequence similarity to mammalian Rho GTPases [13, 14]. In fact, a remarkably large Rho family of at least 19 members is simultaneously expressed in *E. histolytica* trophozoites (Figure 4.1). Although studies are limited, certain *E. histolytica* Rho family GTPases have been shown to regulate its actin cytoskeletal dynamics, which are in turn linked to pathogenic processes. Overexpression of a constitutively active point mutant of the GTPase EhRacA in *E. histolytica* leads to altered phagocytic activity and surface receptor capping - a phenomenon vital for evasion of the host immune response [15]. Overexpression of EhRacG alters membrane turnover, uroid formation, and surface receptor capping [16]. Not surprisingly, Rho-activating GEFs have also emerged as critical players in *E. histolytica* pathogenesis [17, 18]. Similarly, *E. histolytica* possesses homologs to key Rho GTPase effectors, such as p21-activated kinase (PAK) and Diaphanous-like formins [19, 20].

Although none of the formin family members have been shown to interact with Rho GTPases from *E. histolytica*, three of them (EhFormin1-3) contain Rho G̲TPase b̲inding d̲omains (GBDs), suggesting that they serve as a link between Rho GTPase activation and actin polymerization. Interestingly, overexpression of EhFormin1 in *E. histolytica* trophozoites revealed its co-localization with actin assemblies promoted by serum factors, association with microtubules during mitosis, and aberration of cell division [20].

As the first identified [13] and highly expressed Rho GTPase family member in *E. histolytica* (Figure 4.1), EhRho1 serves as an exemplary small G-protein signaling molecule from *E. histolytica*. Protein sequence and biochemical analyses have suggested divergent guanine nucleotide binding motifs and a resistance to inhibitory ADP-ribosylation by *Clostridium botulinum* C3 exoenzyme, a hallmark of mammalian Rho GTPases [21]. Thus, Godbold et al. have suggested that EhRho1 may be a Ras-like GTPase rather than a true functional ortholog of mammalian Rho GTPases [21]. In the current study, we have resolved this functional categorization of EhRho1 *via* structural models of EhRho1 in two nucleotide states, obtained by X-ray diffraction crystallography. EhRho1 possesses both Rho- and Ras-like structural features. While possessing a conserved structural mode of activation with mammalian G-proteins, multiple divergent residues in the nucleotide-binding pocket were seen to contribute to a fast basal exchange rate relative to mammalian Rho GTPases. EhRho1 binds to EhFormin1 and to a novel EhRhoGDI in a nucleotide-dependent fashion. Finally, expression of constitutively active EhRho1 in mammalian cells induces stress fiber formation, implicating EhRho1 in the regulation of actin cytoskeletal dynamics in *Entamoeba histolytica*.

4.3   EXPERIMENTAL PROCEDURES

4.3.1 Bioinformatic analysis of the *E. histolytica* Rho GTPase family

To identify Rho GTPase genes in the sequenced genome of *E. histolytica* [22] the amino acid sequence of EhRho1 (GenBank accession no. XP_654488.2) was used as a template for a sequence similarity search in the NCBI database with the BLAST algorithm [23]. Resultant candidate Rho proteins and associated transcripts were located in the publicly available microarray data [24]. An average expression unit value was calculated for each transcript using replicate data for unperturbed *E. histolytica* HM1:IMSS trophozoites. Rho genes that showed <3 expression units (considered to be not expressed based on a frequency distribution of all transcripts (Figure 4.1)) and protein sequences that were >80% identical to another *E. histolytica* Rho GTPase were excluded from further analysis to avoid inclusion of closely related isoforms and potential microarray probe cross-reactivity. The remaining 19 Rho GTPase protein sequences were aligned with ClustalW2 [25] and a dendrogram generated with NJPlot [26]. Relative expression levels were assigned to each Rho GTPase transcript based upon a frequency distribution of all genes included on the microarray platform *versus* relative expression unit values (Figure 4.1).

4.3.2 Protein expression and purification

Open reading frames of *EhRho1* lacking the C-terminal CAAX prenylation motif and its preceding polybasic region (amino acids 1-191), *EhRacg*, *EhRhogdi* (UniProt identifier O76754), and the *EhFormin1* GBD-FH3 domains (amino acids 69-445) were separately amplified from *E. histolytica* genomic DNA (obtained from Dr. William Petri Jr., University of Virginia, Charlottesville, VA) by polymerase chain reaction (PCR) using Phusion

polymerase (New England BioLabs, Ipswich, MA) and primers from Invitrogen (*EhRho1* (aa 1-191) : sense, 5'- ATGCTTGCATTTTCTGATATGAAC-3'; antisense 5'-CTAATTTGAGAAGATACAATC-3'; *EhRacg*: sense, 5'-ATGAGACCAGTGAAACTT GTC-3'; antisense, 5'-CTATTTAGCAGCTTTAGCAAGAAC-3'; *EhRhogdi*: sense, 5'-ATGTCAGCAGCAGACATTGTTAAAAAC-3'; antisense, 5'-TTAATCCCAATCCTT GGC-3'; *EhFormin1*: sense, 5'-ATGCCACCTGAAGAAGTTG-3'; antisense, 5'-CTAAGTGACTTGAAGAGATATTTGC-3'. PCR amplicons were resolved on, and extracted from, a 1% (w/v) agarose gel using the Qiagen Gel Extraction Kit. All three amplicons were subcloned using ligation-independent cloning [27] into a Novagen (San Diego, CA) pET vector-based prokaryotic expression construct ("pET-His LIC-C" or "pET-GST LIC-C") to form N-terminal tobacco etch virus (TEV) protease-cleavable, hexahistidine- or GST-tagged fusions. Point mutations to *EhRho1* and *Hs RHOA* were made using PCR and the overlap extension method [28]. Clones of human *RHOA* and mouse *mDia1* were a kind gift from Dr. John Sondek (University of North Carolina, Chapel Hill, NC).

For expression of hexahistidine-tagged EhRho1, EhRacG, Hs RhoA, mDia1 (a.a. 69-450), and EhFormin1 and GST-tagged RhoGDI, BL21(DE3) *E. coli* were grown to an OD_{600nm} of 0.8 at 37°C, followed by induction with 500 μM isopropyl-β-D-thiogalactopyranoside (IPTG) for 14-16 hours at 20°C. Bacterial cells were pelleted by centrifugation and resuspended in N1 buffer composed of 50 mM Tris pH 7.5, 250 mM NaCl, 10 mM imidazole, 1 mM DTT, 5% (v/v) glycerol, and for Rho GTPases, 5 mM $MgCl_2$ and 50 μM GDP. Bacteria were lysed at 10,000 kPa using pressure homogenization with an Emulsiflex (Avestin; Ottawa, Canada). Cellular lysates containing hexahistidine-tagged

proteins were cleared with centrifugation at 100,000 x g for 60 minutes at 4°C, and the resulting supernatant was applied to a nickel-nitrilotriacetic acid (NTA) resin FPLC column (FF HisTrap crude; GE Healthcare, Piscataway, NJ), washed with N1 plus 30 mM imidazole before elution in N1 buffer with 300 mM imidazole. Lysates containing GST-tagged proteins were applied to a glutathione resin (FF GSTrap; GE Heathcare), washed with N1 buffer lacking imidazole, and eluted with imidazole-free N1 supplemented with 10 mM reduced glutathione. For GST-RhoGDI, His$_6$-EhFormin1, and His$_6$-mDia1, eluted protein was pooled and resolved using a calibrated size exclusion column (HiLoad 16/60 Superdex 200, GE Healthcare) in S200 buffer (50 mM HEPES pH 8.0, 250 mM NaCl, 2.5% (v/v) glycerol, and 1 mM DTT). For the Rho GTPases, protein was pooled and dialyzed into imidazole-free N1 overnight at 4°C in the presence of His$_6$-tobacco etch virus (TEV) protease to cleave the N-terminal affinity tag. The dialysate was then passed over a second NTA column to remove TEV protease and uncleaved protein. For exchange assays, Rho GTPases were resolved by size exclusion in Rho S200 buffer (50 mM Tris pH 8.0, 200 mM NaCl, 10 mM MgCl$_2$, and 5% (v/v) glycerol). For crystallization, EhRho1 was loaded with either GDP or GTPγS by incubation in EDTA-containing exchange buffer (50 mM Tris pH 7.5, 250 mM NaCl, 5 mM DTT, 15 mM EDTA, and 10-fold molar excess of either GDP or GTPγS) at room temperature for 45 minutes, followed by addition of excess nucleotide-stabilizing magnesium ion (50 mM MgCl$_2$). Nucleotide-loaded EhRho1 was then resolved by size exclusion chromatography in crystallization buffer (50 mM Tris pH 8.0, 250 mM NaCl, 1 mM MgCl$_2$, 5 mM DTT, and either 50 μM GDP or 5 μM GTPγS). All proteins were concentrated to 0.5 – 2 mM and snap frozen in a dry ice/ethanol bath for storage at -80°C. Protein concentration was determined by A$_{280nm}$ measurements upon denaturation in 8 M guanidine hydrochloride,

based on predicted extinction coefficients for each protein (http://us.expasy.org/tools/protparam.html).

4.3.3 Crystallization of EhRho1·GDP and EhRho1·GTPγS and structure determination

Crystals of EhRho1 (residues 1-191) bound to GDP were obtained by vapor diffusion from hanging drops at 18°C. EhRho1·GDP at 15 mg/mL in crystallization buffer was mixed 1:1 with (and equilibrated against) crystallization solution containing 1.5 M ammonium sulfate and 100 mM Tris pH 8.0. Rhomboidal crystals grew to ~200 x 100 x 100 μm over 5 days, exhibiting the symmetry of space group $P2_12_12_1$ (a = 50.3 Å, b = 54.5 Å, c = 132.3 Å, α = β = γ = 90°) and containing two monomers in the asymmetric unit. For data collection at 100K, crystals were serially transferred for ~1 minute into crystallization solution supplemented with 25% (v/v) glycerol in 5% increments and plunged into liquid nitrogen. A native data set was collected at the SER-CAT 22-ID beamline at the Advanced Photon Source (Argonne National Laboratory). Data were processed using the HKL-2000 program [29]. The crystal structure model of human RhoA·GDP (PDB accession 1FTN [9]) was modified by removal of water, magnesium, and nucleotide and by trimming of side chains not conserved in EhRho1 using Chainsaw [30]. The resulting model was used as a molecular replacement search model in the program Phaser [31]. Refinement was carried out using phenix.refine [32] interspersed with manual revisions of the model using the program Coot [33]. Refinement consisted of conjugate-gradient minimization and calculation of individual anisotropic displacement and translation/libration/screw (TLS) parameters [34]. The current model contains two EhRho1 monomers bound to GDP and magnesium; residues at the N-

and C-termini (1-14 for both chains and 187-194 for chain A) could not be located in the electron density.

Crystals of EhRho1 bound to GTPγS were obtained by vapor diffusion from hanging drops at 18°C. EhRho1·GTPγS at 14 mg/mL in crystallization buffer was mixed 1:1 with (and equilibrated against) crystallization solution containing 25% PEG 4000, 150 mM ammonium acetate, and 100 mM sodium acetate pH 4.6. Rod crystals grew to ~200 x 75 x 50 mm over 3 days and exhibited the symmetry of spacegroup P1 (a = 36.4 Å, b = 39.5 Å, c = 63.6 Å, α = 81.8°, β = 80.8°, γ = 65.4°) with two monomers in the asymmetric unit. For data collection at 100K, crystals were serially transferred for ~30 seconds into crystallization solution supplemented with 30% saturated sucrose in 10% increments and plunged into liquid nitrogen. A native data set was collected at the SER-CAT 22-BM beamline at the Advanced Photon Source (Argonne National Laboratory). Data processing and refinement were carried out similarly to EhRho1·GDP, above. However, a molecular replacement solution was obtained using the crystal structure model of human RhoA bound to GTPγS (PDB accession 1A2B [35]) modified to exclude water, magnesium, nucleotide, and non-conserved side chain atoms. The current model contains two EhRho1 monomers bound to GTPγS and magnesium; residues at the N- and C-termini (1-20 and 192-194 for chain A and 1-20 and 189-194 for chain B) could not be located in the electron density. Despite low diffraction data completeness in the high resolution shells for EhRho1·GTPγS, strong electron density arose for GTPγS and missing side chains upon molecular replacement solution and no systematic defects were identified in the electron density map (see Figure 4.2). To ensure model accuracy, the diffraction data were processed in parallel with a 2.5 Å resolution cutoff, producing a model with no observed differences from the high-resolution

inclusive data. However, inclusion of the less complete, high resolution data (2.5 – 1.8 Å) substantially increased the electron density quality. For data collection and refinement statistics, see Table 4.1. All structural images were rendered with PyMOL (Schrödinger LLC, Portland, OR) unless otherwise indicated.

4.3.4 Surface plasmon resonance (SPR) binding assays

SPR-based measurements of protein-protein interactions were performed on a Biacore 3000 (GE Healthcare). Approximately 8000 resonance units (RUs) of purified His_6-EhFormin1 and 10,000 RUs of His_6-mDia1 GBD-FH3 domain tandem were separately immobilized on a nickel-charged NTA biosensor chip (GE Healthcare) using covalent capture coupling as previously described [36]. An irrelevant hexahistidine protein (Hs $G\alpha_{i1}$) was loaded on an independent surface as a negative control. SPR experiments were performed in running buffer containing 50 mM HEPES pH 7.4, 150 mM NaCl, 0.05% NP-40 alternative (Calbiochem), 50 μM EDTA, 1 mM $MgCl_2$, and 5 μM of either GDP or GTPγS. Thirty μL injections of increasing concentrations of GDP- or GTPγS-loaded EhRho1, EhRacG, and Hs RhoA were separately performed at 10 μL/min with a 300 second dissociation phase (using the KINJECT command). The surface was regenerated between Rho GTPase injections by injection of running buffer supplemented with 30 mM EDTA. To correct for non-specific binding and changes in the refractive index upon injection of samples, the observed sensorgram for the flow cell containing the irrelevant protein Hs $G\alpha_{i1}$ was subtracted from all curves using BIAevaluation software v3.0 (GE Healthcare). Equilibrium binding analyses were conducted as previously described [37] using GraphPad Prism v5.0 to determine binding affinities.

4.3.5 GST-EhRhoGDI affinity co-precipitation

The full-length *EhRho1* gene was synthesized following codon optimization for expression in mammalian cells (GeneArt, Invitrogen). *EhRho1* and mutants were subcloned with an HA epitope tag into pcDNA 3.1 and transfected into HEK293T cells using Lipofectamine 2000 (Invitrogen), 1 μg of DNA per 6-well dish. Cells expressing wild-type or mutant EhRho1 were lysed in buffer containing 20 mM Tris pH 7.5, 100 mM NaCl, 1 mM $MgCl_2$. 1% (v/v) NP-40, 0.25% (v/v) deoxycholate, and protease inhibitors. Soluble fractions of the cell lysates were incubated with 15 μg of either recombinant GST-RhoGDI or GST alone and glutathione agarose (Sigma) at 4°C overnight with gentle agitation. Bound proteins were separated by centrifugation, washed 3 times with lysis buffer, and eluted with denaturing SDS-PAGE loading buffer. Proteins/lysates were separated by SDS-PAGE, transferred to nitrocellulose membranes, and subjected to western blotting with anti-HA as described previously [38]. Subsequently, membrane-bound proteins were stained with Ponceau S (Bio-Rad).

4.3.6 Fluorescent guanine nucleotide exchange assays

The fluorescence of BODIPY FL GDP (Invitrogen) was monitored in real-time using a cuvette-based LS 55 spectrometer (Perkin Elmer, Waltham, MA) under thermostat control. Excitation and emission wavelengths were 502 ± 2.5 nm and 511 ± 2.5 nm, respectively. Exchange buffer alone (20 mM Tris pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 2 mM DTT, 5% (v/v) glycerol, 0.008% NP-40 alternative, and 75 nM BODIPY FL GDP) was monitored at 15°C until a stable signal was achieved. At time zero, 400 nM Rho GTPase was added to the

cuvette with mixing. The increase in relative fluorescence of BODIPY FL GDP upon incorporation into the Rho GTPase nucleotide-binding pocket was monitored at 30 second intervals for 100 minutes. Since saturation exchange was not reached due to the characteristically slow release of GDP from Rho GTPases in the absence of guanine nucleotide exchange factor (GEF), 30 mM EDTA was added to the cuvette at the end of each time course, inducing rapid exchange; the fluorescence intensity after equilibration with excess EDTA was thus defined as 100% exchange. Each curve was fitted with an exponential association function to yield a rate constant, k_{obs}, using GraphPad Prism v5.0.

4.3.7 Actin stress fiber quantification

Rat-2 fibroblasts were transfected with HA-tagged constitutively active human RhoA(G14V), EhRho1(Q78L), or empty vector control (pcDNA3.1) and plated onto fibronectin-coated coverslips (15 μg/mL) overnight. Transfection efficiency was 80-90%. Cells were fixed in 3.7% paraformaldehyde, and permeablized with 0.02% (v/v) Triton X-100. F-actin structures were stained with Texas Red phalloidin (Molecular Probes) and exogenously expressed Rho GTPases were visualized using anti-HA immunofluorescence. To quantify stress fibers, HA-positive cells were scored by a blinded observer for the presence or absence of stress fibers, with the criteria being: organized, thickened, parallel actin bundles throughout the majority of the cytoplasm. Results are plotted as percent cells positive for stress fiber phenotype, scoring ~200 cells from many fields in duplicate experiments. Statistical significance was determined by Student's t-test.

4.4 RESULTS

4.4.1 Comparison of human and *E. histolytica* Ras superfamily GTPases

EhRho1 (GenBank accession no. XP_654488.2) is most similar to RhoA (Uniprot P61586) among the human Ras superfamily of GTPases. Sequence identity of only 47% between these two proteins likely reflects unique features gained or lost across a relatively large evolutionary distance. A multiple sequence alignment of Rho and Ras family GTPases from *E. histolytica* and humans (Figure 4.3) identified unique and conserved residues in EhRho1. EhRho1 displays a conserved cysteine at position 35, which has been implicated as an important site for reactive oxygen and nitrogen species-mediated activation of human Rho GTPases [39]. Switch 2 is remarkably conserved from Entamoeba to humans, but switch 1 diverges across species. Notably, EhRho1 displays an isoleucine at position 45 in place of the otherwise highly conserved, nucleotide-contacting phenylalanine. Mutation of the analogous residue in human Rho GTPases results in ~100-fold faster nucleotide exchange than wildtype in the absence of a GEF, resulting in constitutive activity [40]. In fact, the corresponding mutation in Hs Cdc42, F28L, can induce cellular transformation in fibroblasts [41]. This unique residue has led others to postulate that EhRho1 is constitutively active [21]. Two other putative nucleotide-contacting residues in EhRho1 differ from the Rho/Ras consensus, namely, Ser-166 and Val-167. These sequence features of EhRho1 taken together have suggested a unique mechanism of nucleotide exchange and thus activation [21]. To investigate the structural determinants of EhRho1 activation, we obtained high-resolution structural models of EhRho1 in two nucleotide states by X-ray diffraction crystallography.

4.4.2 Structures of EhRho1 in the active and inactive states

Since Rho GTPases possess a conserved C-terminal prenylation site, the CaaX box, and an adjacent polybasic region that are not typically well-ordered in a crystalline state, we crystallized a truncated form of EhRho1 (residues 1-191) bound to a divalent magnesium ion and either GDP or the non-hydrolyzable GTP analog, GTPγS. Both forms of EhRho1 produced well-diffracting crystals, with diffraction data extending to 1.95 Å and 1.80 Å resolution, respectively. Structural statistics of both complexes are listed in Table 4.1. Phases were resolved by molecular replacement using the crystal structure models of human RhoA bound to GDP [9] and GTPγS [35], respectively (PDB id 1FTN and 1A2B).

EhRho1 exhibits the highly conserved G-domain fold characteristic of Ras superfamily GTPases [10], consisting of a 6-stranded β-sheet surrounded by 5 α-helices. The core fold of EhRho1 is similar to that of both human RhoA (PDB 1A2B) and H-Ras (PDB 5P21), with total Cα r.m.s.d. values of 1.1 Å and 2.0 Å, respectively. Superposition of the two EhRho1 conformations (Figure 4.4A) and a plot of Cα r.m.s.d per-residue between the EhRho1 structures (Figure 4.4B) highlights switch regions 1 and 2 as the most mobile areas associated with activation. The switch 1 loop is drawn closer to the nucleotide in the GTPγS-bound form than when bound to GDP, and the N-terminal portion of switch 2 rearranges upon binding GTPγS due to polar contacts with the γ-phosphoryl group.

The switch regions of both monomers in the EhRho1·GDP structural model form minimal crystal contacts with neighboring molecules. Only Glu-80 on switch 2 of chain A forms a polar contact with a neighboring molecule. Chain B switch residues Tyr-49, Thr-55, and Tyr-81 also form hydrogen bonds with residues of another asymmetric unit. Despite these differing crystal contacts, both EhRho1·GDP monomers display switch conformations

distinct from those of EhRho1·GTPγS, indicating that the observed switch conformation shift is not likely due to crystal packing.

The crystal structures of EhRho1 also revealed the absence of a signature Rho insert helix in the β5-α4 loop, a key feature conserved among all other known Rho family GTPases (Figure 4.4C) [9, 42]. Instead, the β5-α4 loop lacks uniform secondary structure, as seen in Ras family GTPases [43]. The insert helix is not vital for the interaction of human Rho GTPases with the majority of interacting partners [44, 45]. However it is required for some effector interactions, such as Cdc42-mediated activation of phospholipase D1 [46] and activation of NADPH oxidase by Rac [47]. The absence of an insert helix, in combination with its insensitivity to the Rho-specific *Clostridium botulinum* C3 exotoxin [21], led us to ask whether EhRho1 was actually a mis-classified Ras family GTPase. Accordingly, we looked for interaction of EhRho1 with *E. histolytica* homologs of classical effectors.

4.4.3 EhRho1 interacts with an mDia homolog, EhFormin1

The *E. histolytica* genome encodes a family of eight Formins [20]. EhFormins 1-3 contain a Rho GTPase binding domain (GBD) in tandem with a formin homology 3 (FH3) domain, homologous to mammalian mDia and *Drosophila melanogaster* Diaphanous [12]. EhFormin1, also known as EhDia, is known to regulate actin polymerization and cell cycle progress in *E. histolytica* trophozoites [20]. We cloned and purified the GBD-FH3 tandem of EhFormin1 (Figure 4.3B) and used surface plasmon resonance (SPR) to quantify binding to EhRho1 in each nucleotide state. EhFormin1 was observed to selectively bind EhRho1 (K_D = 1.7 ± 0.1 μM), but not the related GTPase EhRacG (Figure 4.5). The interaction was dependent on the nucleotide state of EhRho1 (Figure 4.5B), characteristic of a

GTPase/effector interaction. The affinity of EhRho1·GTPγS/EhFormin1 was an order of magnitude lower than that observed for Hs RhoA·GppNHp/mDia1 ($K_D = 104 \pm 37$ nM) under identical conditions (Figure 4.5C, D). However, this low μM affinity EhRho1/EhFormin1 interaction is likely relevant in a cellular context, given the observed stringent nucleotide state selectivity and previous evidence for colocalization of both proteins in the *E. histolytica* uropod [48]. We conclude that EhRho1 engages similar effectors to mammalian Rho family GTPases.

4.4.4 EhRho1 interacts with a newly-identified RhoGDI

Rho guanine dissociation inhibitors (GDIs) maintain a pool of soluble, inactive Rho GTPases by extracting prenylated GTPases from cellular membranes [11]. The *E. histolytica* genome encodes a single gene with a conserved RhoGDI domain, which we refer to as EhRhoGDI (UniProt identifier O76754). We cloned and purified EhRhoGDI as a GST fusion and used the recombinant protein to co-precipitate EhRho1 expressed in HEK293T cells (Figure 4.6). Mammalian RhoGDIs contain a geranylgeranyl-binding pocket and favor binding to the inactive, GDP-bound conformation of Rho GTPases [11]. To examine the nucleotide state selectivity and Rho prenylation dependency of the EhRho1/EhRhoGDI interaction, we compared co-precipitation of a GTPase-deficient, constitutively active EhRho1 mutant (Q78L) and EhRho1 lacking the putatively-prenylated CaaX motif cysteine (C212S) to that of wildtype EhRho1. While wild type EhRho1 was seen to robustly interact with GST-EhRhoGDI in cellular lysates, the complex was disrupted by either EhRho1 constitutive activity or mutation of the putative prenylation site (Figure 4.6B). Consistent with a requirement for prenylation of EhRho1 to engage EhRhoGDI, we also did not observe

binding between GST-EhRhoGDI and C-terminally truncated EhRho1 produced recombinantly from *E. coli* (not shown).

4.4.5 EhRho1 stimulates stress fiber formation in mammalian cells

One of the most prominent and immediate effects of human RhoA activation in fibroblasts is the formation of filamentous actin bundles known as stress fibers [7]. To determine the behavior of EhRho1 in a cellular context, we expressed GTPase deficient, constitutively active EhRho1(Q78L) or constitutively active human RhoA in Rat-2 cells and examined stress fiber formation with phalloidin staining. Both constitutively active EhRho1 and Hs RhoA expression significantly induced stress fiber formation compared to mock-transfected Rat-2 cells (Figure 4.7). Thus, EhRho1 modulates cellular actin structures, likely through signaling pathways similar to human RhoA. These findings, together with EhRho1 interaction with the canonical effectors EhFormin1 and EhRhoGDI, imply that EhRho1 functions as true Rho family member, despite its Ras-like lack of an insert helix (Figure 4.4) and insensitivity to the Rho-specific C3 exoenzyme [21].

4.4.6 Unique nucleotide interactions in EhRho1

A structure-based multiple sequence alignment of EhRho1 with mammalian Rho GTPases (Figure 4.3A) revealed a potentially activating Ile-45 in the position occupied by a conserved phenylalanine in all other Rho GTPases (Phe-28 in Cdc42). Additionally, the non-conserved Ser-166 and Val-167 of EhRho1 replace residues of Rho GTPases that contact the guanine ring of GDP/GTP [9], suggesting a unique mechanism of nucleotide binding and exchange. The electron density map of the EhRho1·GDP complex shows clearly defined backbone and

side chains surrounding the bound nucleotide (Figure 4.8A), affording a detailed analysis of nucleotide contacts.

The aromatic side chain of Phe-30 of Hs RhoA makes π-orbital interactions with the guanine ring of GDP (Figure 4.8B) [49], and similar contacts are seen in Hs Cdc42 and Hs Rac [49, 50]. Mutation of this phenylalanine to leucine in each Rho family GTPase results in markedly faster basal nucleotide exchange [40] (Figure 4.9) and increased disorder in switch 1, presumably allowing an easier exit route for GDP [49]. In contrast, switch 1 of EhRho1·GDP is well defined in the crystal structure with continuous, strong electron density (Figure 4.8A), despite substitution of a non-aromatic hydrophobic side chain at position 45. The "top" face of the GDP guanine ring forms Van der Waals interactions with the hydrophobic portion of Lys-133 in an identical fashion to Hs RhoA (Figure 4.8). Also conserved is a hydrogen bond network between Asp-135 and the 2-amino and 6-keto groups of the guanine ring. However, the residue triad in positions 165-167, Ser-Ser-Val of EhRho1, differs from the conserved Ser-Ala-Lys/Leu of other Rho GTPases (Figure 4.3A). Since the switch 1 phenylalanine side chain (typically available to form aromatic contacts with the "bottom" face of the guanine ring) is replaced by Ile-45 in EhRho1, we hypothesized that other non-conserved residues (Ser-166 and Val-167) may compensate to control the rate of spontaneous nucleotide exchange. Ser-166 is clearly defined in the electron density and does not engage the nucleotide through its side chain. However, an alternate Ser-166 rotamer could easily bring the non-conserved hydroxyl group within hydrogen bonding distance of the 6-keto or N7 position of the guanine ring, thus restraining the nucleotide. Similarly, the bulky hydrophobic side chain of Val-167 may serve to engage the guanine ring opposite Lys-133 to stabilize nucleotide binding. To address these hypotheses, we sequentially mutated

these EhRho1 residues to the corresponding Hs RhoA side chains and tested their effects on guanine nucleotide exchange.

4.4.7 Non-conserved residues in EhRho1 contribute to a restrained nucleotide exchange rate

The intrinsic rate of nucleotide exchange on G-proteins can be monitored in real time by tracking the increase in fluorescence of a dye-labeled nucleotide upon binding into the nucleotide pocket (reviewed in [51]). The resulting binding curves yield kinetic information about the nucleotide exchange reaction, with the rate-limiting step being release of tightly bound nucleotide. We first examined the basal exchange rates of EhRho1 and HsRhoA to determine whether EhRho1 is constitutively active, given its isoleucine at position 45. EhRho1 was found to exchange nucleotide more quickly than Hs RhoA (and other human Rho GTPases [40]), with an ~4-fold higher k_{obs} (Figure 4.9A). However, EhRho1's rate of exchange is not comparable to the constitutively activating F30L Hs RhoA mutant, which exhibits an ~20-fold faster exchange than wildtype Hs RhoA and saturation within a few minutes. We reasoned that EhRho1 must have additional mode(s) of nucleotide retention, potentially through the non-conserved residues Ser-166 and Val-167 (Figure 4.8A), as previously described. Substitution of both residues for the analogous Hs RhoA residues (S116A/V167K) increased the exchange rate an additional ~2-fold over wildtype EhRho1, implicating these two residues as important for maintaining a controlled exchange rate despite the presence of Ile-45 on switch 1 (Figure 4.8A). However, the EhRho1 S166A/V167K mutant still exchanged only ~8-fold faster than wildtype Hs RhoA, implying

other controlling mechanisms in EhRho1 that are not present in the extremely fast-exchanging Hs RhoA F30L mutant [40].

Finally, we asked whether the additional methyl group of isoleucine-45 relative to leucine might contribute to the controlled nucleotide exchange of EhRho1 compared to Hs RhoA F30L. Mutation of Ile-45 to leucine increased the rate of exchange ~2-fold, comparable to the RhoA-like S166A/V167K double mutation (Figure 4.9B). Interestingly, insertion of a phenylalanine at this position (the conserved residue among other Rho GTPases) drastically dampened the fast EhRho1 exchange rate; EhRho1 I45F exhibited exchange indistinguishable from wildtype Hs RhoA (Figure 4.9B). In conclusion, Ile-45 of EhRho1 is sufficient to confer ~4-fold faster exchange over other known Rho GTPases, without resulting in an uncontrolled, constitutively active G-protein as is seen in Phe->Leu mutations of human Rho GTPases.

4.5   DISCUSSION

The genome of the single-cell protist *E. histolytica* encodes a strikingly large family of expressed Rho GTPases (19 members in *E. histolytica* vs. 20 in humans) (Figure 4.1), given the relative simplicity of a unicellular parasite compared to the diverse array of cell and tissue types in mammals. Mammalian Rho GTPases are known to have a complex choreography of spatio-temporal regulation of multiple family members within a single cell during such processes as cell migration (e.g. [52]). Amoeboid motility is also complex, though poorly understood, and requires membrane detachment from cytoskeletal components (blebbing) and rapid subsequent restructuring of actin [53]. In addition to motility, *E. histolytica* is dependent on a dynamic actin cytoskeleton for multiple pathogenic processes

such as attachment, destruction, and phagocytosis of host cells, chemotaxis, and shedding of host antibodies from the cell membrane [4, 54]. To begin to understand the role *E. histolytica*'s Rho family GTPases in regulating these processes via nucleotide cycling, we have now provided structural snapshots of a representative family member, EhRho1, yielding insights to its mechanisms of nucleotide-dependent activation and effector engagement. Like its mammalian homologs, EhRho1 exhibits a conformational change upon exchanging its bound GDP for a nucleotide bearing a third phosphoryl group, dominated by mobile two switch regions. Since the highly mobile N-terminal portion of switch 2 is uniformly conserved with the related EhRacG (and others), we hypothesize that switch 1 plays a dominant role in dictating nucleotide state-specific binding to EhRho1 effectors, such as EhFormin1. Surprisingly, EhRho1 lacks a signature Rho insert helix, conserved among mammalian Rho GTPases, resulting in a secondary structure pattern more closely akin to H-Ras. Although EhRho1 has highest sequence similarity to Hs RhoA, we and others [21] questioned whether it might be mis-classified as a Rho GTPase.

Our present observations of nucleotide-state selective binding to EhFormin1 and EhRhoGDI, together with stress fiber induction in mammalian fibroblasts, indicate correct labeling of EhRho1 as a *bona fide* Rho GTPase. One may speculate that EhRho1, in lacking the characteristic Rho insert helix yet binding to classical Rho GTPase effectors, resembles an ancestral small G-protein family, and thereby represents an early split of the Rho GTPase subfamily from the greater Ras superfamily [6].

EhFormin1 localizes with actin filaments in *E. histolytica* trophozoites upon exposure to serum factors (19) and is co-enriched with EhRho1 in the uropod, suggesting roles for the EhRho1/EhFormin1 complex in amoebic motility and/or surface receptor capping [48].

EhFormin1 also co-localizes with microtubules during mitosis, and its overexpression leads to delays in cell cycle progression [20]. It will be interesting to determine whether EhRho1 mediates actin and/or microtubule dynamics in response to extracellular cues in *E. histolytica* trophozoites. Like its mammalian counterparts [10], EhRho1 might also exert control over cell division, either through EhFormin1 regulation of microtubule structures, or through other mechanisms such as transcriptional regulation.

Another unique feature of EhRho1 is its relatively fast rate of spontaneous nucleotide exchange in the absence of an exchange factor (GEF). Future experiments will determine whether exchange on EhRho1 is modulated by EhRhoGDI. The moderately fast, spontaneous exchange is mediated by three non-conserved residues surrounding the guanine moiety binding site. The switch 1 residue Ile-45, reminiscent of the constitutively active Hs Cdc42 F28L mutant [41], is sufficient to endow EhRho1 with an exchange rate intermediate between a Rho-like phenylalanine substitution and a rapid exchanging leucine substitution. In addition, Ser-166 and Val-167 on the β6-α5 loop contribute to nucleotide binding stability to a greater extent than the analogous Ala-161 and Lys-162 in Hs RhoA. Together, these data suggest that EhRho1 has evolved a relatively high rate of basal activity through the loss of a highly conserved switch one phenylalanine residue. However, uncontrolled exchange and constitutive activity are avoided, in part through compensatory nucleotide-interacting residues. Interestingly, neither the I45L mutation nor the S166A/V167K double mutation recapitulated the uncontrolled exchange seen in the Hs RhoA F30L mutation [41]. In the crystal structures of EhRho1, the nucleotide binding pocket is nearly identical to that of Hs RhoA except for the residues highlighted in this study. We hypothesize that EhRho1 protein dynamics in solution, rather than amino acid sequence, may contribute to a more moderate

nucleotide exchange. NMR studies of Cdc42 have identified increased switch 1 mobility upon introduction of the F28L mutation [55]. Since the switch 1 region of EhRho1 is well ordered in the crystal structures, its conformational restrictions may contribute to relative nucleotide binding stability.
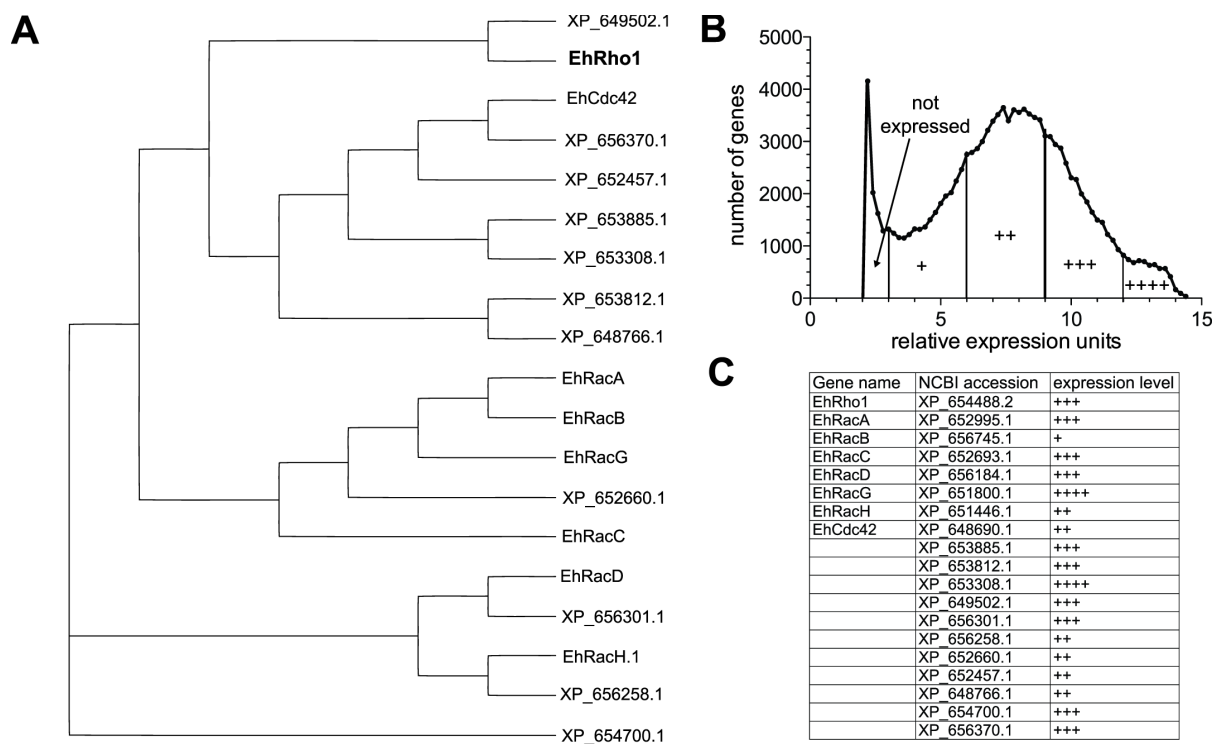
Figure 4.1. Identification of 19 expressed Rho family GTPases in *E. histolytica*. (A) Nineteen predicted Rho family GTPase open-reading frames encoded within the genome of *E. histolytica* were aligned to produce a dendrogram. Translations of mRNAs whose expression was not detected by microarrays were excluded. **(B)** A frequency distribution of relative expression units, derived from available microarray data on *E. histolytica* HM1:IMSS trophozoites [24] was used to determine the relative transcription levels of each Rho GTPase gene. **(C)** Named genes previously reported in the literature and 11 additional transcribed Rho GTPases are listed with NCBI accession numbers and their relative expression levels.
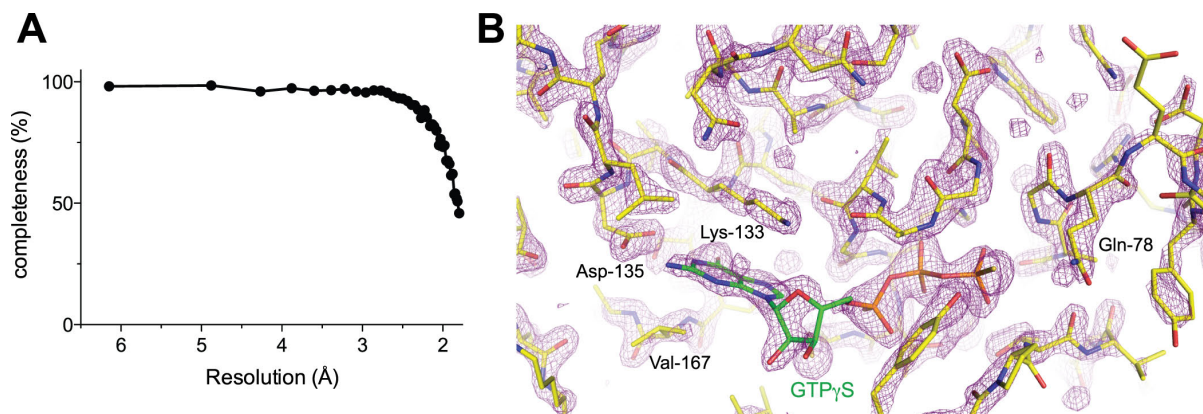
Figure 4.2. Electron density map derived from EhRho1·GTPγS diffraction data. (A)
Diffraction data completeness was plotted against the high resolution limit of each shell. Data
completeness is near 100% at lower resolutions, with a sharp decrease at ~2.5 Å. Data from
2.5 to 1.80 Å were included in the refinement because in each shell, I/σI was greater than 2.5
and R_{sym} was less than 0.28. Inclusion of the high-resolution data improved the electron
density map **(B)** shown for the bound GTPγS and surrounding EhRho1 stick model. The map
was contoured to σ = 2.3.

Figure 4.3. Sequence similarity among Rho and Ras family GTPases from *Entamoeba histolytica* and humans. (A) A multiple sequence alignment highlights conserved switch regions and nucleotide-interacting residues (closed circles) derived from the crystal structures of EhRho1 reported here. The characteristic Rho insert helix region is indicated in gray, and arrowheads mark residues mutated in this study. **(B)** Purified Rho GTPases and the tandem Rho GTPase binding (GBD) and formin homology 3 (FH3) domains of EhFormin1 (residues 69-445) were resolved by SDS-PAGE and stained with Coomassie blue. EhFormin1 GBD-FH3 co-purified with a likely C-terminally truncated form (asterisk).

A *E.h.* Rho1 • GDP or GTPγS

B *H.s.* RhoA • GTPγS

C *H.s.* H-Ras • GppNHp
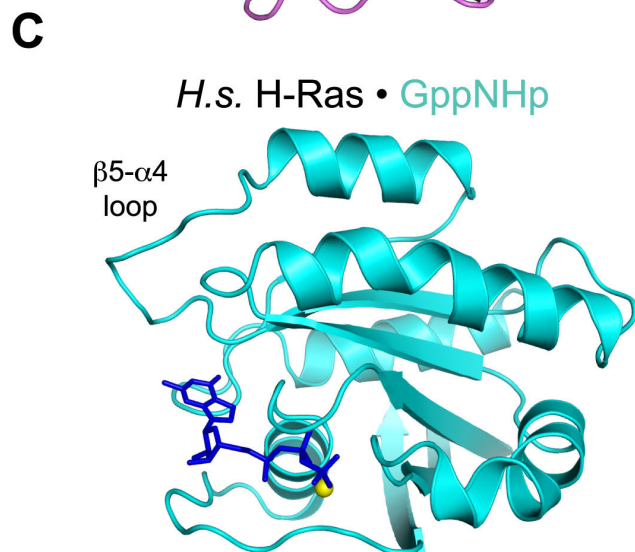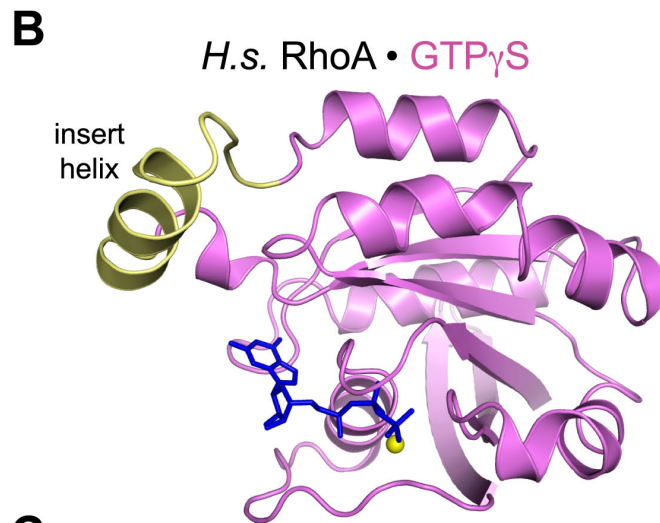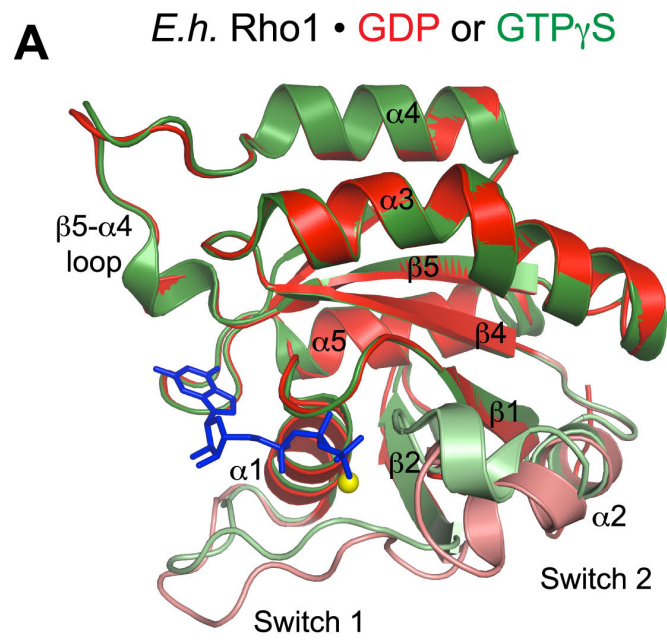
Figure 4.4. The structural models of EhRho1 in two nucleotide states reveal a conserved mechanism of nucleotide-dependent activation with Rho- and Ras-like characteristics. **(A)** Models of the crystal structures of EhRho1 bound to magnesium and GDP (*red*) or the non-hydrolyzable GTP analog GTPγS (*green*) are superimposed. See Table 4.1 for crystallographic data collection and refinement statistics. GTPγS is modeled as blue sticks, and magnesium as a yellow sphere. The well conserved switch regions, shown in light shades, account for the majority of nucleotide state-dependent conformational shifts. EhRho1 is seen to lack the signature Rho insert helix found on all members of the mammalian Rho GTPase family, represented by human RhoA **(B),** with the insert in yellow (PDB accession 1A2B). Instead, EhRho1 shares a non-helical β5-α4 loop with Ras GTPases, represented by human H-Ras **(C)** (PDB accession 5P21).
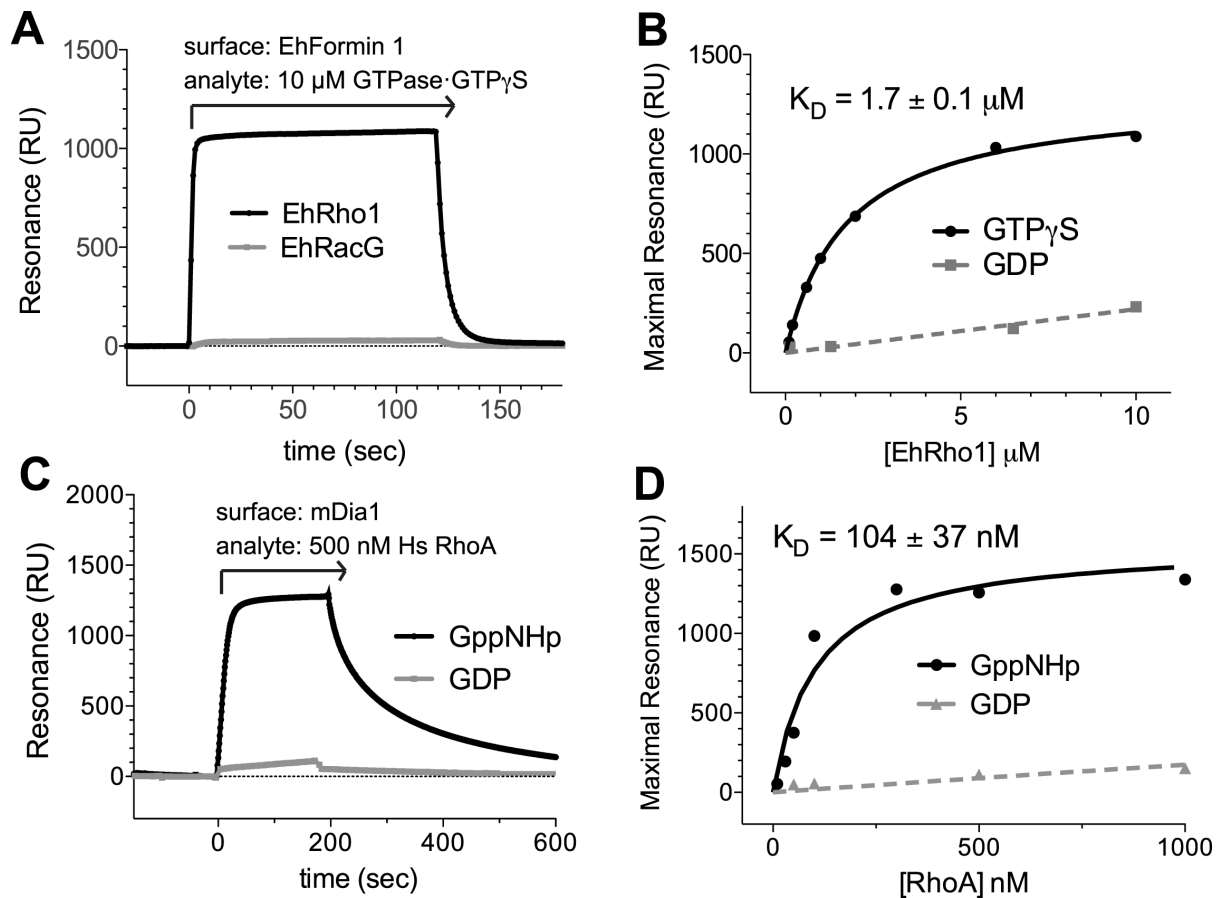
Figure 4.5. EhRho1 interacts with EhFormin1, a homolog of the human Rho effector mDia. (A) The GBD-FH3 tandem of EhFormin1, also referred to as EhDia, was observed to bind EhRho1·GTPγS selectively (to the exclusion of EhRacG·GTPγS), as determined by surface plasmon resonance. **(B)** Injection of increasing amounts of GTPase over immobilized EhFormin1 revealed specific binding to the activated form of EhRho1 with low micromolar affinity. **(C, D)** The homologous interaction between Hs RhoA and mouse mDia1 GBD-FH3 tandem was also highly nucleotide state-selective, with high nanomolar affinity.
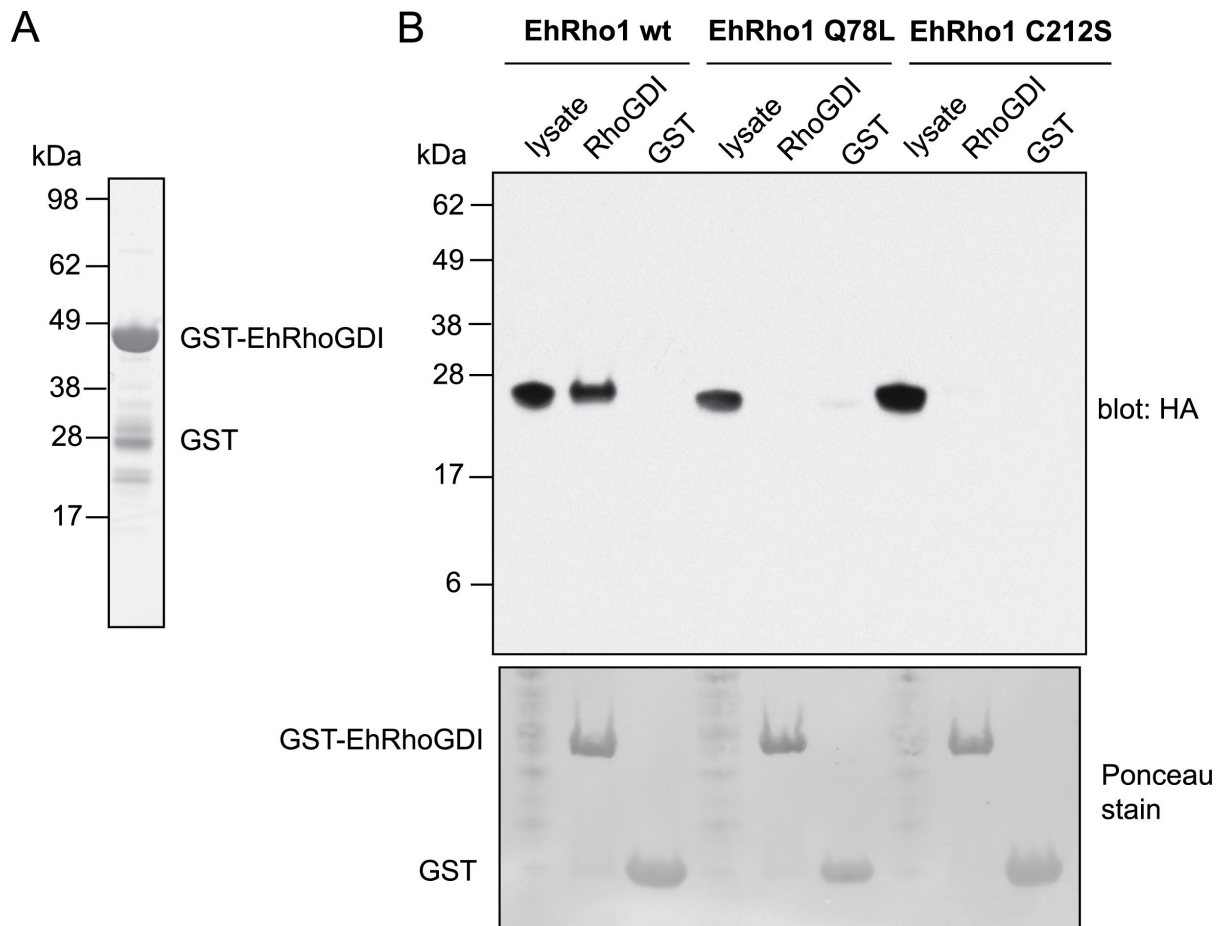
Figure 4.6. EhRho1 interaction with EhRhoGDI is favored by the inactive conformation and prenylation at the CaaX motif. (A) EhRhoGDI was cloned from the *E. histolytica* genome and purified as a GST fusion from *E. coli*. Purified GST-EhRhoGDI was resolved by SDS-PAGE and stained with Coomassie blue. **(B)** HA epitope-taged EhRho1 expressed in HEK293T cells co-precipitated with GST-RhoGDI immobilized on glutathione agarose, but not GST alone, as indicated by immunoblotting with anti-HA antibody. The EhRho1/EhRhoGDI interaction was disrupted by mutation of the key glutamine for GTP hydrolysis (Q78L), rendering EhRho1 constitutively active. In addition, mutation of the CaaX motif cysteine that is geranylgeranylated in mammalian Rho GTPases (C212S in EhRho1) abolished binding to EhRhoGDI.
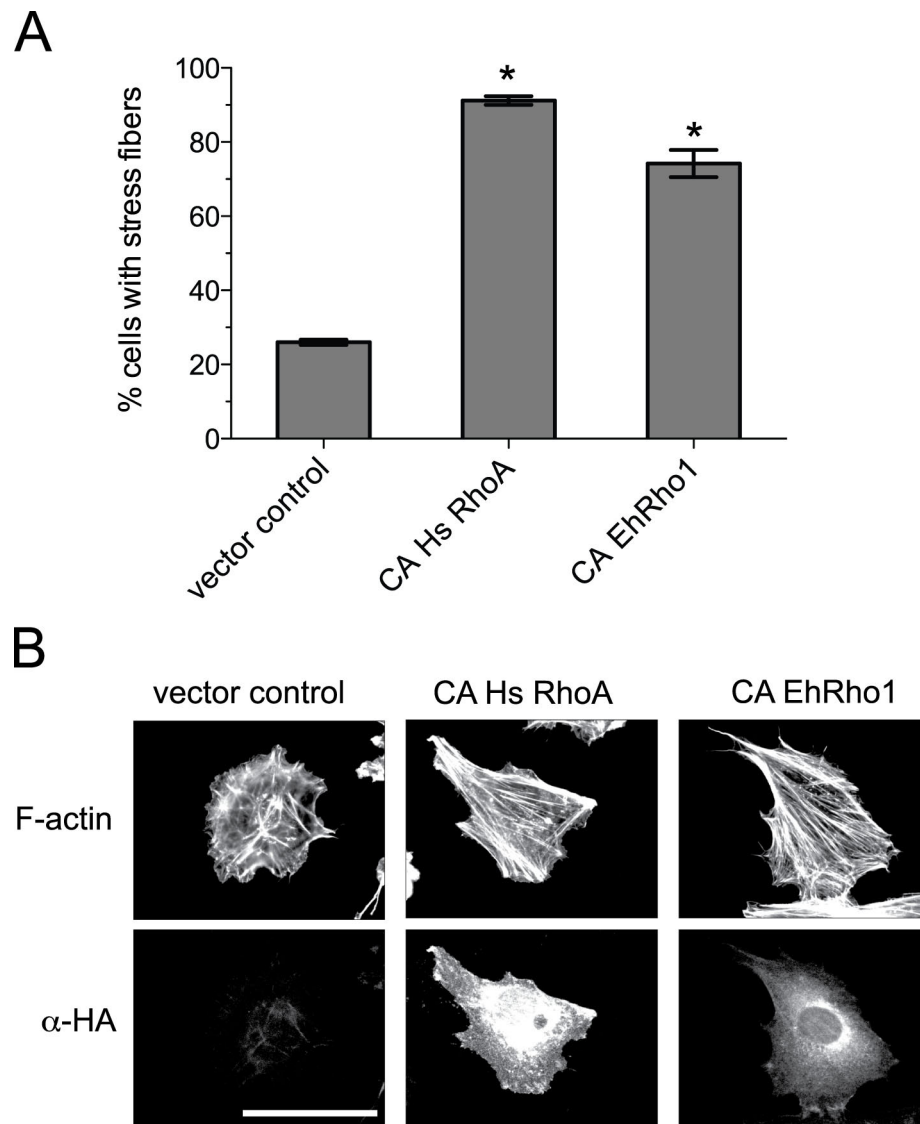
Figure 4.7. EhRho1 induces stress fiber formation in mammalian fibroblasts. Rat-2 fibroblasts were transfected with either pcDNA3.1 (vector control), HA epitope-tagged constitutively active (CA) human RhoA(G14V) or EhRho1(Q78L). Following 24 hours plating on fibronectin-coated coverslips, the cells were fixed and stained with phalloidin and anti-HA antibody. Cells were scored for the presence of actin stress fibers by a blinded observer. **(A)** Fibroblasts expressing either constitutively active Hs RhoA or EhRho1 had a significantly greater percentage of cells with actin stress fibers compared to vector control. Error bars represent S.E.M. for duplicate experiments and * indicates $p < 0.01$ compared to vector control. **(B)** Representative phalloidin staining (upper panels) and anti-HA antibody immunofluorescence (lower panels) images indicate more prominent stress fibers in constitutively active Hs RhoA- and EhRho1-expressing cells compared to vector control. The scale bar represents 50 μm.
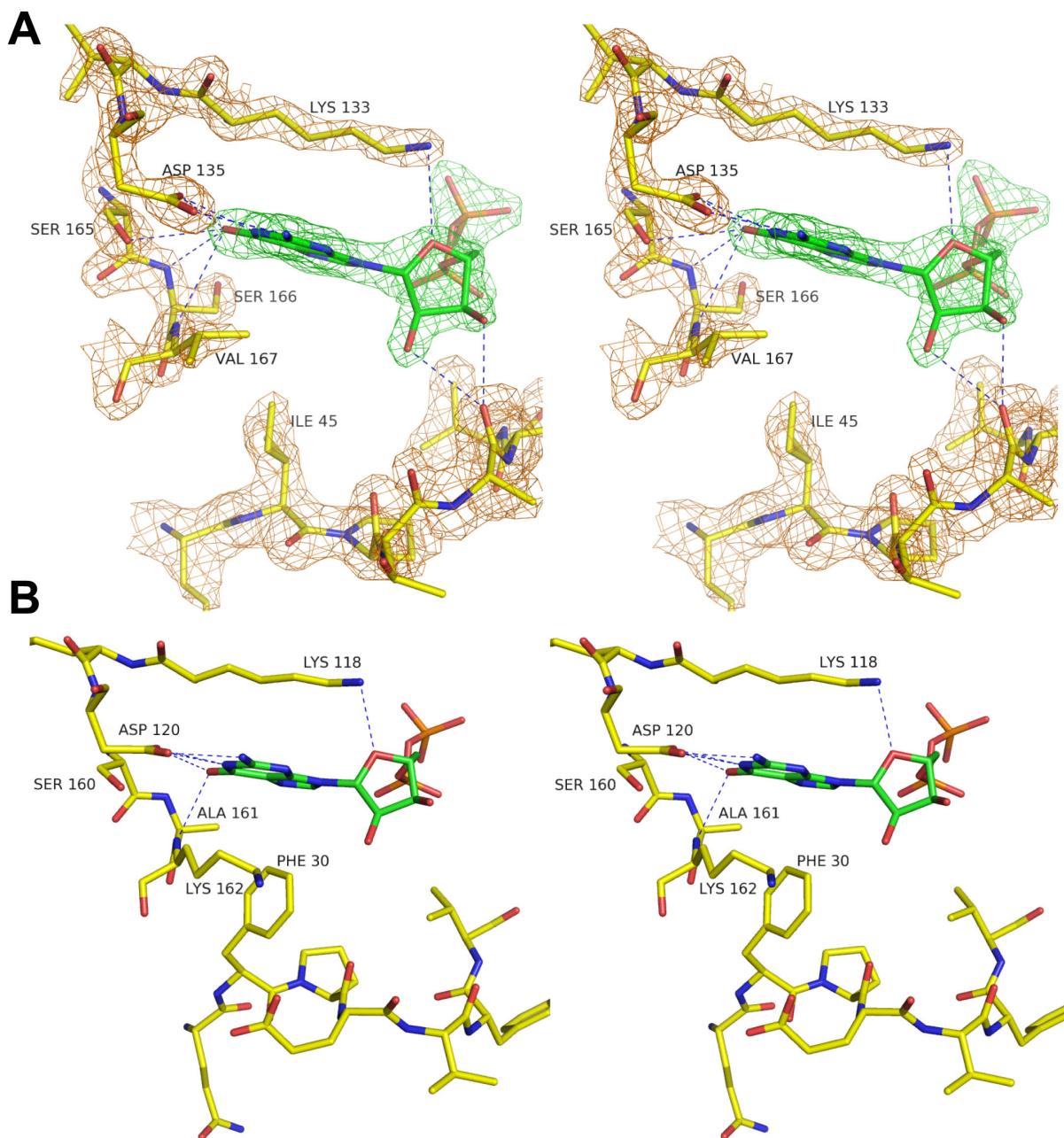
Figure 4.8. Unique guanine nucleotide binding pocket residues of EhRho1. (A) A stereo view of the nucleotide binding pocket from EhRho1·GDP illustrates a divergent set of residues interacting with the guanine ring of GDP (*green*) when compared to the corresponding region of human RhoA **(B)** (PDB accession 1FTN). Position 45 on switch 1 is occupied by an isoleucine (Ile-45) in EhRho1, corresponding to Phe-30 in Hs RhoA; the latter residue makes π-orbital interactions with the guanine ring of GDP [55]. The conserved triad of Ser-160, Ala-161, and Lys-162 in Hs RhoA is replaced by Ser-165, Ser-166, and Val-157 in EhRho1. The electron density map was contoured to σ = 3.0.

181
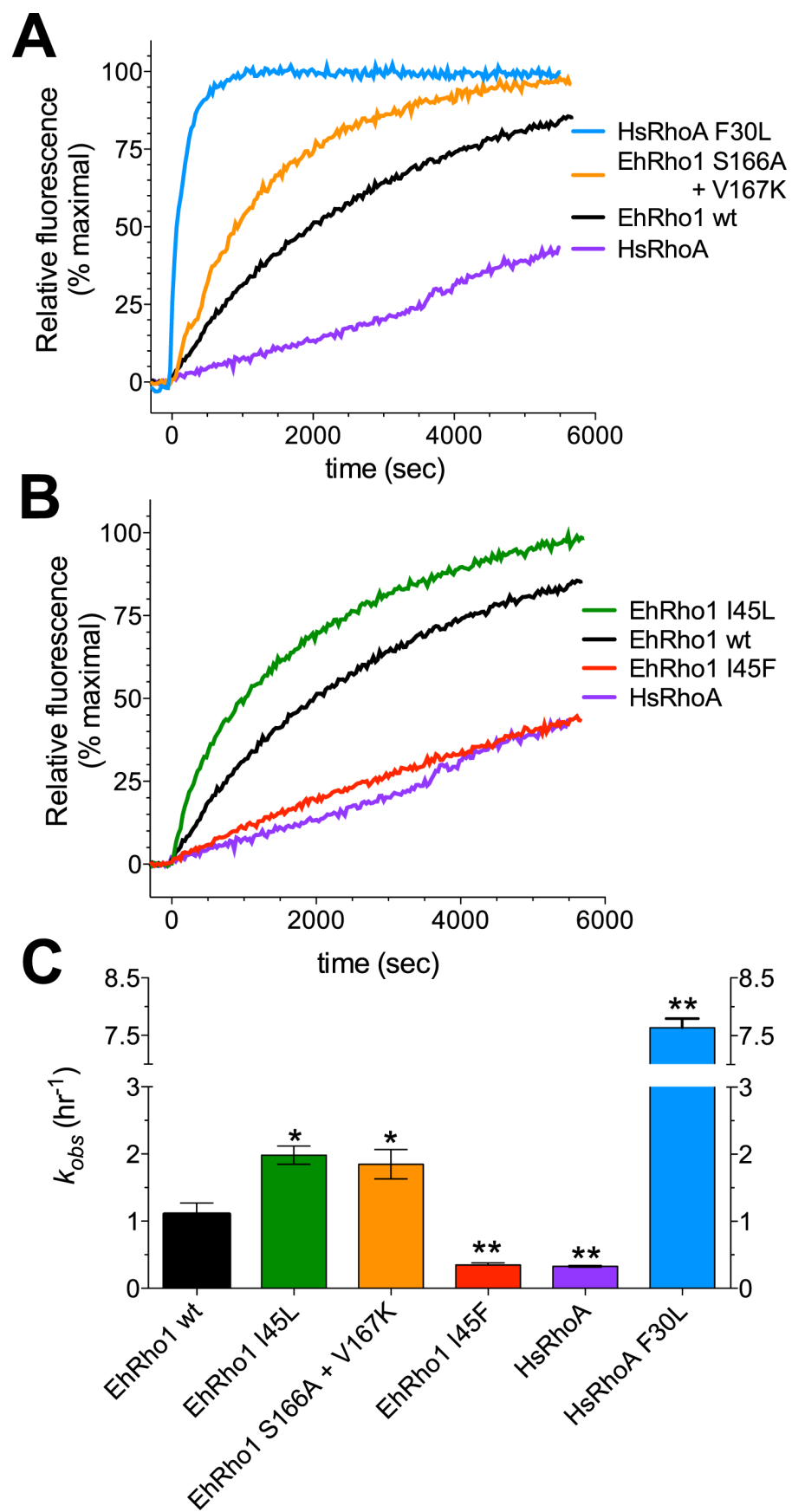
Figure 4.9. Non-conserved nucleotide binding pocket residues moderate an otherwise fast rate of nucleotide exchange on EhRho1. (A) Real-time measurement of Rho GTPase binding to fluorescent BODIPY-GDP demonstrated a ~4-fold higher spontaneous nucleotide exchange rate on EhRho1 compared to Hs RhoA, representative of the Rho family GTPases in mammals. The Hs RhoA F30L mutant exchanged nucleotide at a ~20-fold higher rate than wild type Hs RhoA. Substitution of the non-conserved guanine ring-interacting residues Ser-166 and Val-167 of EhRho1 for the corresponding amino acids from Hs RhoA (Ala and Lys) results in a 2-fold higher rate of exchange compared to wildtype EhRho1, implicating these residues as important for limiting basal nucleotide exchange in wildtype EhRho1. **(B)** Ile-45 in EhRho1 corresponds to the highly conserved Phe-30 of Hs RhoA, of which mutation to leucine results in very rapid exchange and constitutive activity [40]. Interestingly, the Hs RhoA-like substitution on EhRho1, Ile-45 to phenylalanine, produces a slow exchange rate, indistinguishable from wildtype Hs RhoA. Mutation of Ile-45 to leucine produces a 2-fold increase in the rate of exchange. **(C)** The average exchange rate constants, k_{obs}, derived from exponential curve fitting were plotted with S.E.M. error bars for quadruplicate experiments. Asterisks indicate statistically significant differences from EhRho1 wildtype (*, $p < 0.05$; **, $p < 0.01$).

Table 4.1. Data collection and refinement statistics for EhRho1

	EhRho1·GDP	EhRho1·GTPγS
PDB accession code	3REF	3REG
Data Collection		
Space group	$P2_12_12_1$	P1
Unit cell: a, b, c (Å)	50.3, 54.5, 132.3	36.5, 39.5, 63.6
α, β, γ (°)	90, 90, 90	81.8, 80.8, 65.4
Resolution (Å)	30.0 – 1.95 (1.98 – 1.95)[a]	40.0 – 1.80 (1.82 – 1.80)
R_{sym}	0.047 (0.187)	0.043 (0.278)
I/sI	21.9 (2.7)	16.1 (2.8)
Unique reflections	27,267 (633)	23,929 (435)
Completeness (%)	99.6 (99.4)	81.6 (45.9) [b]
Redundancy	2.5 (1.7)	2.0 (1.4)
Wilson B-factor ($Å^2$)	23.0	21.6
Refinement		
Resolution (Å)	28.3 – 1.95 (2.01 - 1.95)	33.0 – 1.80 (1.87 – 1.80)
No. of reflections (work/free)	27,160/1366 (2660/123)	23,884/1214 (1558/86)
Cut-off (s)	0.12	0
R_{work} / R_{free} (%)	15.7 / 19.7 (19.0 / 25.8)	17.2 / 22.2 (21.5 / 29.1)
No. of atoms		
Protein	2468	2658
GDP	56	64
Ions	17	2
Water	213	229
B-factor ($Å^2$)		
Protein	33.8	31.4
GDP	20.4	25.4
Ions	60.3	29.2
Water	36.4	38.8
R.m.s. deviations		
Bond Lengths (Å)	0.009	0.009
Bond Angles (°)	1.26	1.20
Ramachandran		
Favored (%)	95.7	98.5
Generally Allowed (%)	4.3	1.5
Disallowed (%)	0	0
Missing residues	A: 1-14, 188-194	A: 1-20, 192-194
	B: 1-14	B: 1-20, 189-194

[a]Values in parentheses represent the highest resolution shell
Diffraction data were generated from single crystals of EhRho1 in each nucleotide state.
[b]Despite 81.6% completeness of the EhRho1·GTPγS data set, the resulting electron density was of high quality without systematic deficits (see Figure 4.2).

4.6   REFERENCES

1. WHO, *WHO/PAHO/UNESCO report. A consultation with experts on amoebiasis. Mexico City, Mexico 28-29 January, 1997.* Epidemiol Bull, 1997. **18**(1): p. 13-4.

2. Haque, R., C.D. Huston, M. Hughes, E. Houpt, and W.A. Petri, Jr., *Amebiasis.* N Engl J Med, 2003. **348**(16): p. 1565-73.

3. Ravdin, J.I., *Amoebiasis.* Vol. 2. 2000, London: Imperial College Press.

4. Guillen, N., *Role of signalling and cytoskeletal rearrangements in the pathogenesis of Entamoeba histolytica.* Trends Microbiol, 1996. **4**(5): p. 191-7.

5. Voigt, H. and N. Guillen, *New insights into the role of the cytoskeleton in phagocytosis of Entamoeba histolytica.* Cell Microbiol, 1999. **1**(3): p. 195-203.

6. Wennerberg, K., K.L. Rossman, and C.J. Der, *The Ras superfamily at a glance.* J Cell Sci, 2005. **118**(Pt 5): p. 843-6.

7. Ridley, A.J. and A. Hall, *The small GTP-binding protein rho regulates the assembly of focal adhesions and actin stress fibers in response to growth factors.* Cell, 1992. **70**(3): p. 389-99.

8. Hoffman, G.R. and R.A. Cerione, *Flipping the switch: the structural basis for signaling through the CRIB motif.* Cell, 2000. **102**(4): p. 403-6.

9. Wei, Y., et al., *Crystal structure of RhoA-GDP and its functional implications.* Nat Struct Biol, 1997. **4**(9): p. 699-703.

10. Wennerberg, K. and C.J. Der, *Rho-family GTPases: it's not only Rac and Rho (and I like it).* J Cell Sci, 2004. **117**(Pt 8): p. 1301-12.

11. Garcia-Mata, R., E. Boulter, and K. Burridge, *The 'invisible hand': regulation of RHO GTPases by RHOGDIs.* Nat Rev Mol Cell Biol, 2011. **12**(8): p. 493-504.

12. Narumiya, S., M. Tanji, and T. Ishizaki, *Rho signaling, ROCK and mDia1, in transformation, metastasis and invasion.* Cancer Metastasis Rev, 2009. **28**(1-2): p. 65-76.

13. Lohia, A. and J. Samuelson, *Molecular cloning of a rho family gene of Entamoeba histolytica.* Mol Biochem Parasitol, 1993. **58**(1): p. 177-80.

14. Lohia, A. and J. Samuelson, *Heterogeneity of Entamoeba histolytica rac genes encoding p21rac homologues.* Gene, 1996. **173**(2): p. 205-8.

15. Ghosh, S.K. and J. Samuelson, *Involvement of p21racA, phosphoinositide 3-kinase, and vacuolar ATPase in phagocytosis of bacteria and erythrocytes by Entamoeba histolytica:*

suggestive evidence for coincidental evolution of amebic invasiveness. Infect Immun, 1997. **65**(10): p. 4243-9.

16. Guillen, N., P. Boquet, and P. Sansonetti, *The small GTP-binding protein RacG regulates uroid formation in the protozoan parasite Entamoeba histolytica.* J Cell Sci, 1998. **111 (Pt 12)**: p. 1729-39.

17. Aguilar-Rojas, A., et al., *Entamoeba histolytica: inhibition of cellular functions by overexpression of EhGEF1, a novel Rho/Rac guanine nucleotide exchange factor.* Exp Parasitol, 2005. **109**(3): p. 150-62.

18. Arias-Romero, L.E., et al., *EhGEF3, a Novel Dbl Family Member, Regulates EhRacA Activation During Chemotaxis and Capping in Entamoeba histolytica.* Cell Motility and the Cytoskeleton, 2007. **64**: p. 390-404.

19. Labruyere, E., C. Zimmer, V. Galy, J.-C. Olivo-Marin, and N. Guillen, *EhPAK, a member of the p21-activated kinase family, is involved in the control of Entamoeba histolytica migration and phagocytosis.* Journal of Cell Science, 2003. **116**: p. 61-71.

20. Majumder, S. and A. Lohia, *Entamoeba histolytica encodes unique formins, a subset of which regulates DNA content and cell division.* Infect Immun, 2008. **76**(6): p. 2368-78.

21. Godbold, G.D., K.D. Corbett, and B.J. Mann, *A Rho-like small GTPase of Entamoeba histolytica contains an unusual amino acid residue in a conserved GDP-stabilization region and is not a substrate for C3 exoenzyme.* Exp Parasitol, 2002. **101**(2-3): p. 107-10.

22. Loftus, B., et al., *The genome of the protist parasite Entamoeba histolytica.* Nature, 2005. **433**(7028): p. 865-8.

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

24. Gilchrist, C.A., et al., *Regulation of Virulence of Entamoeba histolytica by the URE3-BP Transcription Factor.* MBio, 2010. **1**(1).

25. Larkin, M.A., et al., *Clustal W and Clustal X version 2.0.* Bioinformatics, 2007. **23**(21): p. 2947-8.

26. Perriere, G. and M. Gouy, *WWW-query: an on-line retrieval system for biological sequence banks.* Biochimie, 1996. **78**(5): p. 364-9.

27. Stols, L., et al., *A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site.* Protein Expr Purif, 2002. **25**(1): p. 8-15.

28. Ho, S.N., H.D. Hunt, R.M. Horton, J.K. Pullen, and L.R. Pease, *Site-directed mutagenesis by overlap extension using the polymerase chain reaction.* Gene, 1989. **77**(1): p. 51-9.

29. Otwinowski, Z.a.W.M., *Processing of X-ray Diffraction Data Collected in Oscillation Mode*, in *Methods in Enzymology*. 1997, Academic Press: New York. p. 307-326.

30. Stein, N., *CHAINSAW: a program for mutating pdb files used as templates in molecular replacement.* J. Appl. Cryst., 2008. **41**: p. 641-643.

31. McCoy, A.J., et al., *Phaser crystallographic software.* J Appl Crystallogr, 2007. **40**(Pt 4): p. 658-674.

32. Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.

33. Emsley, P., B. Lohkamp, W.G. Scott, and K. Cowtan, *Features and development of Coot.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.

34. Painter, J. and E.A. Merritt, *Optimal description of a protein structure in terms of multiple groups undergoing TLS motion.* Acta Crystallogr D Biol Crystallogr, 2006. **62**(Pt 4): p. 439-50.

35. Ihara, K., et al., *Crystal structure of human RhoA in a dominantly active form complexed with a GTP analogue.* J Biol Chem, 1998. **273**(16): p. 9656-66.

36. Kimple, A.J., R.E. Muller, D.P. Siderovski, and F.S. Willard, *A capture coupling method for the covalent immobilization of hexahistidine tagged proteins for surface plasmon resonance.* Methods Mol Biol, 2010. **627**: p. 91-100.

37. Kimple, R.J., et al., *Guanine nucleotide dissociation inhibitor activity of the triple GoLoco motif protein G18: alanine-to-aspartate mutation restores function to an inactive second GoLoco motif.* Biochem J, 2004. **378**(Pt 3): p. 801-8.

38. Willard, M.D., et al., *Selective role for RGS12 as a Ras/Raf/MEK scaffold in nerve growth factor-mediated differentiation.* EMBO J, 2007. **26**(8): p. 2029-40.

39. Raines, K.W., M.G. Bonini, and S.L. Campbell, *Nitric oxide cell signaling: S-nitrosation of Ras superfamily GTPases.* Cardiovasc Res, 2007. **75**(2): p. 229-39.

40. Lin, R., R.A. Cerione, and D. Manor, *Specific contributions of the small GTPases Rho, Rac, and Cdc42 to Dbl transformation.* J Biol Chem, 1999. **274**(33): p. 23633-41.

41. Lin, R., S. Bagrodia, R. Cerione, and D. Manor, *A novel Cdc42Hs mutant induces cellular transformation.* Curr Biol, 1997. **7**(10): p. 794-7.

42. Nassar, N., G.R. Hoffman, D. Manor, J.C. Clardy, and R.A. Cerione, *Structures of Cdc42 bound to the active and catalytically compromised forms of Cdc42GAP.* Nat Struct Biol, 1998. **5**(12): p. 1047-52.

43. Pai, E.F., et al., *Refined crystal structure of the triphosphate conformation of H-ras p21 at 1.35 A resolution: implications for the mechanism of GTP hydrolysis.* EMBO J, 1990. **9**(8): p. 2351-9.

44. Scheffzek, K., I. Stephan, O.N. Jensen, D. Illenberger, and P. Gierschik, *The Rac-RhoGDI complex and the structural basis for the regulation of Rho proteins by RhoGDI.* Nat Struct Biol, 2000. **7**(2): p. 122-6.

45. Abdul-Manan, N., et al., *Structure of Cdc42 in complex with the GTPase-binding domain of the 'Wiskott-Aldrich syndrome' protein.* Nature, 1999. **399**(6734): p. 379-83.

46. Walker, S.J., W.J. Wu, R.A. Cerione, and H.A. Brown, *Activation of phospholipase D1 by Cdc42 requires the Rho insert region.* J Biol Chem, 2000. **275**(21): p. 15665-8.

47. Freeman, J.L., A. Abo, and J.D. Lambeth, *Rac "insert region" is a novel effector region that is implicated in the activation of NADPH oxidase, but not PAK65.* J Biol Chem, 1996. **271**(33): p. 19794-801.

48. Marquay Markiewicz, J., et al., *A proteomic and cellular analysis of uropods in the pathogen Entamoeba histolytica.* PLoS Negl Trop Dis, 2011. **5**(4): p. e1002.

49. Adams, P.D. and R.E. Oswald, *Solution structure of an oncogenic mutant of Cdc42Hs.* Biochemistry, 2006. **45**(8): p. 2577-83.

50. Lapouge, K., et al., *Structure of the TPR domain of p67phox in complex with Rac.GTP.* Mol Cell, 2000. **6**(4): p. 899-907.

51. Rojas, R.J., R.J. Kimple, K.L. Rossman, D.P. Siderovski, and J. Sondek, *Established and emerging fluorescence-based assays for G-protein function: Ras-superfamily GTPases.* Comb Chem High Throughput Screen, 2003. **6**(4): p. 409-18.

52. Machacek, M., et al., *Coordination of Rho GTPase activities during cell protrusion.* Nature, 2009. **461**(7260): p. 99-103.

53. Maugis, B., et al., *Dynamic instability of the intracellular pressure drives bleb-based motility.* J Cell Sci, 2010. **123**(Pt 22): p. 3884-92.

54. Meza, I., P. Talamas-Rohana, and M.A. Vargas, *The cytoskeleton of Entamoeba histolytica: structure, function, and regulation by signaling pathways.* Arch Med Res, 2006. **37**(2): p. 234-43.

55. Adams, P.D., A.P. Loh, and R.E. Oswald, *Backbone dynamics of an oncogenic mutant of Cdc42Hs shows increased flexibility at the nucleotide-binding site.* Biochemistry, 2004. **43**(31): p. 9968-77.

CHAPTER 5

Entamoeba histolytica Rho1 REGULATES ACTIN POLYMERIZATION THROUGH A DIVERGENT, DIAPHANOUS-RELATED FORMIN[1]

5.1   OVERVIEW

Entamoeba histolytica requires a dynamic actin cytoskeleton for intestinal and systemic pathogenicity. Diaphanous-related formins represent an important family of actin regulators that are activated by Rho GTPases. The *E. histolytica* genome encodes a large family of Rho GTPases and three diaphanous-related formins, of which EhFormin1 is known to regulate mitosis and cytokinesis in trophozoites. We demonstrate that EhFormin1 modulates actin polymerization through its formin homology 2 (FH2) domain. Despite a highly divergent diaphanous autoinhibitory domain, EhFormin1 is autoinhibited by an N- and C-terminal intramolecular interaction, but activated upon binding of EhRho1 to the N-terminal domain tandem. A crystal structure of the EhRho1·GTPγS/EhFormin1 complex illustrates an EhFormin1 conformation that diverges from mammalian mDia1 and lacks a secondary interaction with a Rho insert helix. The structural model also highlights residues required for specific recognition of the EhRho1 GTPase and suggests that the molecular mechanisms of EhFormin1 autoinhibition and activation differ from mammalian homologs.

---

[1] Bosch, D.E., Yang, B., and Siderovski, D.P. (2012) *Entamoeba histolytica* Rho1 regulates actin polymerization through a divergent, diaphanous-related formin. *Biochemistry*. 51(**44**):8791-801.

5.2   INTRODUCTION

Entamoeba histolytica is the causative agent of amoebic colitis and systemic amoebiasis [1]. Infection by the parasite is spread endemically among poor populations of developing countries, although outbreaks among travelers and susceptible populations occur in the United States [1]. Water-borne *E. histolytica* cysts cycle to the trophozoite form in the human host, in some cases leading to destruction of the intestinal mucosa (amoebic colitis). If untreated, trophozoites may enter the blood stream, leading to systemic amoebiasis characterized by liver, lung, and brain abscesses [2]. Many cellular processes critical to *E. histolytica* pathogenesis, such as chemotaxis, adherence to intestinal epithelium, cell killing, phagocytosis, and penetration of the mucosa are dependent on a highly dynamic actin cytoskeleton [3-5]. *E. histolytica* expresses a relatively large number of Rho family small GTPases [6, 7], conserved signaling molecules that are vital to coordination of actin cytoskeletal rearrangements [8]. Rho GTPases undergo a conformational change dominated by switch regions upon exchange of GTP for GDP, allowing engagement of specific downstream effectors [9]. Overexpression of constitutively active versions of the Rho family GTPases EhRacA or EhRacG in *E. histolytica* trophozoites impairs pathogenic processes such as phagocytosis and surface receptor capping [10, 11]. However, signaling mechanisms by which the actin cytoskeleton of *E. histolytica* is regulated have not been elucidated in molecular detail.

Formins constitute a major class of proteins that directly regulate actin filament formation and thus cellular morphology, adhesion, and motility [12]. Formins promote nucleation and polymerization of unbranched actin filaments [13]. The highly conserved formin homology 2 domain (FH2) forms a head-to-tail dimer that binds to the barbed ends of

actin filaments, catalyzing assembly through a processive capping mechanism [12, 14, 15]. The FH2 domain is commonly preceded by an unstructured, proline-rich formin homology 1 domain (FH1) that engages profilin/actin complexes, thus recruiting G-actin monomers for incorporation into a growing filament [12]. Members of the Diaphanous-related formin (DRF) subfamily also possess an N-terminal Rho GTPase binding domain (GBD) and a formin homology 3 domain (FH3) that, in turn, is composed of an Armadillo repeat-containing Diaphanous inhibitory domain (DID) and a dimerization domain [13]. C-terminal to the FH2 domain of DRFs is a Diaphanous autoinhibitory domain (DAD) that forms intramolecular interactions with the DID, maintaining the formin in an inactive state, as best characterized structurally for the DRF mDia1 [16, 17]. This autoinhibited conformation is released upon binding of specific Rho family GTPases to the GBD-FH3 domain tandem, likely due to active DAD displacement from its DID binding site by Rho-induced contingent folding of the GBD and by the Rho GTPase itself [18, 19]. Although mDia1 primarily engages one of its activating GTPases, RhoC, through the switch regions, the last Armadillo repeat of the DID also weakly contacts the signature Rho insert helix of RhoC [20]. Mutation of residues at this secondary interface leads to reduced RhoC/mDia1 affinity, but it is unclear whether the interaction is important for formin activation *per se* [14].

   E. histolytica possesses a family of eight formins, among which EhFormin1-3 are Diaphanous-related [21]. EhFormin1 and -2 are expressed in trophozoites and associated with pseudopodia, pinocytic and phagocytic vesicles, and F-actin in response to serum. Both formins are also co-localized with the microtubular assembly during mitosis [21]. Overexpression of EhFormin1 increases the number of binucleated cells and nuclear DNA content, suggesting roles for EhFormin1 in mitosis and cytokinesis [21]. A recent proteomic

characterization of *E. histolytica* cysts indicated that EhFormin1 is expressed during both the encysted and trophozoite life cycle stages [22]. We recently showed [7] that the GBD-FH3 domain tandem of EhFormin1 binds EhRho1 in a nucleotide state-dependent fashion, which is typical of Rho GTPase/effector interactions. Furthermore, expression of constitutively active EhRho1 in fibroblasts induced stress fiber formation [7], suggesting that EhRho1 might regulate actin filament formation in *E. histolytica* trophozoites through EhFormin1 or other effectors. Crystal structures of EhRho1 in different nucleotide states highlighted a lack of the signature Rho insert helix [7], suggesting that the EhRho1/EhFormin1 interaction might differ from that of RhoC/mDia1 [20].

In the current study, we demonstrate that EhFormin1 regulates actin polymerization through its FH2 domain. Despite considerable sequence divergence of the DAD motif and the DID motif surface expected to bind the DAD, EhFormin1 is autoinhibited by interaction between its N- and C-terminal domains. As in the case of mDia1, highly selective binding of EhRho1 to the GBD-FH3 tandem is sufficient to activate EhFormin1. Finally, a crystal structure of the EhRho1·GTPγS/EhFormin1 complex reveals a primary interface between EhRho1 and the EhFormin1 GBD with similarities to that of RhoC/mDia1. However, the absence of a Rho insert helix within EhRho1 and a large conformational difference in the DID of EhFormin1 compared to mDia1 illustrate the lack of a secondary EhRho1/EhFormin1 binding site, in contrast with mammalian homologs.

5.3   EXPERIMENTAL PROCEDURES

5.3.1 Protein purification

EhRho1, EhRacC, EhRacD, and EhRacD were cloned from *E. histolytica* genomic DNA by PCR amplification as hexahistidine-tagged open-reading frame fusions, expressed in BL21(DE3) *E. coli*, purified by nickel affinity and gel filtration chromatography, and loaded with GTPγS as described previously for EhRho1 [7]. EhFormin1 (UniProt C4M622) was cloned from genomic DNA, and fragments were PCR amplified and subcloned with a tobacco etch virus (TEV) protease-cleavable hexahistidine tag using ligation independent cloning [23]. Mutagenesis was performed using the two PCR method [24]. All EhFormin1 fragments were expressed in B834 *E. coli* induced with 500 μM isopropyl-β-D-thiogalactopyranoside (IPTG) for 14-16 hours at 20°C. For crystallization of the GBD-FH3 domain tandem, a selenium-containing derivative was produced by induction in minimal media containing selenomethioine (Molecular Dimensions, Apopka, FL). Bacterial cells were pelleted by centrifugation and resuspended in N1 buffer composed of 50 mM Tris pH 7.5, 250 mM NaCl, 10 mM imidazole, 1 mM DTT, and 5% (v/v) glycerol. Bacteria were lysed using pressure homogenization with an Emulsiflex (Avestin; Ottawa, Canada). Cellular lysates were cleared with centrifugation at 100,000 x *g* for 60 minutes at 4°C, and the resulting supernatant was applied to a nickel-nitrilotriacetic acid (NTA) resin FPLC column (FF HisTrap crude; GE Healthcare, Piscataway, NJ), washed with N1 buffer plus 20 mM imidazole before elution in N1 buffer with 300 mM imidazole. EhFormin1 GBD-FH3 tandem selenomethioine protein was pooled and dialyzed overnight in imidazole-free N1 buffer with His$_6$-TEV protease to cleave the N-terminal affinity tag. The dialysate was then passed over a second NTA column to remove TEV protease and uncleaved protein. All

EhFormin1 fragments were resolved using a calibrated size exclusion column (HiLoad 16/60 Superdex 200, GE Healthcare) in S200 buffer (50 mM HEPES pH 8.0, 150 mM NaCl, and 5 mM DTT). All proteins were concentrated to 0.5 - 2 mM and snap frozen in a dry ice/ethanol bath for storage at -80°C.

Actin for *in vitro* polymerization assays was purified from rabbit skeletal muscle acetone powder as described previously [25], and further purified with gel filtration chromatography. The actin was polymerized and conjugated with pyrene as described previously [26]. Following a final gel filtration step, the pyrene-actin was stored at 2 mg/mL and 4°C in G buffer (2 mM Tris pH 8.0, 0.2 mM $CaCl_2$, 0.5 mM ATP, and 5 mM DTT). Protein concentration was determined by A_{280nm} measurements upon denaturation in 8 M guanidine hydrochloride, using predicted extinction coefficients for each protein (http://us.expasy.org/tools/protparam.html).

5.3.2 Actin co-sedimentation

Rabbit skeletal muscle actin for co-sedimentation assays was purchased from Cytoskeleton, Inc. (Denver, CO). EhFormin1 FH2 domain and F-actin co-sedimentation assays were conducted as previously described for talin [27]. Briefly, actin was diluted to 0.4 mg/mL in buffer containing 5 mM Tris pH 8.0, 0.2 mM $CaCl_2$, 0.2 mM ATP, and 0.5 mM DTT and polymerized by addition of 50 mM KCl and 2 mM $MgCl_2$, followed by incubation at room temperature for 1 hr. EhFormin1 fragments were incubated alone or with a 2:1 molar excess of polymerized actin in binding buffer (10 mM Tris pH 7.0, 1 mM ATP, 0.2 mM DTT, 1 mM EGTA, 0.1 mM $CaCl_2$, and 2 mM $MgCl_2$) for 1 hour at room temperature. Samples were centrifuged at 100,000 x *g* and 20°C for 15 minutes. Proteins in the supernatant and pellet

fractions of each experiment were resolved by SDS-PAGE and stained with Coomassie Brilliant Blue.

5.3.3 Actin polymerization *in vitro*

Monomeric pyrene-actin (~20% pyrene labeled, ~80% unlabeled) was diluted to 40 μM in buffer containing 25 μM Tris pH 7.4 and 5 mM DTT. 25 μL of diluted pyrene-actin (10 μM final concentration) and various amounts of EhFormin1 fragments and/or EhRho1 were brought to a volume of 95 μL in S200 buffer (50 mM HEPES pH 8.0, 150 mM NaCl, and 5 mM DTT). Buffer conditions were held constant in all comparison experiments, since ionic strength and pH are known to influence actin polymerization kinetics [28]. Fluorescence of the pyrene moiety was monitored throughout the experiment at 30-second intervals using a FluoroLog modular spectrofluorometer (Horiba, Ann Arbor, MI) with excitation and emission wavelengths of 365 nm and 407 nm, respectively. Following establishment of a stable baseline fluorescence (~5 minutes), polymerization was initiated by addition of 5 μL polymerization buffer (1 mM Tris pH 7.4, 500 mM KCl, 20 mM $MgCl_2$, 2 mM ATP, and 5 mM DTT). Polymerization was allowed to proceed for at least 1 hour. The relative rate of polymerization was estimated by measuring the slope (fluorescence units/second) of the actin polymerization curve at 50% of the maximal fluorescent signal, as previously described [28]. All slope measurements were averaged across at least 3 replicate experiments and statistical significance determined by Student's t-test.

5.3.4 Surface plasmon resonance

SPR-based measurements of protein-protein interactions were performed on the Biacore 3000 of UNC's Center for Structural Biology (GE Healthcare), as described previously [7]. Approximately 10,000 resonance units (RUs) of purified His$_6$-EhFormin1 GBD-FH3 tandem and 5,000 RUs of His$_6$-EhFormin1 fusion were separately immobilized on a nickel-NTA biosensor chip (GE Healthcare) using covalent capture coupling as previously described [29]. An empty surface served as a negative control. Experiments were performed in running buffer containing 50 mM HEPES pH 7.4, 150 mM NaCl, 0.05% NP-40 alternative (Calbiochem), 50 μM EDTA, and 1 mM MgCl$_2$. For assessment of kinetic binding properties, 3 injections of 5 μM EhRho1·GTPγS were performed at 10 μL/min with a 300 second dissociation phase. k_{obs} was obtained by fitting the average of three injections with a single phase exponential association function using GraphPad Prism v5.0. Similarly, k_{off} was obtained by fitting the average data immediately following the injections with a single phase exponential dissociation function. Since k_{obs} is dependent on the concentration of analyte, k_{on} was derived by $k_{on} = (k_{obs} - k_{off})$ / (analyte concentration). An affinity constant was derived from the kinetic data by $K_D = k_{off} / k_{on}$. For equilibrium binding analyses, multiple injections were performed with increasing concentrations of GTPγS-loaded EhRho1, EhRacC, EhRacD, and EhRacG, and an affinity constant derived as described previously [7].

5.3.5 Crystallization and structure determination

A complex of EhRho1·GTPγS and the selenomethionine derivative of the EhFormin1 GBD-FH3 tandem (a.a. 69-418) was assembled by mixing the two proteins at a 1:1 molar ratio to a total concentration of 15 mg/mL in crystallization buffer (50 mM Tris pH 8.0, 250 mM NaCl,

2.5% (v/v) glycerol, 5 mM DTT, 50 μM GTPγS, and 1 mM MgCl$_2$) and incubation for 30 minutes at room temperature. Crystals of EhRho1·GTPγS / EhFormin1 were obtained by vapor diffusion from hanging drops at 18°C. The protein solution was mixed 1:1 with and equilibrated against crystallization solution containing 18% PEG 3350, 100 mM Tris pH 8.5, and 200 mM MgCl$_2$. Clusters of six hexagonal rod crystals grew to ~200 x 100 x 100 μm over 3 days, but diffraction was limited to >3 Å resolution. Crystal clusters were used to microseed similar crystallization experiments using the method described previously [30]. Microseeded experiments yielded single hexagonal rod crystals (~300 x 150 x 150 μm) over 5 days, exhibiting the symmetry of space group P6$_1$ (a = b = 138.6 Å, c = 57.8 Å, α = β = 90°, γ = 120°) and containing one EhRho1/EhFormin1 dimer in the asymmetric unit. For data collection at 100K, crystals were serially transferred for ~1 minute into crystallization solution supplemented with 30% (v/v) glycerol in 10% increments and plunged into liquid nitrogen. Single wavelength (0.9795 Å) anomalous diffraction data were collected at the GM/CA-CAT 23-ID-B beamline at the Advanced Photon Source (Argonne National Laboratory). Data were processed using the HKL-2000 program [31]. Heavy atom searching, experimental phasing, and automated model building were performed with Phenix AutoSol [32]. Heavy atom searching identified 13 of 13 possible sites, and refinement yielded an estimated Bayes correlation coefficient of 51 to 2.6 Å resolution. After density modification, the estimated Bayes correlation coefficient increased to 58. ~75% of the model was constructed automatically, and the remaining portion was built manually throughout the refinement. The current model (Table 5.1) contains a single EhRho1/EhFormin1 dimer with EhRho1 bound to GTPγS and magnesium. Refinement was carried out against peak anomalous data with phenix.refine [32], keeping Bijvoet pairs separate, interspersed with

manual model revisions using the program Coot [33]. Refinement consisted of conjugate-gradient minimization and calculation of individual atomic displacement and translation/libration/screw (TLS) parameters [34]. Residues 1-20 and 185-186 of EhRho1 and residues 69-72 and 378-418 of EhFormin1 could not be located in the electron density. The model exhibits excellent geometry as determined by MolProbity [35]. A Ramachandran analysis of protein residue backbone angles identified 93% favored, 7% allowed, and 0% disallowed. Coordinates and structure factors are deposited in the RCSB Protein Data Bank (id 4DVG).

5.4   RESULTS

5.4.1 *E. histolytica* Formin1 modulates actin filament formation

The Diaphanous-related formins catalyze actin polymerization through the FH2 domain [12]. To examine the potential interaction of EhFormin1 with actin, two FH2 domain-containing fragments of this protein were generated and expressed from *E. coli*: FH2-DAD (a.a. 731-1182) and FH2 (731-1127) (see Figure 5.1). To determine whether the EhFormin1 FH2 domain interacts with actin, co-sedimentation assays were performed [27] using rabbit skeletal muscle-derived filamentous actin. The EhFormin1 fragments FH2-DAD and FH2 were highly soluble alone, but both co-sedimented with polymerized actin (Figure 5.2A), suggesting a direct actin/FH2 domain interaction independent of the putative DAD motif. Mutation of a conserved surface lysine (K964, Figure 5.3) required for actin binding by other formin FH2 domains [36] abolished actin co-sedimentation, suggesting a conserved mode of FH2 domain/actin interaction (Figure 5.2A). To determine whether the EhFormin1/actin interaction altered actin polymerization kinetics, *in vitro* polymerization assays were

performed with pyrene fluorophore-labeled actin (~20% labeled). The FH2 and FH2-DAD fragments each decelerated actin polymerization in a concentration-dependent fashion (Figure 5.2B, C), as observed by the slope of the actin polymerization curve [28]. Some isolated FH2 domains, such as that of Cdc12 in fission yeast, slow overall actin polymerization *in vitro* despite a positive effect on actin filament formation in a cellular context [37]. Like Cdc12, the FH2 domain of EhFormin1 may cap actin filament barbed ends and require an FH1 domain-associated profilin to accelerate polymerization [38]. However, no FH1 domain-containing EhFormin1 fragments could be obtained as a soluble recombinant protein from *E. coli*, precluding a direct test of this hypothesis. Relatively high concentrations of all EhFormin1 fragments were required to significantly alter actin polymerization kinetics, likely reflecting a low affinity EhFormin1/actin interaction. Alignment of the EhFormin1 FH2 domain sequence with Diaphanous-related formins from other species [36] (28% identity and 48% similarity to yeast Bni1) indicates only moderate conservation of residues known to participate in the actin interaction (Figure 5.3). Similarly, the *E. histolytica* actin sequence differs significantly from rabbit skeletal muscle actin at FH2-domain interaction sites (not shown), suggesting that high EhFormin1 fragment concentrations may be required to overcome cross-species sequence and/or structural divergence.

5.4.2 EhFormin1 is autoinhibited by N- and C-terminal interactions

Other Diaphanous-related formins are maintained in an inactive conformation by interactions between the D̲iaphanous a̲utoinhibitory d̲omain (DAD motif) and a surface of the Armadillo repeat portion of the FH3 domain, also called the D̲iaphanous i̲nhibitory d̲omain (DID) [16,

17]. Inspection of the C-terminus of EhFormin1 revealed considerable sequence divergence from the core DAD motif conserved among other known formins (MDxLLExL) (Figure 5.4C). Accordingly, we wondered whether the DID/DAD autoinhibitory interaction would be conserved in the case of EhFormin1. Surprisingly, the EhFormin1 FH2-DAD long fragment was seen to interact with the N-terminal GBD-FH3 domain tandem, as determined by surface plasmon resonance (Figure 5.4A). This intramolecular interaction required the divergent putative DAD motif, since the FH2 domain alone did not bind the GBD-FH3 domain tandem (Figure 5.4A,B). To determine the effects of the GBD-FH3 domain interaction on FH2 domain-catalyzed actin filament formation, *in vitro* polymerization assays were conducted in the presence or absence of a molar excess of GBD-FH3 tandem protein. Addition of the GBD-FH3 tandem selectively affected the DAD motif-containing construct, returning the rate of actin polymerization in its presence to one indistinguishable from actin alone (Figure 5.4D). These results suggest that the C-terminus of EhFormin1 forms a DAD motif-dependent interaction with the N-terminal GBD-FH3 domain region that prevents the modulation of actin polymerization rate by the FH2 domain.

5.4.3 Interaction of the EhFormin1 GBD-FH3 domain tandem with EhRho1 reverses autoinhibition of the FH2 domain

In previous work [7], we demonstrated that EhRho1 binds the GBD-FH3 domain tandem of EhFormin1 (a.a. 69-445) selectively in its GTPγS-bound, activated conformation. Since some minor degradation of that particular GBD-FH3 fragment was reported during expression and purification (*e.g.*, see Fig. 1B of [7]), we generated multiple alternative constructs containing the GBD-FH3 region, finding that amino acids 69-418 of EhFormin1 were highly stable as a

recombinant protein fragment. The residues removed from this smaller fragment (*i.e.*, a.a. 419-445) correspond to the dimerization domain of mDia1 [20]. The single-celled *E. histolytica* parasite expresses ~20 Rho family GTPases, raising the possibility of highly specific interactions between Rho GTPases and their signaling effectors. The EhFormin1 GBD-FH3 domain tandem (a.a. 69-418) was observed to interact selectively with EhRho1 to the exclusion of three other Rho GTPases tested (Figure 5.5A).

Given apparent endoproteolytic sensitivity of its unstructured, proline-rich FH1 domain, recombinant full length (and thus autoinhibited) EhFormin1 could not be produced and purified. To circumvent this problem, we produced a construct (hereafter referred to as "EhFormin1 fusion") consisting of the N-terminal GBD-FH3 domain tandem and the C-terminal FH2-DAD fragment, connected by a 40-residue linker to simulate the presumably flexible FH1 domain (see Figure 5.1). While either the EhFormin1 fusion or EhRho1·GTPγS alone had no measurable effect on *in vitro* actin polymerization, the EhFormin1 fusion/EhRho1·GTPγS complex was observed to inhibit actin filament formation (Figure 5.5B) to a similar degree as the corresponding FH2-DAD construct (Figure 5.2C). This finding suggests that EhRho1 selectively engages the N-terminus of EhFormin1, freeing the C-terminal FH2 domain to regulate actin filament formation. Using SPR, EhRho1·GTPγS was shown to bind this EhFormin1 fusion protein with ~3 µM affinity (Figure 5.6).

5.4.4 Structural features of the EhRho1/EhFormin1 complex

The sequence of the GBD-FH3 domain tandem within EhFormin1 is highly divergent compared to other known Diaphanous-related formins, with mDia1 being the closest mammalian homolog (Figure 5.7). EhRho1 also differs significantly from mammalian Rho

GTPases and other *E. histolytica* Rho GTPases (Figure 5.7), particularly given its lack of a signature Rho insert helix [7]. We sought a crystal structure of EhRho1 bound to EhFormin1 to allow structure-based comparison with the mammalian RhoC/mDia1 complex and to elucidate the determinants of a highly selective Rho/effector interaction. Well-diffracting crystals of EhRho·GTPγS/EhFormin1 GBD-FH3 domain tandem were obtained with aid of microseeding (see Experimental Procedures). Molecular replacement attempts with structural models of either EhRho1 (PDB id 3REF) or the mDia GBD-FH3 tandem (PDB id 3EG5) did not produce electron density maps suitable for accurate modeling, likely given the divergent conformation of the GBD-FH3 domain tandem. A selenomethioine derivative of the EhFormin1 GBD-FH3 fragment was therefore generated and crystallographic phases determined by single-wavelength anomalous dispersion (SAD). For data collection and refinement statistics, see Table 5.1.

The overall structure of the EhRho1·GTPγS/EhFormin1 complex resembles that of human RhoC·GppNHp/mDia1 [20] except that the EhFormin1 fragment used for crystallography (a.a. 69-418) lacks the dimerization domain, and thus lacks the dimeric quaternary structure seen for mDia1 (Figure 5.8). The Armadillo repeats of the DID domain in EhFormin1 are rotated ~40° away from EhRho1 relative to the conformation seen in the mammalian homolog (Figure 5.9). EhRho1 engages the EhFormin1 GBD and the N-terminal portion of the DID Armadillo repeats through its two mobile switch regions and its α3 helix, as also seen in mammalian homologs (Figure 5.8). However, the GTPase binding domain of EhFormin1 differs from that of mDia by a shortened second helix and an elongated α3 helix with clearly defined and continuous electron density. As a result of a shortened α2 helix, the GBD of EhFormin1 remains farther away (~4 Å) from the putative DAD-binding site within

the DID (Figure 5.10) which may indicate a slightly different mechanism of Rho-induced activation, since the contingently folded GBD of mDia1 is thought to contribute to DAD displacement by directly obstructing the DAD-binding site of the DID [18]. EhFormin1 also has an elongated $\alpha12$-$\alpha13$ loop relative to mDia1 (Figure 5.8). This loop is near the DAD-binding site of the DID in mDia1 [18]; together with the highly divergent nature of the putative DAD motif within EhFormin1 (Figure 5.4), the presence of this elongated loop suggests that a unique mode of DAD-mediated autoinhibition may exist in the case of EhFormin1.

The conformation of EhRho1 in the complex is nearly identical with that we have previously reported [7] of free EhRho1·GTPγS (Cα r.m.s.d. 0.4 Å, PDB id 3REG), and is also similar to RhoC in the homologous RhoC·GppNHp/mDia1 complex (Cα r.m.s.d. 1.5 Å, PDB id 1Z2C) (Figures 5.9, 5.11). The Rho insert helix of RhoC approaches the C-terminal end of the mDia1 DID (Figures 5.8B, 6), leading to the hypothesis that a secondary binding site (beyond the switch region/GBD interaction) may be important for mDia1 activation [14, 20]. In contrast, the lack of this insert helix within EhRho1 and the relative rotation of the EhFormin1 DID indicate that such a secondary interaction is absent in the *E. histolytica* orthologues (Figures 5.8A, 6). An EhRho1 molecule from the adjacent asymmetric unit is interposed between the non-uniformly structured $\beta5$-$\alpha4$ loop of EhRho1 and the last Armadillo repeat of the EhFormin1 DID, raising the possibility that crystal contacts could be responsible for the rotation of the DID domain away from EhRho1 relative to the RhoC/mDia1 complex. However, this is unlikely given the large magnitude of rotation and the absence of significant contacts between the interposed EhRho1 and the DID domain;

crystal contacts exist only between EhRho1 molecules in adjacent asymmetric units (see Figure 5.12).

The EhRho1/EhFormin1 interface is dominated by hydrophobic interactions between the two switch regions of the GTPase and the GBD (Figure 5.13A). The critical EhRho1 residues at this interaction site, such as Phe54 of switch 1 and Leu84 of switch 2, are conserved across the related GTPases (Figure 5.7A), suggesting that these hydrophobic interactions, while important for binding, do not necessarily determine the observed specificity of EhFormin1 for EhRho1. The non-conserved Asp91 of EhRho1 is positioned to potentially form ionic contacts with EhFormin1 α2 helix residues Lys108 and Lys112 (Figure 5.13A). The EhRho1 α3 helix contacts the first two Armadillo repeats of the EhFormin1 DID domain, primarily through residues His120 and Tyr121 (Figure 5.13B). The imidazole ring of His120 is oriented for a hydrogen bond interaction with Glu151 of EhFormin1. Interestingly His120 and Tyr121 are conserved in RacD, but not RacC or RacG (Figure 5.7A), suggesting that these two interface residues may be important for specificity. Indeed, mutation of EhRho1 His120 to Gln, the analogous residue in EhRacG, drastically reduced affinity for the EhFormin1 GBD-FH3 tandem as measured by SPR (Figure 5.13D). Finally, the EhRho1 switch 2 residue Arg83 is inserted into a groove between the GBD and the DID domain, a region of slightly negative charge as indicated by vacuum electrostatic calculations (Figure 5.13C). Arg83 is within hydrogen bonding distance to multiple exposed peptide backbone carbonyl groups in this region (Phe156 and Arg157; see Figure 5.14). Arg83 of EhRho1 corresponds to Arg68 of human RhoC, which engages its formin effector in a strikingly similar fashion (Figure 5.9B) and is required for high affinity interaction [18]. However, RhoC Arg68 forms hydrogen bonds with the side chain of mDia N217 while

EhRho1 Arg83 is exclusively within hydrogen-bonding distance of backbone carbonyl groups (Figure 5.9, 5.14). This arginine is also present in EhRacG, but not other Rho family GTPases (Figure 5.7A), implicating this particular residue as a likely determinant of specificity for the EhRho1/EhFormin1 interaction. Mutation of EhRho1 Arg83 to the corresponding Gln in EhRacC and EhRacD drastically reduced affinity for EhFormin1 (K_D >100 μM compared to ~3 μM for wild type EhRho1; Figure 5.13D).

5.5   DISCUSSION

The isolated FH2 domain of EhFormin1 was observed to slow actin polymerization *in vitro*, a phenomenon also exhibited by the corresponding domain from fission yeast Cdc12 [37], suggesting a possibly similar actin barbed-end capping interaction. Other formin regions mediate interactions with proteins that can also influence its activation state as an agent of actin polymerization [39]. Perhaps the best studied is the interaction of actin-bound profilin with the proline-rich FH1 domain [40]. Profilins can increase the rate of actin filament elongation by formins, possibly by increasing the local concentration of actin monomers to be included in the growing filament [15]. In the case of fission yeast Cdc12, the FH1 domain and associated profilin are required, in combination with the FH2 domain, for acceleration of actin polymerization [38]. While *in vitro* actin polymerization assays with EhFormin1 fragments provide important mechanistic insights into activity of the isolated protein, it is important to note that the full-length protein in a cellular context is likely also modulated by subcellular localization and interactions with multiple other proteins. A limitation of the pyrene actin polymerization assay used in this study is its inability to differentiate the effects of EhFormin1 on nucleation of new actin filaments *versus* accelerated elongation of existing

205

filaments. Nucleation and elongation are catalyzed with varying efficiency among formins [12], and further studies are thus necessary to elucidate fully the mechanisms by which EhFormin1 modulates actin polymerization.

DRFs are commonly autoinhibited by an intramolecular interaction between the N-terminal DID domain and the C-terminal DAD domain, consisting of a core MDxLLExL motif followed by a polybasic region [19]. However, the C-terminus of EhFormin1 contains a highly divergent segment (MExAANxG) corresponding to the core DAD motif, suggesting a potentially unique mode of regulation. Despite poor conservation of the putative DID/DAD interface, the N- and C-terminal fragments of EhFormin1 were seen to bind one another in a DAD-dependent fashion, resulting in inhibition of FH2 domain-mediated modulation of actin polymerization. Furthermore, a fusion protein containing the N-terminal GBD-FH3 tandem and the C-terminal FH2-DAD tandem mimicked the presumably autoinhibited state of full-length EhFormin1, having no measureable effect on actin polymerization kinetics; binding of activated EhRho1·GTPγS to the GBD-FH3 region reversed the apparent autoinhibition of the EhFormin1 fusion, suggesting that the EhRho1/EhFormin1 interaction is sufficient to free the FH2 domain for modulation of actin polymerization.

The crystal structure in this study provides only the second exemplary snapshot of an interaction between a Rho GTPase and formin effector. Thus a comparison of the *E. histolytica* and human Rho/formin complexes reveals consistent structural features that are conserved across species and likely of shared importance for Rho-mediated activation of formins, as well as differences that may reflect properties of individual proteins, such as formin specificity for particular Rho GTPases. Our structural model of the EhRho1/EhFormin1 complex reveals a similar interface to that of human RhoC/mDia1

206

despite a distant evolutionary relationship as evidenced by substantial sequence divergence. The GTPase binding domain of EhFormin1 is quite similar to that of mDia1, with a predominant hydrophobic patch engaging the switch regions of EhRho1 [20]. Arg83 of EhRho1 projects into a relatively negatively charged groove between the GBD and DID domains in a manner highly homologous to Arg68 of RhoC [20]. However, the EhRho1 Arg83 residue forms hydrogen bonds exclusively with backbone carbonyl groups rather than asparagine side chains, as seen in the mammalian homolog [20]. The EhRho1 switch 2 residue Arg83, together with the buried side chains of His120 and Tyr121 on the α3 helix of EhRho1, are important determinants of Rho GTPase specificity in binding EhFormin1. EhRacC and EhRacD lack the critical switch 2 arginine, while EhRacG lacks the histidine-tyrosine tandem, and all three of these Rac subfamily GTPases are unable to bind EhFormin1. Furthermore, mutation of EhRho1 residues Arg83 to the corresponding EhRacC/D glutamine or His120 to the EhRacG-like glutamine each reduced the affinity of EhRho1 for EhFormin1 by >100-fold. Such strict specificity of Rho GTPase and effector interactions is likely of particular importance in *E. histolytica* where at least 19 Rho family GTPases are apparently expressed in a single cell [7].

Several lines of evidence suggest that the mechanisms of EhFormin1 autoinhibition and its activation by EhRho1 may differ significantly from that of the well-studied mDia1 homolog [16, 17]. There is poor EhFormin1 sequence conservation in the putative core DAD motif and DAD-binding surface on the DID, and EhFormin1 has a uniquely elongated α12-α13 loop near the putative DID/DAD interaction site. Furthermore, the GBD of EhFormin1 has a shortened α2 helix that would not directly obstruct the DAD-binding surface on the DID as modeled for mDia1 (Figure 5.10) [18]. Finally, EhRho1 lacks a Rho insert helix and

does not approach the C-terminus of the DID domain, indicating that a secondary EhRho1/EhFormin1 interaction between these two regions does not exist, in contrast with RhoC/mDia1.

Previous studies of EhFormin1 in the context of *E. histolytica* trophozoites [21], together with our findings ([7] and this paper), suggest that an EhRho1/EhFormin1 signaling axis may be important for formation of complex actin structures within pseudopodia and phagocytic and pinocytic vesicles, particularly in response to extracellular cues such as serum factors. EhFormin1 also apparently exerts effects on trophozoite mitosis and cytokinesis [21]. Interestingly, both EhRho1 and EhFormin1 are enriched in *E. histolytica* uropods [41], suggesting that the GTPase/effector pair may also regulate actin polymerization at the trailing edge during trophozoite migration and/or surface receptor capping critical for immune response evasion [5]. Knowledge of the structural determinants defining the EhRho1/EhFormin1 interaction, and of its differences from mammalian GTPase/formin complexes (Fig. 5.13), should assist in understanding the contributions of this actin polymerization pathway to *E. histolytica* infectivity and invasiveness.
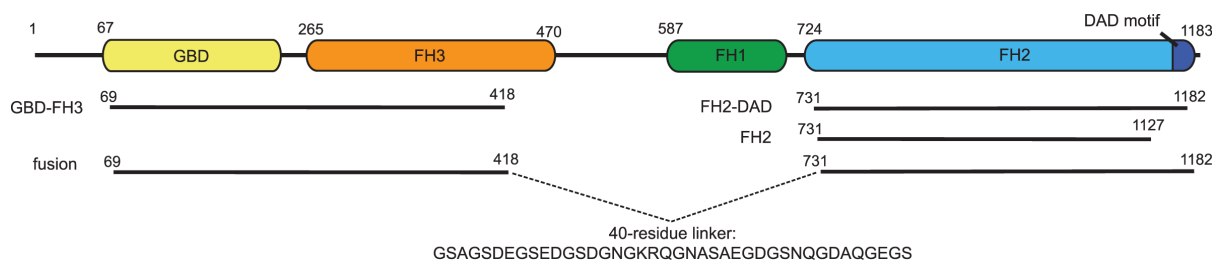
Figure 5.1. EhFormin1 domain structure and constructs used in this study. The domain structure of EhFormin1 is represented with residue numbers indicating domain borders. The EhFormin1 fusion consists of residues 69-418 and 731-1182 with an intermediate 40-residue synthetic linker designed to be soluble and flexible.
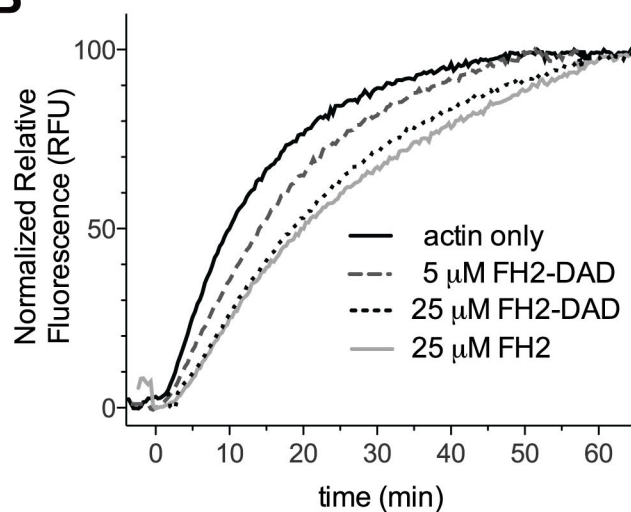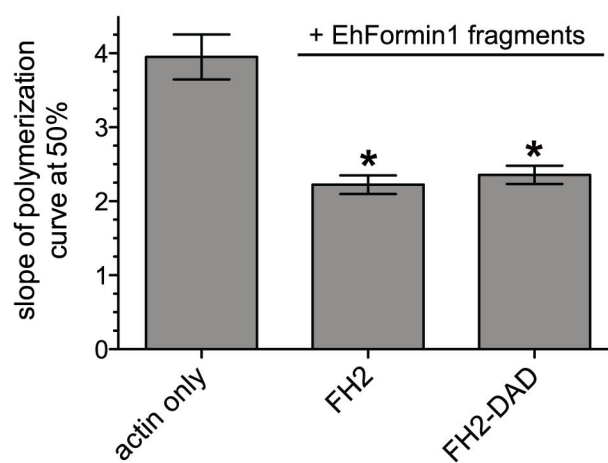
Figure 5.2. The FH2 domain of EhFormin1 modulates actin filament formation. A) Actin co-sedimentation assays demonstrate that the EhFormin1 FH2 domain fragment (a.a. 731-1127) and the FH2-DAD combination (a.a. 731-1182) both associate with pre-formed

filamentous actin derived from rabbit skeletal muscle. Mutation of the conserved Lys964 that is critical for other formin FH2/actin interactions abolished co-sedimentation. *S* and *P* represent the soluble and pellet fractions following high-speed centrifugation, respectively. **B)** Indicated fragments of EhFormin1, each containing the FH2 domain, modulate actin polymerization *in vitro*, as measured by pyrene-actin polymerization assays. **C)** The EhFormin1 FH2 and FH2-DAD fragments are both seen to slow actin polymerization, as quantified by measuring the slope of each fluorescence curve at 50% complete polymerization. Error bars represent standard error for three or more replicate experiments. * indicates a statistically significant difference from actin only ($p < 0.05$).

```
                                                          lasso
                  720        730       740       750       760       770        780
EhFormin1        -----FGFRPAAPV---INGLVDTLPKPKSKVKNFMWQKMNDKNLSGTLFTKLDTLDKTKQKLDLSLIEAAFKVPEK--------VKP
S.c. Bni1        MPASQIKSAVTSPLLPQSPSLFEKYPRPHKKLKQLHWEKLD---CTDNSIWGTGKAEKFADDLYEKGVLADLEKAFA--------ARE
mDia1            -----GFGVPAAPV-LPFGLTPKKVYKPEVQLRRPNWSKFVAEDLSQDCFWTKVKEDRFENNELFAKLTLAFSAQTKTSKAKKDQEGG
mDia2            -----VFPLLSGPIELPYGMKQKKLYKPDIPMKRINWSKIEPKELSENCVWLKLKEEKYENADLFAKLALTFPSQMK---GQRNTEAA
D.m. Diaphanous  -----MMMGPMVPV-LPHGLKPKKKWDVKNPMKRANWKAIVPAKMSDKAFWVKCQEDKLAQDDFLAELAVKFSSKPV----KKEQKDA

                      790       800       810       820       830       840             850
EhFormin1        VADAAPKETKKQGPICILDAKNNQTFTILLK-GFKNKTPEEVCAAINSVDQTVFEETSTIKTMLKA---------------LEDAKEE
S.c. Bni1        IKSLASKRKEDLQKITFLSRDISQQFGINLH-MYSSLSVADLVKKILNCDRDFLQTPSVVEFLSKSEIIEVSVNLARNYAPYSTDWEG
mDia1            EEKKS-VQKKKVKELKVLDSKTAQNLSIFLG-SFR-MPYQEIKNVILEVNEAVLTES-MIQNLIKQ---------------MPEP-EQ
mDia2            EENRSGPPKKKVKELRILDTKTAQNLSIFLG-SYR-MPYEEIKNIILEVNEEMLSEA-LIQNLVKY---------------LPDQ-NA
D.m. Diaphanous  VDKPTTLTKKNV-DLRVLDSKTAQNLAIMLGGSLKHLSYEQIKICLLRCDTDILSSN-ILQQLIQY---------------LPPP-EQ

                      860       870       880       890       900       910       920        930
EhFormin1        IGQVEEHIKEK-GTDNLGSAEM--FVHALNGVSNITLKLNSFQSKMELPVKLEEVGPNIKHVIKASNQLLDSKKFIHLMEVILLMGNF
S.c. Bni1        VRNLEDAKPPEKDPNDLQRADQIYLQLMVNLESYWGSRMRALTVVTSYEREYNELLAKLRKVDKAVSALQESDNLRNVFNVILAVGNF
mDia1            LKMLSELKEE---YDDLAESEQ--FGVVMGTVPRLRPRLNAILFKLQFSEQVENIKPEIVSVTAACEELRKSENFSSLLELTLLVGNY
mDia2            LRELAQLKSE---YDDLCEPEQ--FGVVMSTVKMLRPRLTSILFKLTFEEHVNNIKPSIIAVTLACEELKKSESFKRLLELILLVGNY
D.m. Diaphanous  LKRLQEIKAK---GEPLPPIEQ--FAATIGEIKRLSPRLHNLNFKLTYADMVQDIKPDIVAGTAACEEIRNSKKFSKILELILLLGNY

                 post
                 940    950       960       970       980       990       1000      1010      1020
EhFormin1        LNSGTAKGNVAGFKFETLNKTLDTKTGDNKRTLLHIIQEYAAEKLKEEVDGWQNEVTEIPLAAKVPGAQFESDIKSLEKMYNDIQTAV
S.c. Bni1        MN--DTSKQAQGFKLSTLQRLTFIKDTTNSMTFLNYVEKIVRLNYPSFND-FLSELEPVLDVVKVSIEQLVNDCKDFSQSIVNVERSV
mDia1            MNAGSRNAGAFGFNISFLCKLRDTKSADQKMTLLHFLAELCENDHPEVLK-FPDELAHVEKASRVSAENLQKSLDQMKKQIADVERDV
mDia2            MNSGSRNAQSLGFKINFLCKIKDTKSADQKSTLLHFLAEICDEKYRDILK-FPDELEHVESAGKVSAQILKSNLVAMEQSILHLEKNI
D.m. Diaphanous  MNSGSKNEAAFGFEISYLTKLSNTKDADNKQTLLHYLADLVEKKFPDALN-FYDDLSHVNKASRVNMDAIQKAMRQMNSAVKNLETDL

                 1030      1040      1050      1060      1070      1080      1090      1100
EhFormin1        GKIPEDGS-------KFHPIMKKFLEETRGKLDAVAKDFAEMNEIYKKAIQYMALDIAKPPPPEEFFTTLSTFISNWNRVTEENKKLQ
S.c. Bni1        EIGNLSDSSKFHPLDKVLIKTLPVLPEARKKGDLLEDEVKLTIMEFESLMHTYGEDSGDKFAKISFFKKFADFINEYKKAQAQNLAAE
mDia1            QNFPAATDE----KDKFVEKMTSFVKDAQEQYNKLRMMHSNMETLYKELGDYFVFDPKK-LSVEEFFMDLHNFRNMFLQAVKEN---Q
mDia2            KNFPPAESH----HDKFVEKMMSFTQNAREQYDKLSTMHSNMLKLYESLGEYFIFDPNT-VNMEEFFGDLNTFRTLFLEALKEN---H
D.m. Diaphanous  QNNKVPQCD----DDKFSEVMGKFAEECRQQVDVLGKMQLQMEKLYKDLSEYYAFDPSK-YTMEEFFADIKTFKDAFQAAHNDN---V

                                                        DAD motif
                 1110      1120      1130      1140      1150      1160      1170      1180
EhFormin1        AQREKEAKKAEAAAAKKAGKKKTSLAPI----PGEEGNMVKGRRQNPLAMEAAANVGGLKKTGFTFKKKTTKVLKTMDI-----
S.c. Bni1        EEERLYIKHKKIVEEQQKRAQEKEKQKENSNSPSSEGNEEDEAEDRRAVMDKLLEQ---LKNAGPAKSDPSSARKRALVRKKY-
mDia1            KRRETEEKMRRAKLAKEKAEKERLEKQQ----KREQLIDMNAEGDETGVMDSLLEA---LQSGAAFR--RKRGPRQVNRK-----
mDia2            KRKEMEEKSRRAKLAKEKAEQEKLERQK----KKQLIDINKEGDETGVMDNLLEA---LQSGAAFRDRRKRIPRNPDNRRPPL
D.m. Diaphanous  RVREELEKKRRLQEAREQSAREQQERQQ----RKKAVVDMDAPQTQEGVMDSLLEA---LQTGSAFGQRNRQARRQRPAGAER-
```

Figure 5.3. Multiple sequence alignment of FH2 domains and DAD motifs. The FH2 domain and putative DAD motif of EhFormin1 is aligned with other known autoinhibited Diaphanous-related formins. *S.c.* indicates *Saccharomyces cerevisiae* and *D.m.* indicates *Drosophila melanogaster*, while mDia isoforms are derived from *Mus musculus*. Black dots indicate residues that contact actin the crystal structure of the *S.c.* Bni1/actin complex [36] (PDB id 1Y64).
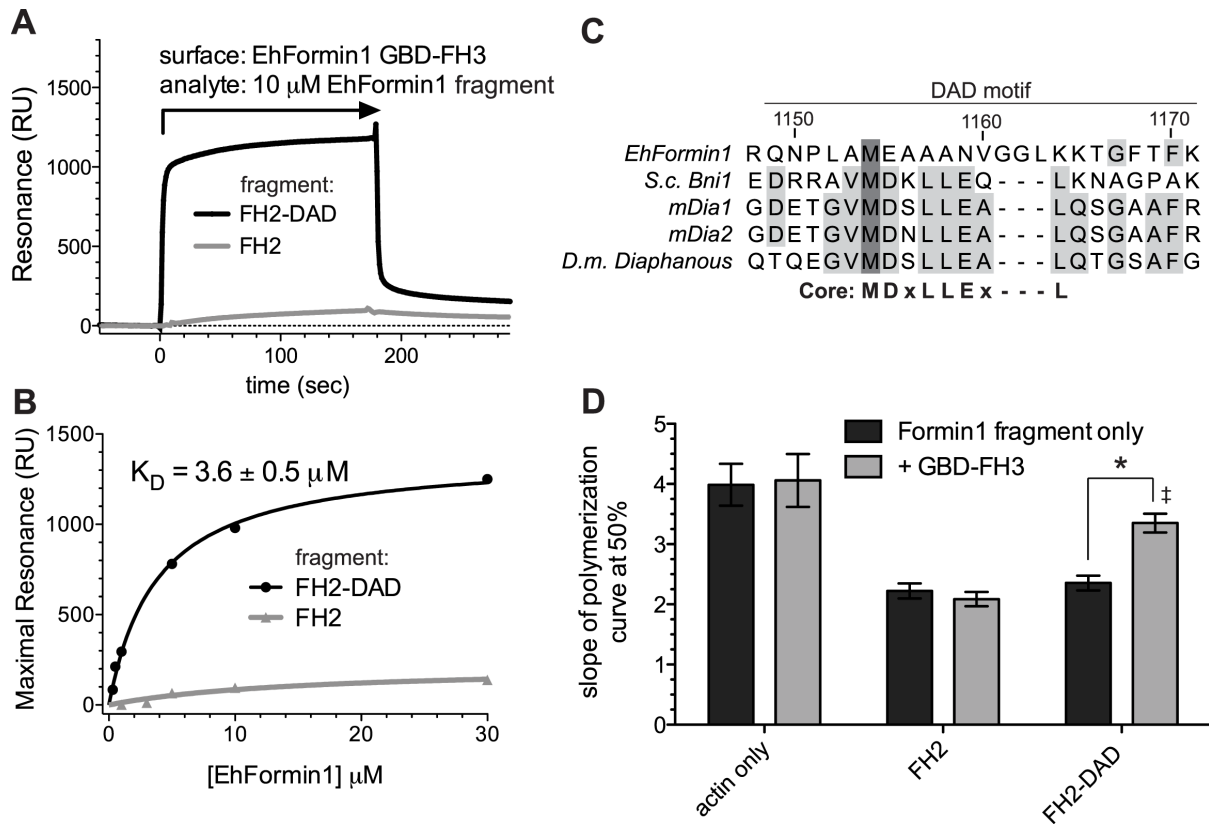
Figure 5.4. EhFormin1 is autoinhibited by N- and C-terminal domain interactions. A) The N-terminal GBD-FH3 domain tandem (a.a. 69-418) of EhFormin1 binds to the C-terminal FH2-DAD tandem (a.a. 731-1182), but not the FH2 domain alone (a.a. 731-1127) as determined by surface plasmon resonance. **B)** Equilibrium binding analyses revealed a DAD motif-dependent low micromolar affinity interaction with the EhFormin1 GBD-FH3 fragment. **C)** The DAD motif region, with a core motif (MDxLLExL) highly conserved among other known Diaphanous-related formins, is divergent in EhFormin1. *S.c.* indicates *Saccharomyces cerevisiae* and *D.m.* indicates *Drosophila melanogaster*, while mDia isoforms are derived from *Mus musculus*. **D)** *In vitro* actin polymerization assays indicate a DAD-motif dependent interaction between the N- and C-terminal fragments of EhFormin1 that prevents modulation of actin polymerization. Addition of a molar excess of GBD-FH3 domain tandem (100 µM) had no effect on the isolated FH2 domain from EhFormin1 or on actin alone. However, the GBD-FH3 tandem prevented deceleration of actin polymerization by the FH2-DAD fragment. Error bars represent standard error for at least 3 replicate experiments. * indicates a statistically significant difference ($p < 0.05$) and ‡ indicates an indistinguishable slope compared to actin only.
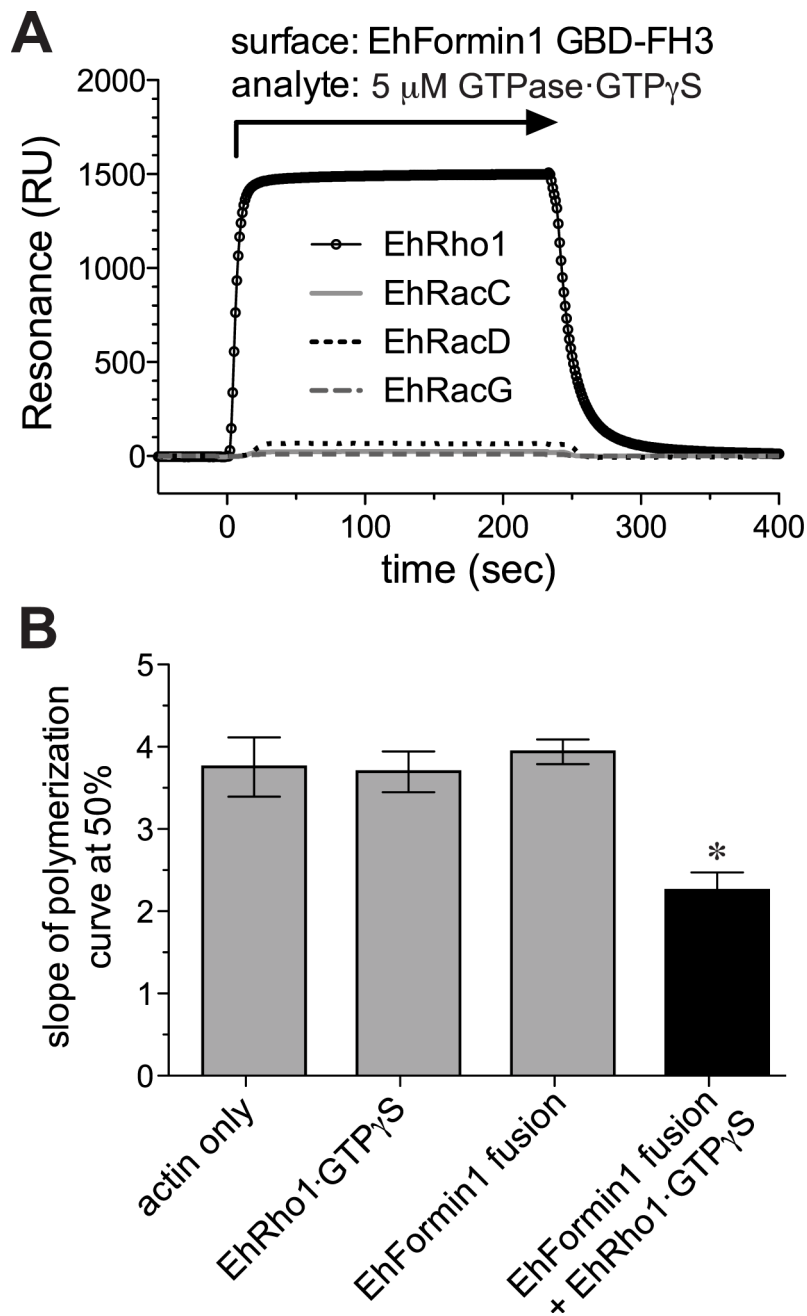
Figure 5.5. EhRho1 activates EhFormin1 through interaction with the GBD-FH3 domain tandem. A) The GBD-FH3 domain tandem (a.a. 69-418) of EhFormin1, immobilized on a nickel-NTA biosensor surface, selectively engaged EhRho1·GTPγS to the exclusion of multiple other Rho family GTPases from *E. histolytica*, as measured by surface plasmon resonance. **B)** The EhFormin1 fusion protein (25 μM) is apparently autoinhibited, having no measured effect on actin polymerization kinetics *in vitro*. While EhRho1·GTPγS alone (100 μM) did not perturb actin polymerization, it was capable of activating the EhFormin1 fusion, resulting in deceleration of actin polymerization comparable to that of the FH2-DAD fragment alone (Figure 5.2B, C). Error bars represent standard error for at least 3 replicate experiments. * indicates statistical significance ($p < 0.05$) compared to actin only.

A)

B)

	k_{obs} (s^{-1})	k_{on} (s^{-1}M^{-1})*	k_{off} (s^{-1})	K$_D$ kinetic (μM)**	K$_D$ equilibrium (μM)
EhRho1·GTPγS/ EhFormin1fusion	0.188 [± 0.010]	17,800	0.0994 [± 0.0016]	2.81	2.87 [± 1.50]

*derived from $k_{on} = (k_{obs} - k_{off})$ / (analyte concentration)
**derived from K$_D$ = k_{off} / k_{on}
***values in brackets represent 95% C.I.

Figure 5.6. EhRho1·GTPγS/EhFormin1 fusion binding parameters. A) As measured by SPR, EhRho1·GTPγS bound the engineered fusion of the GBD-FH3 and FH2-DAD fragments with low μM affinity. For details of the EhFormin1 fusion and other constructs used in this study, see Figure 5.1. **B)** Kinetic and equilibrium binding parameters for the EhRho1·GTPγS/EhFormin1 fusion protein interaction were derived from SPR experiments.

A

β1 P-loop α1 switch 1 β2 β3

```
            10          20          30          40          50          60          70
EhRho1  MLAFSDMNTGAGKIENGKKALKIVVVGDGAVGKTCLLLAFSKGEIPTAYVPTVFENFSHVMKYKNEEFILHLWD
EhRacC  -------------MSEKPTSIKLVVVGDGAVGKTCLLICYTTNEFPKDYIPTVFDNYVVSLTAGTRQIQLALWD
EhRacD  -------------AAPTDAKSVKLVVVGDGSVGKTCLLICYTTNEFPKDYVPTVFDNYMAPMTVDGEPINLGLWD
EhRacG  --------------MR----PVKLVIVGDGAVGKTCMLISYTTNAFPNEYIPTVFENYNSSLVVDDVKINLGLWD
```

switch 2 α2 β4 α3 β5

```
            80          90         100         110         120         130         140
EhRho1  TAGGQEEYDRLRPLSYADSDVVLLCFAVNNRTSFDNISTKWEPEIKHYIDTAKTVLVGLKVDLRKD---------
EhRacC  TAGGQEEYDQLRPLSYSSASIFLICFSVTSSVSYDNVITKWHPEVIHFAPKVPIILVGTKLDTRNDPAIVKRLTE
EhRacD  TAGGQEDSEQLRPLSYPNTDLFLLCFSVISRTSFNNISSKWLPEIKHYEPKCKMMVVGTNTDCRNDEAMIRKLAD
EhRacG  TAGGQEDYDRLRPLSYPSTDVFLVCFSVIAPASYENVEGKWKPEIDQHCPNVPIILVGTKIDIRDDPEQVKRLAE
```

α4 β6 α5

```
           150         160         170         180         190         200         210
EhRho1  -GSDDVTKQEGDDLCQKLGCVAYIEASSVAKIGLNEVFEKSVDCIFSNKPVPKASVTTQAKSQEVCFFIFLLVV
EhRacC  QGMTVINTAKGEELKNRIKAVKYIECSAKTSENLKTVFDEAVKTVLMNKPQQ---------RSKCALL------
EhRacD  ENQKPITTEEGEKLAKDIKAICYMECSALTRSGLNQVFDEAIHIVLNKNQSS------KKSSKKCALL------
EhRacG  KNIVPIQPPQGDELAKKIGAVKYIECSALTQANLKLVFEEAVRAVLAKAAKEPTGKKEKGGKKGCSLF------
```

B

α1 α2 α3 α4

```
          70          80          90         100         110         120         130
EhFormin1  MPPEEVVQ----QKFAIVAKEMKIDNPELITIPNQWKLVQEYE-KKQKKDIRI-------QLNAQKTGNWR--NAIT
mDia1      DPTAQSLQDISDEQVLVLFEQMLVDMN----------LNEEKQQPLREKDIVIKREMVSQYLHTSKAGMNQKESSRS
```
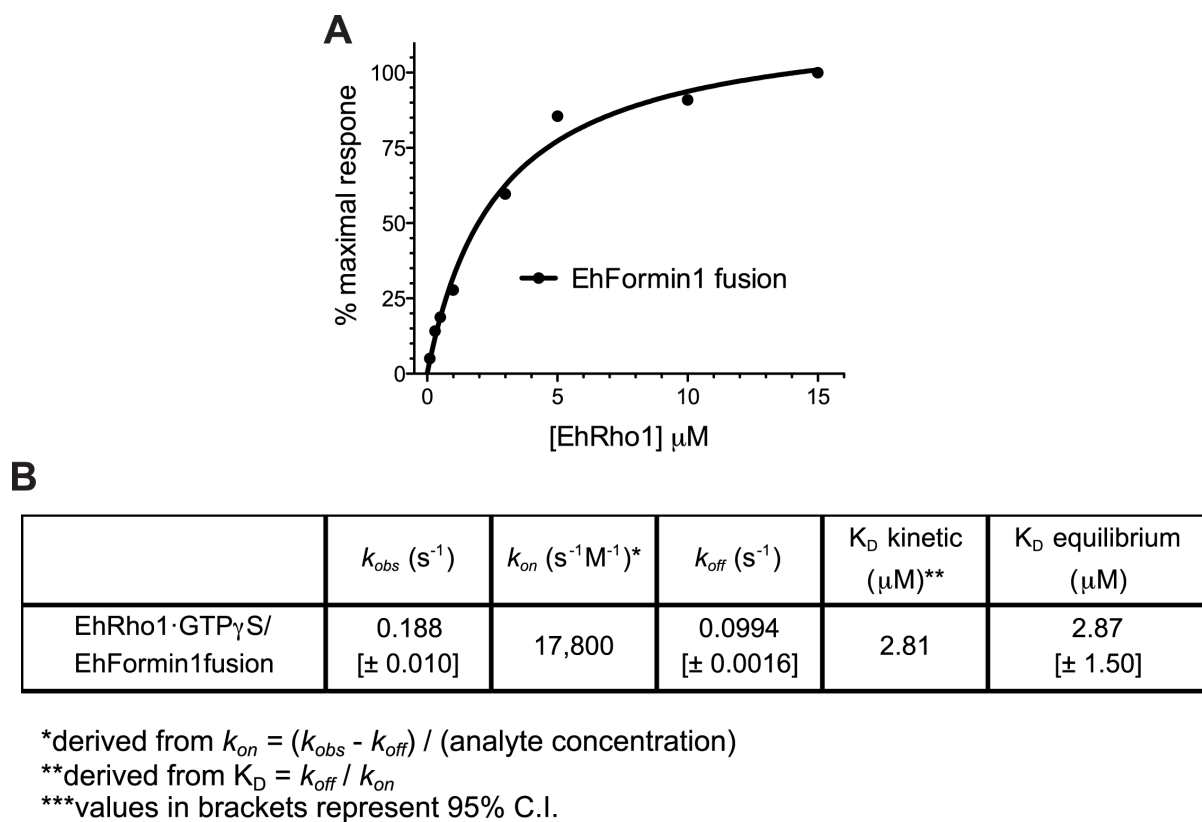
α5 α6 α7 α8 α9

```
          140         150         160         170         180         190         200
EhFormin1  DPKYLADLLKTRDDMDLLN---EMVVVFRSSSVSFIKTFVSVGGLANLMAIYKKKIEAENSNTAIDEERKCCEVLRY
mDia1      AMMYIQELRSGLRDMHLLSCLESLRVSLNNNPVSWVQTFGA-EGLASLLDILKRLHDEKEETSGNYDSRNQHEIIRC
```

α10 α11 α12 α13

```
          210         220         230         240         250         260         270         280
EhFormin1  VFAEE---DATVALIEIDGGVELLLKGMNSKRITPDNQLDILLEITLTSSMVEHPSQEGLYLGGDVCVMNAFSNLVSE
mDia1      LKAFMNNKFGIKTMLETEEGILLLVRAMDPA--VPNMMIDAA---KLLSALCILPQPEDMN----ERVLEAMTER-AE
```

α14 α15 α16 α17 α18

```
          290         300         310         320         330         340         350
EhFormin1  GVDMKKFLSFFSLFSKSKSEKFKHASLVLINNLID-QPELEHRMDVRNSFIEIGLVNELENMKNTEWMKIDKIKDSIN
mDia1      MDEVERFQPLLDGLKSGTSIALKVGCLQLINALTPAEELDFRVHIRSELMRLGLHQVLQELREIEN---EDMKVQLC
```

```
          360         370         380         390         400         410         420         430
EhFormin1  DFFDSWEEDKKEVESRFDDLKQVVD-F-ESTKSLNNYVTEQMDKFECNDILTNVHKEILFFAKVKDDNLLLIKKIVIL
mDia1      VFDEQGDEDFFDLKGRLDDIRMEMDDFGEVFQIILNTVKDSKAEPHFLSILQH-------LLLVRNDYEARPQYYKLI
```

```
          440         450         460
EhFormin1  TGIIRQISLQVTKNNNEKSNKKKIESGDNEEER
mDia1      EECVSQIVLH--KNGTDPDFKC-----RHLQID
```

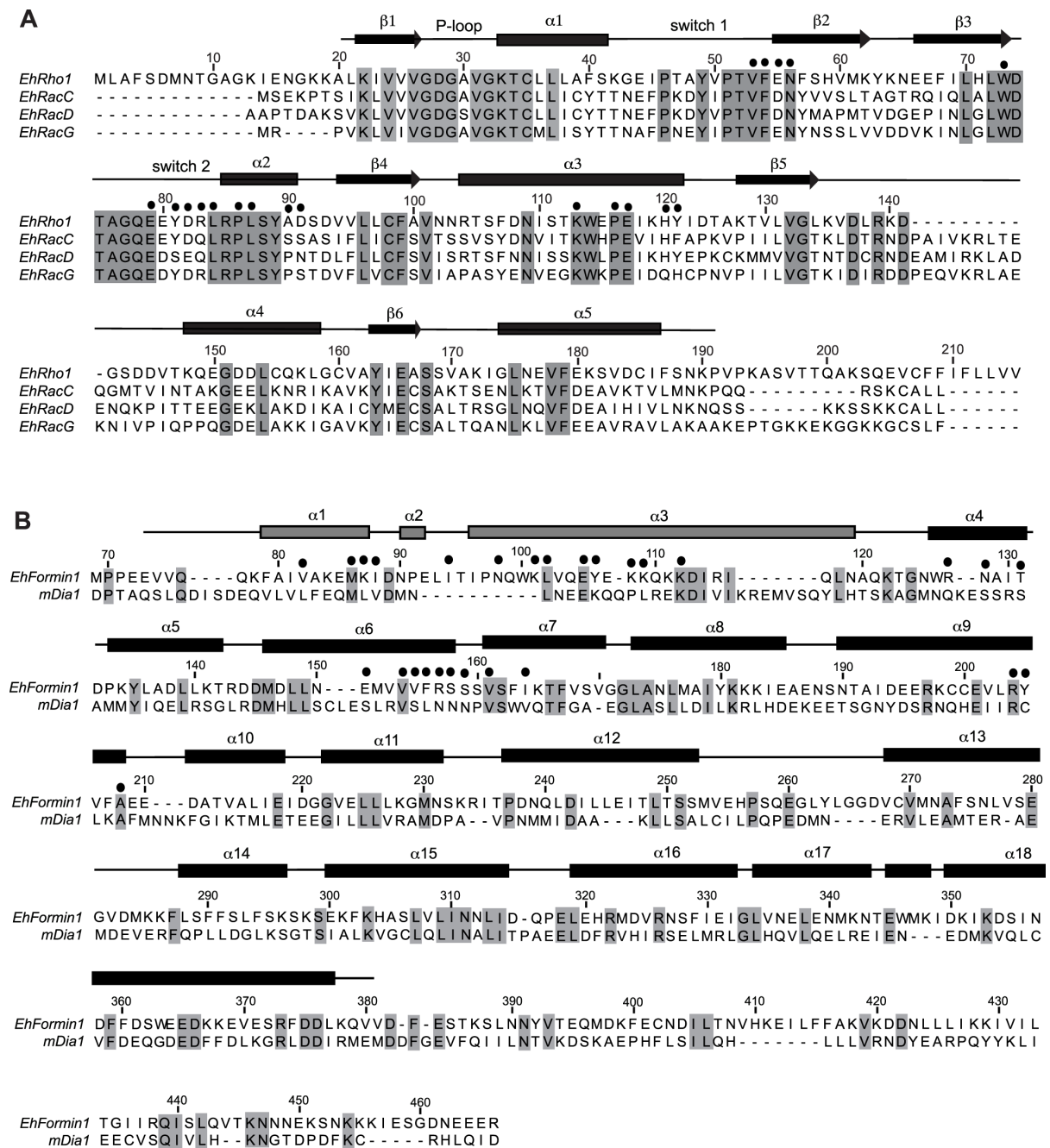**Figure 5.7. Sequence comparisons of Rho family GTPases and Diaphanous-related formins. A)** The protein sequence of EhRho1 is aligned with three additional *E. histolytica* Rho GTPases that do not engage EhFormin1 (Figure 5.5A). Secondary structure from the EhRho1·GTPγS/EhFormin1 crystal structure (PDB id 4DVG; this paper) is diagrammed above the primary sequence. Black dots indicate residues of EhRho1·GTPγS within 1 Å of EhFormin1 in the complex. **B)** The protein sequence of EhFormin1 GBD-FH3 domain tandem is aligned with its closest mammalian homolog, mDia1. Secondary structure from the EhRho1·GTPγS/EhFormin1 crystal structure (PDB id 4DVG; this paper) is diagrammed

above the primary sequence. Black dots indicate residues in EhFormin1 within 1 Å of EhRho1·GTPγS in the complex. All alignments were performed with ClustalW2.
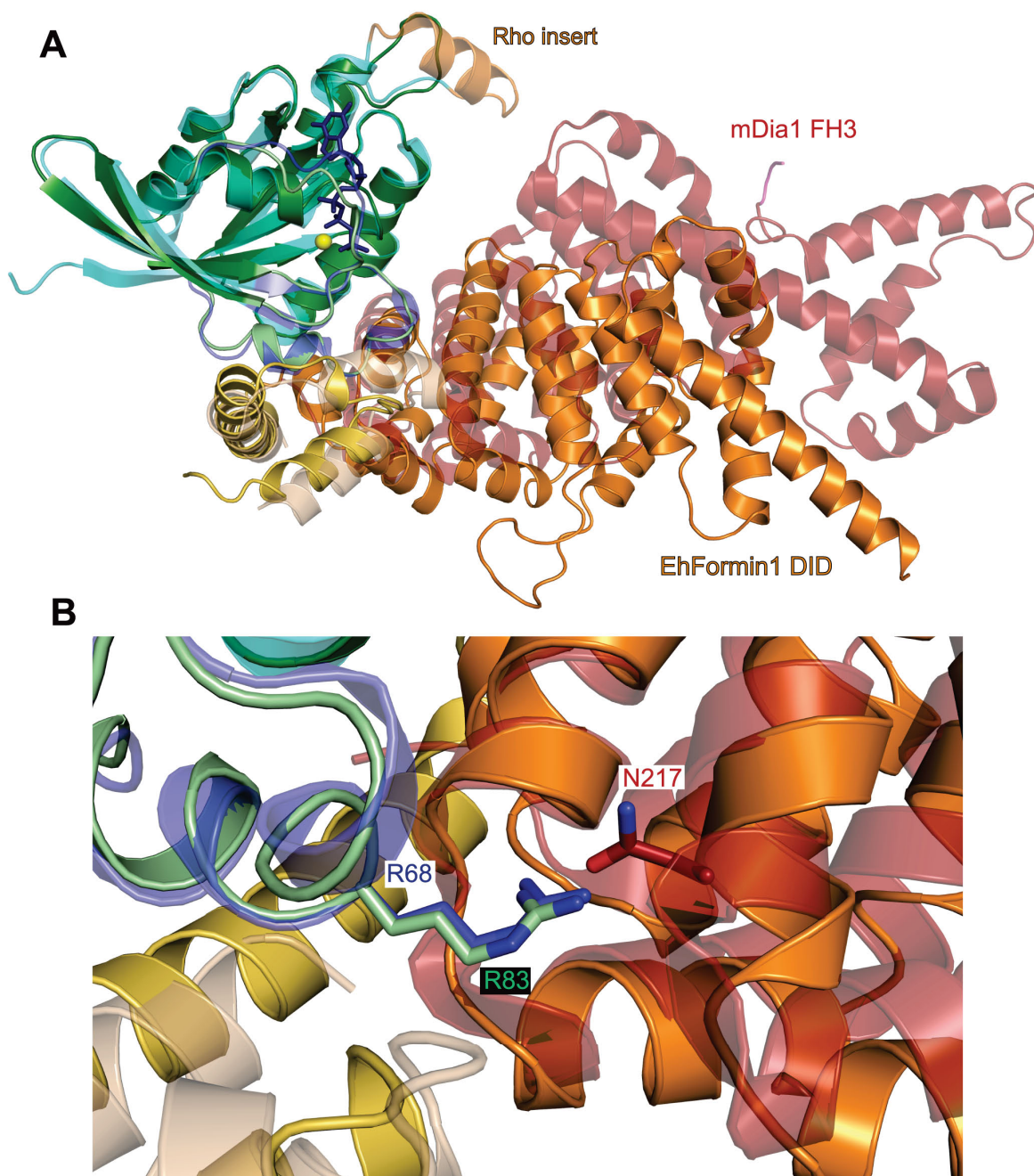
Figure 5.8. The crystal structure of EhRho1·GTPγS bound to the GBD-FH3 tandem of EhFormin1. A) EhRho1 (*dark green*) in its activated state bound to GTPγS (*blue sticks*) engages EhFormin1 primarily through its mobile switch regions (*light green*) and its α3 helix. The GTPase binding domain (GBD) of EhFormin1 is shown in *yellow* and the Armadillo repeats of the DID domain in *orange*. Magnesium is shown as a yellow sphere. **B)** The homologous mammalian complex between RhoC·GppNHp (*cyan*) and mDia1 (*wheat* and *red*) is posed in a similar configuration [20] (PDB id 1Z2C). EhRho1 lacks a signature Rho insert helix (highlighted in *orange* on RhoC) and the EhFormin1 DID domain Armadillo repeats are rotated away from EhRho1 relative to this mammalian complex. The EhFormin1 GBD-FH3 tandem construct used in this study also lacks the dimerization domain portion of the FH3 domain, as shown in the RhoC/mDia1 complex.

Figure 5.9. Structural comparison of EhRho1/EhFormin1 with mammalian RhoC/mDia1. A) The EhRho1·GTPγS/EhFormin1 complex is superimposed with RhoC·GppNHp/mDia1 (PDB id 1Z2C) using the Cα atoms of the respective Rho GTPases, and colored as in Figure 5.8. The Rho GTPase and GBD domains are similar across species, but the DID domain exhibits a different conformation, with an ~40° relative rotation of C-terminal Armadillo repeats. **B)** The switch 2 arginines, Arg-83 in EhRho1 and Arg-68 in human RhoC, adopt nearly identical orientations, inserting between the GBD (*yellow/wheat*) and DID domain (*orange/red*). However, Arg-68 of RhoC forms hydrogen bonds with Asn-217 of mDia1 as well as backbone carbonyl groups, while Arg-83 of EhRho1 exclusively contacts main chain carbonyl groups.
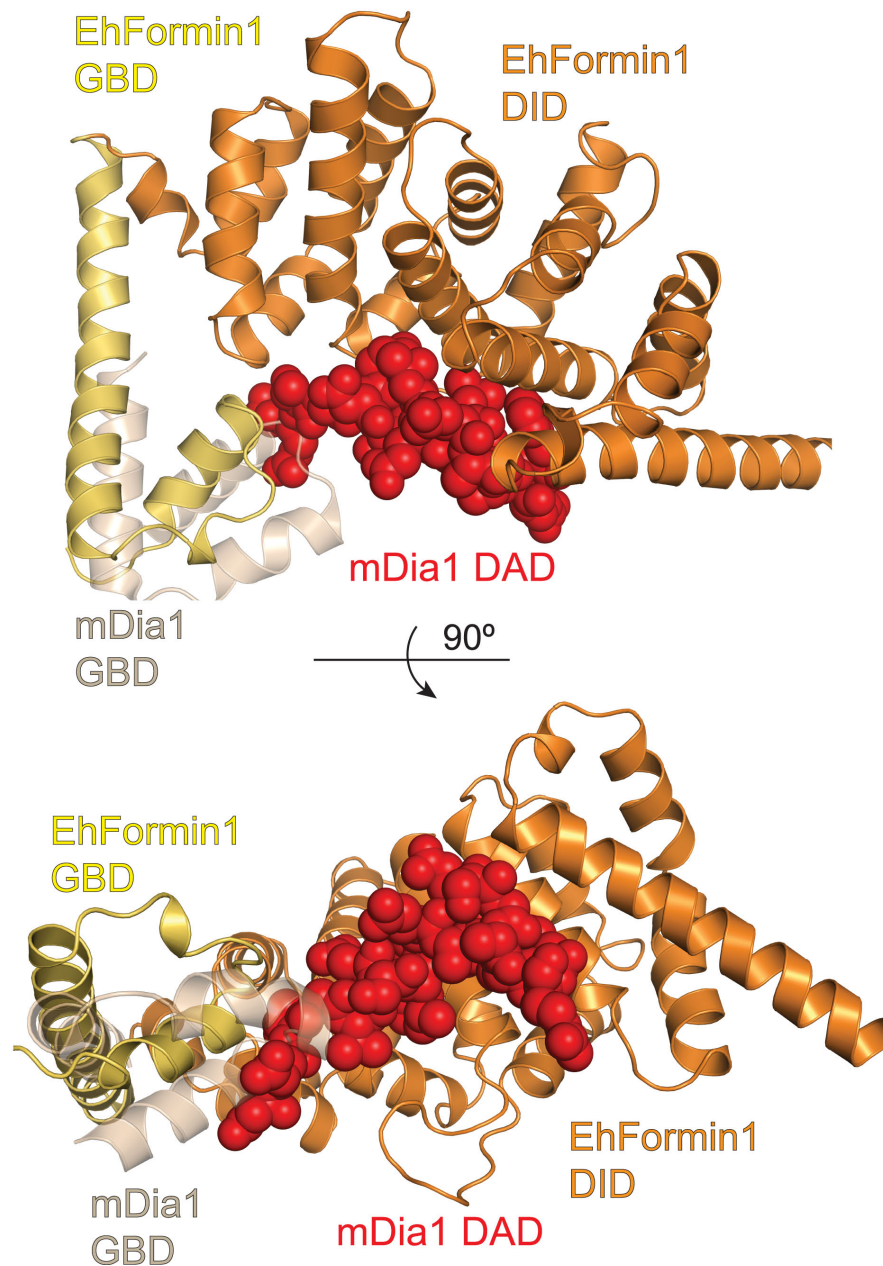
Figure 5.10. The GBD of EhFormin1 does not impinge on the putative DAD motif-binding region of the DID, in contrast to the known structures of mDia1. The GBD-DID domain tandems derived from the structures of EhRho1·GTPγS/EhFormin1 (PDB id 4DVG) and RhoC·GppNHp/mDia1 (PDB id 1Z2C) were superimposed with the mDia1 DID/DAD motif complex (PDB id 2F31) based on their DID domains. EhFormin1 is shown in *yellow* and *orange*, and the GBD motif from mDia1 in the RhoC/mDia1 complex in *wheat*. The DAD motif of mDia1 (*red spheres*) is docked on the putative DAD-binding site of the EhFormin1 DID domain. As shown previously [19], the GBD domain of mDia1 overlaps with the DAD-binding site, suggesting that the GBD participates in ejecting the DAD motif upon Rho-mediated activation. The GBD of EhFormin1 displays a shorter α2 helix and does not clash with the docked DAD motif, suggesting that an alternative mechanism of activation is operative with the *E. histolytica* DRF homolog EhFormin1.
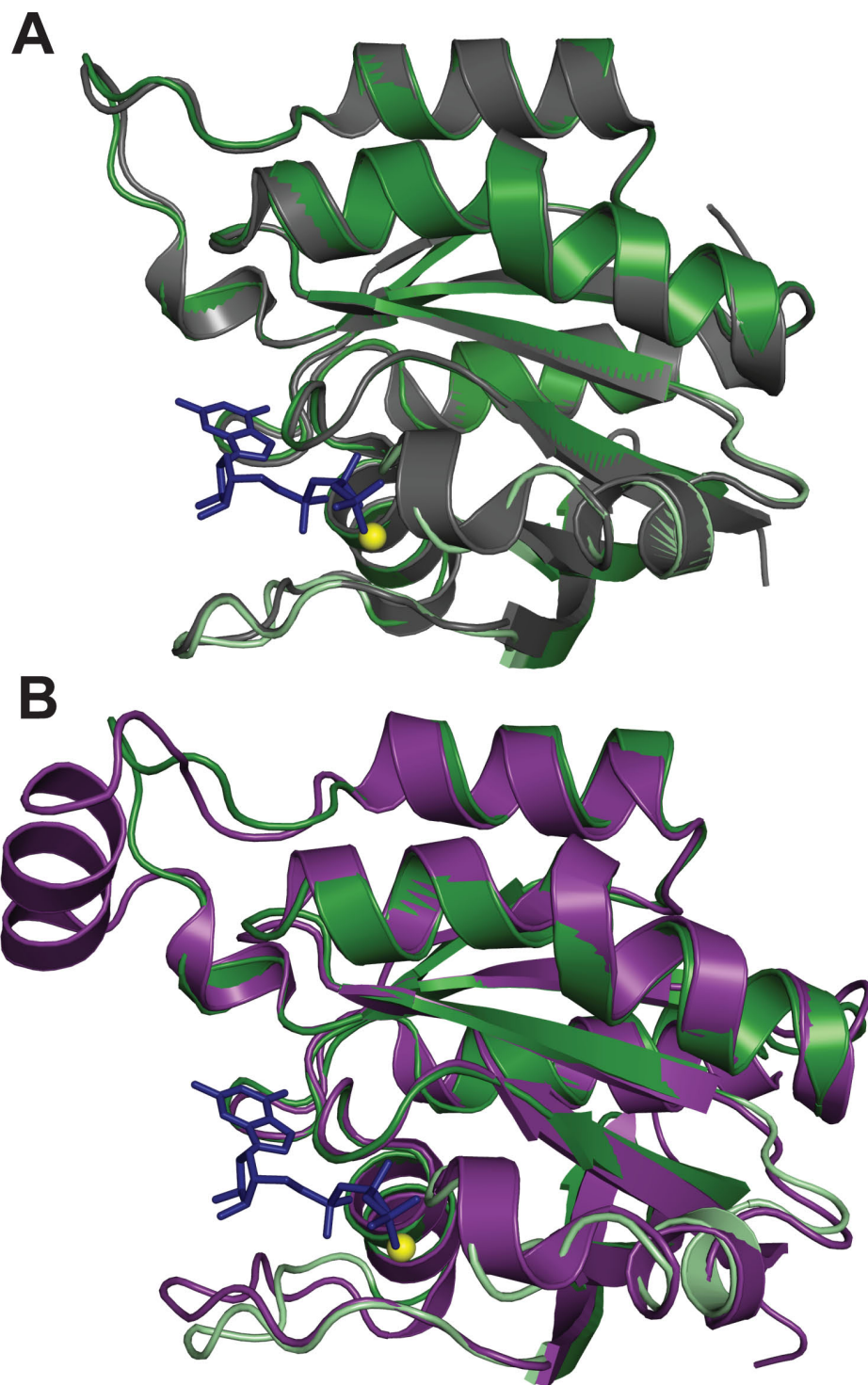
Figure 5.11. The conformation of EhRho1 in the EhRho1·GTPγS/EhFormin1 complex resembles that of free EhRho1·GTPγS and of human RhoC. A) EhRho1 from the EhRho1·GTPγS/EhFormin1 complex of this study (*green*) is superimposed with free EhRho1·GTPγS [7] (gray, PDB id 3REG); the Cα r.m.s.d. is 0.4 Å. **B)** EhRho1 from this study is superimposed with RhoC from the RhoC/mDia1 complex (*purple*, PDB id 1Z2C); the Cα r.m.s.d. is 1.5 Å.
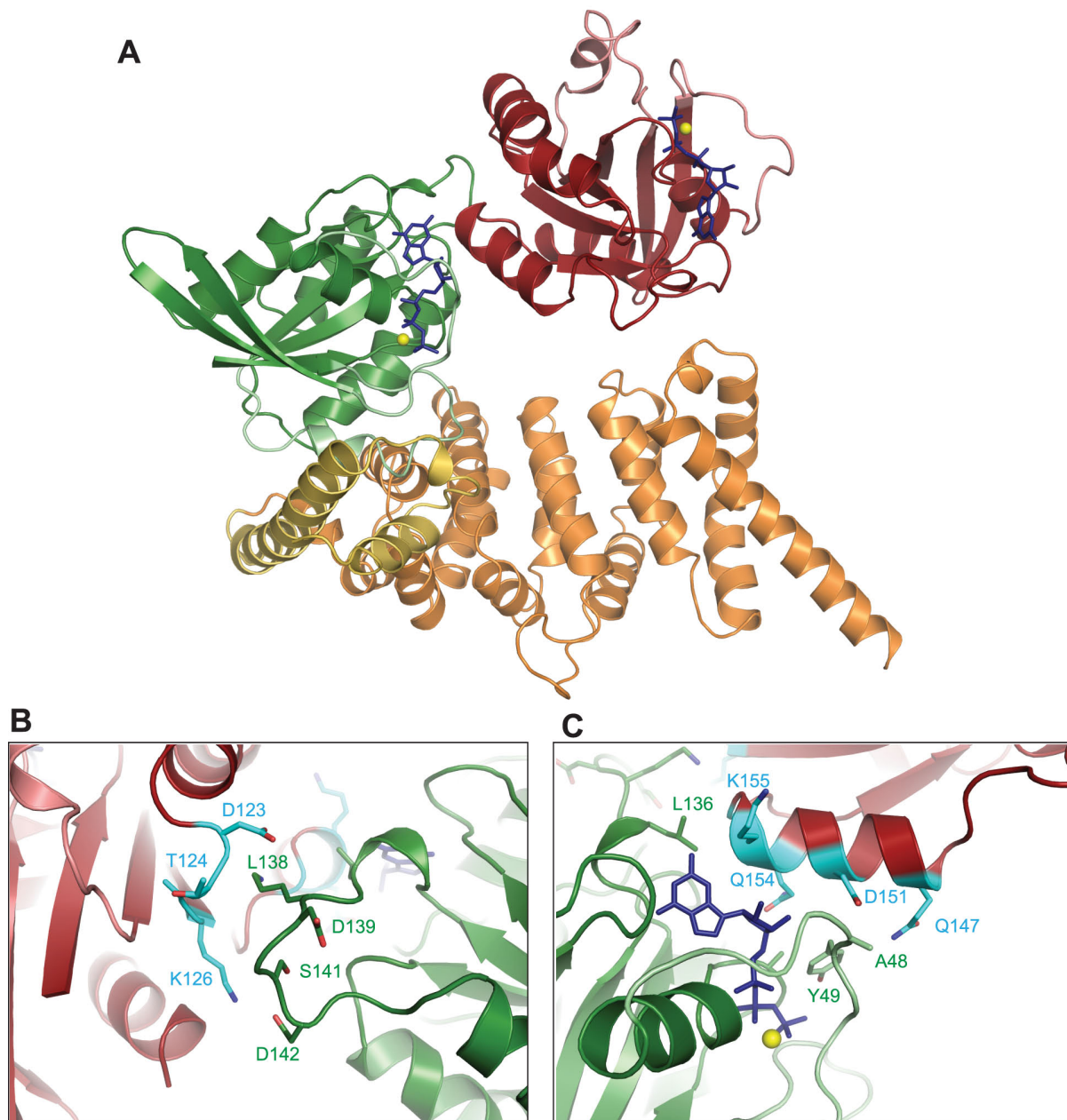
Figure 5.12. Crystal contacts between EhRho1 molecules. A) The structure of the EhRho1·GTPγS/EhFormin1 complex is shown as in Figure 5.8, with an EhRho1 GTPase from the adjacent asymmetric unit modeled in red. The neighboring EhRho1 does not form significant contacts with the DID domain. **B, C)** Two areas of contact between EhRho1 molecules in adjacent asymmetric units are shown in greater detail. Crystal contacts may influence the conformation of EhRho1 switch 1 (panel A) and/or the β5-α4 loop (panel B) in the complex.
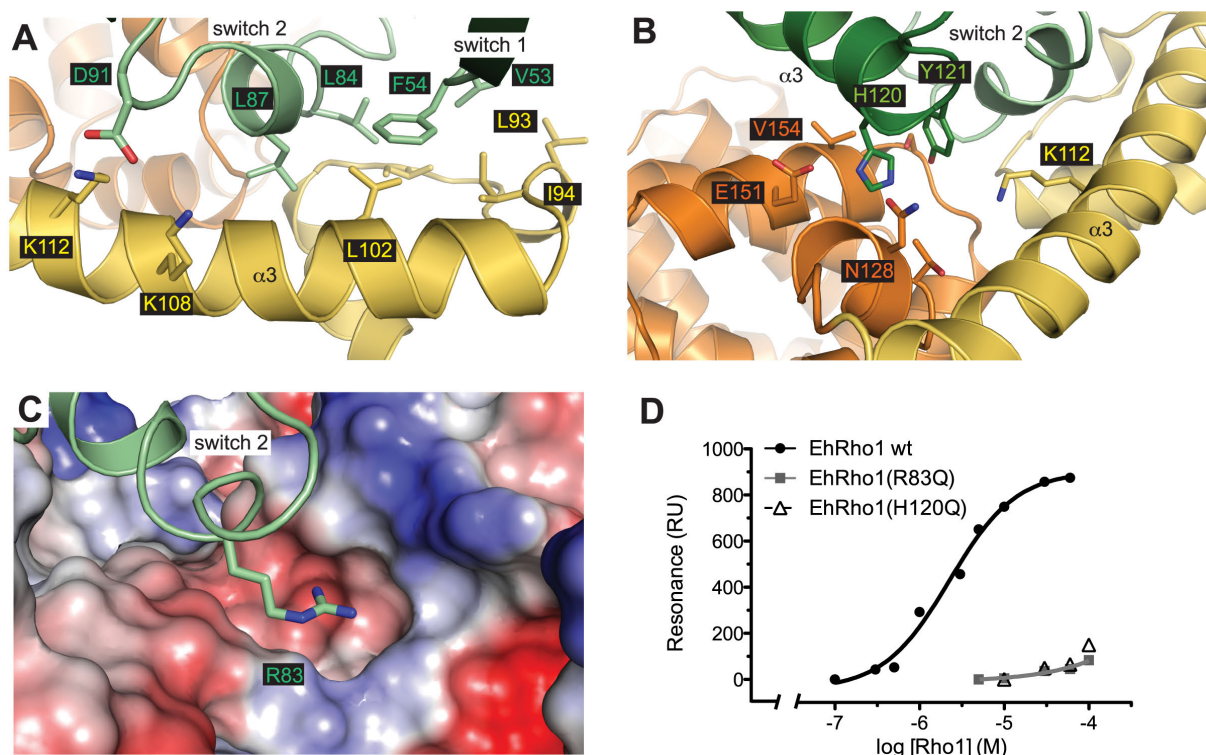
Figure 5.13. Structural determinants of EhRho1/EhFormin1 binding specificity.
A) The EhRho1·GTPγS/EhFormin1 interface is dominated by a hydrophobic surface burial involving the switch regions of EhRho1 (*green*) and the EhFormin1 GBD (*yellow*). Key hydrophobic residues are shown in sticks. EhRho1 Asp91 is also in position to form an ionic interaction with Lys108 or Lys112 on EhFormin1. **B)** The α3 helix of EhRho1 (*green*) also contributes to the binding interface, with His120 and Tyr121 inserted between the GBD α3 helix (*yellow*) and the N-terminal portion of the DID domain (*orange*). The imidazole ring of His120 is oriented for hydrogen bonding with Glu151 of EhFormin1. **C)** EhRho1 switch 2 residue Arg83 inserts into a groove between the GBD and DID domains of EhFormin1 (shown as an electrostatic surface). The guanidinium group of Arg83 resides near an area of relative negative charge (red surface) and is within hydrogen bonding distance of multiple backbone carbonyl groups (see Figure 5.14). **D)** Wild type EhRho1 binds the EhFormin1 GBD-FH3 tandem, as measured by surface plasmon resonance. Mutation of either Arg83 to the corresponding glutamine in EhRacC and EhRacD or His120 to an EhRacG-like glutamine resulted in a >100-fold affinity reduction.
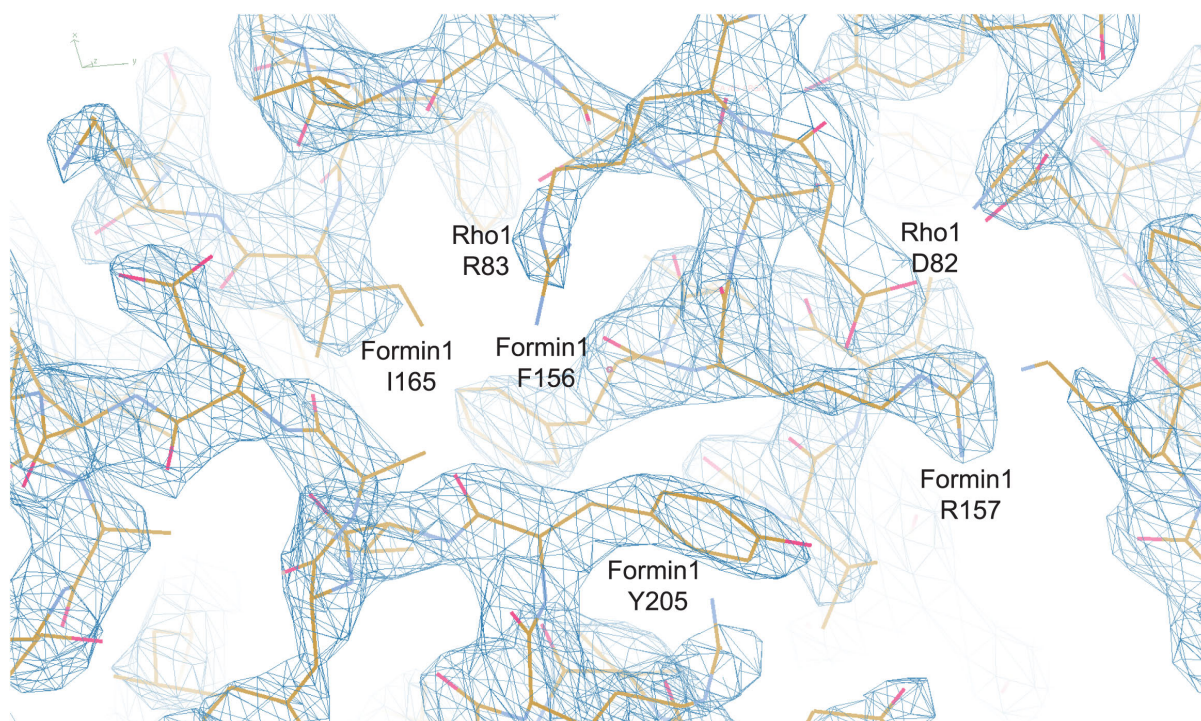
223

Figure 5.14. A representative electron density map of the region surrounding EhRho1 Arg83. A $2F_o$-F_c electron density map was modeled using Coot [33]. Arg83 is within potential hydrogen bonding distance of the carbonyl groups from EhFormin1 Phe156 and Arg157. The electron density map was contoured to σ = 2.2.

Table 5.1. Data collection and refinement statistics for the EhRho1·GTPγS/EhFormin1 complex.

	EhRho1·GTPγS/EhFormin1
PDB accession code	4DVG
Data collection	
Space group	$P6_1$
Cell dimensions	
a, b, c (Å)	138.6, 138.6, 57.8
α, β, γ (°)	90, 90, 120
	Peak
Wavelength (Å)	0.97954
Resolution (Å)	40.0 - 2.6 (2.63 – 2.60)*
R_{merge} (%)	8.5 (58.0)**
$I / \sigma I$	15.2 (2.1)
Wilson B-factor	65.0
Completeness (%)	98.4 (86.0)
Redundancy	9.6 (6.4)
Refinement	
Resolution (Å)	35.7 - 2.6 (2.7 – 2.6)
No. reflections	19211 (2420)
Cutoff criterion	$F_{obs}/\sigma F_{obs} > 0$
R_{work} / R_{free} (%)	21.5 / 25.1 (33.4 / 37.2)
No. atoms	
Protein	3747
Ligand/ion	33
Water	4
B-factors (Å^2)	
Protein	79.5
Ligand/ion	57.4
Water	36.1
R.m.s deviations	
Bond lengths (Å)	0.010
Bond angles (°)	1.233

*Values in parentheses are for highest-resolution shell.
**All data were collected from a single crystal.

5.6   REFERENCES

1.   Haque, R., C.D. Huston, M. Hughes, E. Houpt, and W.A. Petri, Jr., *Amebiasis.* N Engl J Med, 2003. **348**(16): p. 1565-73.

2.   Ravdin, J.I., *Amoebiasis.* Vol. 2. 2000, London: Imperial College Press.

3.   Voigt, H. and N. Guillen, *New insights into the role of the cytoskeleton in phagocytosis of Entamoeba histolytica.* Cell Microbiol, 1999. **1**(3): p. 195-203.

4.   Guillen, N., *Role of signalling and cytoskeletal rearrangements in the pathogenesis of Entamoeba histolytica.* Trends Microbiol, 1996. **4**(5): p. 191-7.

5.   Meza, I., P. Talamas-Rohana, and M.A. Vargas, *The cytoskeleton of Entamoeba histolytica: structure, function, and regulation by signaling pathways.* Arch Med Res, 2006. **37**(2): p. 234-43.

6.   Lohia, A. and J. Samuelson, *Heterogeneity of Entamoeba histolytica rac genes encoding p21rac homologues.* Gene, 1996. **173**(2): p. 205-8.

7.   Bosch, D.E., E.S. Wittchen, C. Qiu, K. Burridge, and D.P. Siderovski, *Unique structural and nucleotide exchange features of the Rho1 GTPase of Entamoeba histolytica.* J Biol Chem, 2011. **286**(45): p. 39236-46.

8.   Ridley, A.J. and A. Hall, *The small GTP-binding protein rho regulates the assembly of focal adhesions and actin stress fibers in response to growth factors.* Cell, 1992. **70**(3): p. 389-99.

9.   Wennerberg, K. and C.J. Der, *Rho-family GTPases: it's not only Rac and Rho (and I like it).* J Cell Sci, 2004. **117**(Pt 8): p. 1301-12.

10.   Ghosh, S.K. and J. Samuelson, *Involvement of p21racA, phosphoinositide 3-kinase, and vacuolar ATPase in phagocytosis of bacteria and erythrocytes by Entamoeba histolytica: suggestive evidence for coincidental evolution of amebic invasiveness.* Infect Immun, 1997. **65**(10): p. 4243-9.

11.   Guillen, N., P. Boquet, and P. Sansonetti, *The small GTP-binding protein RacG regulates uroid formation in the protozoan parasite Entamoeba histolytica.* J Cell Sci, 1998. **111 (Pt 12)**: p. 1729-39.

12.   Chesarone, M.A., A.G. DuPage, and B.L. Goode, *Unleashing formins to remodel the actin and microtubule cytoskeletons.* Nat Rev Mol Cell Biol, 2010. **11**(1): p. 62-74.

13.   Higgs, H.N., *Formin proteins: a domain-based approach.* Trends Biochem Sci, 2005. **30**(6): p. 342-53.

14. Lammers, M., S. Meyer, D. Kuhlmann, and A. Wittinghofer, *Specificity of interactions between mDia isoforms and Rho proteins.* J Biol Chem, 2008. **283**(50): p. 35236-46.

15. Schonichen, A. and M. Geyer, *Fifteen formins for an actin filament: a molecular view on the regulation of human formins.* Biochim Biophys Acta, 2010. **1803**(2): p. 152-63.

16. Nezami, A., F. Poy, A. Toms, W. Zheng, and M.J. Eck, *Crystal structure of a complex between amino and carboxy terminal fragments of mDia1: insights into autoinhibition of diaphanous-related formins.* PLoS One, 2010. **5**(9).

17. Otomo, T., D.R. Tomchick, C. Otomo, M. Machius, and M.K. Rosen, *Crystal structure of the Formin mDia1 in autoinhibited conformation.* PLoS One, 2010. **5**(9).

18. Lammers, M., R. Rose, A. Scrima, and A. Wittinghofer, *The regulation of mDia1 by autoinhibition and its release by Rho*GTP.* EMBO J, 2005. **24**(23): p. 4176-87.

19. Nezami, A.G., F. Poy, and M.J. Eck, *Structure of the autoinhibitory switch in formin mDia1.* Structure, 2006. **14**(2): p. 257-63.

20. Rose, R., et al., *Structural and mechanistic insights into the interaction between Rho and mammalian Dia.* Nature, 2005. **435**(7041): p. 513-8.

21. Majumder, S. and A. Lohia, *Entamoeba histolytica encodes unique formins, a subset of which regulates DNA content and cell division.* Infect Immun, 2008. **76**(6): p. 2368-78.

22. Ali, I.K., et al., *Proteomic Analysis of the Cyst Stage of Entamoeba histolytica.* PLoS Negl Trop Dis, 2012. **6**(5): p. e1643.

23. Stols, L., et al., *A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site.* Protein Expr Purif, 2002. **25**(1): p. 8-15.

24. Ho, S.N., H.D. Hunt, R.M. Horton, J.K. Pullen, and L.R. Pease, *Site-directed mutagenesis by overlap extension using the polymerase chain reaction.* Gene, 1989. **77**(1): p. 51-9.

25. Spudich, J.A. and S. Watt, *The regulation of rabbit skeletal muscle contraction. I. Biochemical studies of the interaction of the tropomyosin-troponin complex with actin and the proteolytic fragments of myosin.* J Biol Chem, 1971. **246**(15): p. 4866-71.

26. Cooper, J.A., S.B. Walker, and T.D. Pollard, *Pyrene actin: documentation of the validity of a sensitive assay for actin polymerization.* J Muscle Res Cell Motil, 1983. **4**(2): p. 253-62.

27.   Srivastava, J. and D. Barber, *Actin co-sedimentation assay; for the analysis of protein binding to F-actin.* J Vis Exp, 2008(13).

28.   Harris, E.S. and H.N. Higgs, *Biochemical analysis of mammalian formin effects on actin dynamics.* Methods Enzymol, 2006. **406**: p. 190-214.

29.   Kimple, A.J., R.E. Muller, D.P. Siderovski, and F.S. Willard, *A capture coupling method for the covalent immobilization of hexahistidine tagged proteins for surface plasmon resonance.* Methods Mol Biol, 2010. **627**: p. 91-100.

30.   Luft, J.R. and G.T. DeTitta, *A method to produce microseed stock for use in the crystallization of biological macromolecules.* Acta Crystallogr D Biol Crystallogr, 1999. **55**(Pt 5): p. 988-93.

31.   Otwinowski, Z.a.W.M., *Processing of X-ray Diffraction Data Collected in Oscillation Mode*, in *Methods in Enzymology*. 1997, Academic Press: New York. p. 307-326.

32.   Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.

33.   Emsley, P., B. Lohkamp, W.G. Scott, and K. Cowtan, *Features and development of Coot.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.

34.   Painter, J. and E.A. Merritt, *Optimal description of a protein structure in terms of multiple groups undergoing TLS motion.* Acta Crystallogr D Biol Crystallogr, 2006. **62**(Pt 4): p. 439-50.

35.   Chen, V.B., et al., *MolProbity: all-atom structure validation for macromolecular crystallography.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 1): p. 12-21.

36.   Otomo, T., et al., *Structural basis of actin filament nucleation and processive capping by a formin homology 2 domain.* Nature, 2005. **433**(7025): p. 488-94.

37.   Shemesh, T. and M.M. Kozlov, *Actin polymerization upon processive capping by formin: a model for slowing and acceleration.* Biophys J, 2007. **92**(5): p. 1512-21.

38.   Kovar, D.R., J.R. Kuhn, A.L. Tichy, and T.D. Pollard, *The fission yeast cytokinesis formin Cdc12p is a barbed end actin filament capping protein gated by profilin.* J Cell Biol, 2003. **161**(5): p. 875-87.

39.   Aspenstrom, P., *Formin-binding proteins: modulators of formin-dependent actin polymerization.* Biochim Biophys Acta, 2010. **1803**(2): p. 174-82.

40.   Kursula, P., et al., *High-resolution structural analysis of mammalian profilin 2a complex formation with two physiological ligands: the formin homology 1 domain of mDia1 and the proline-rich domain of VASP.* J Mol Biol, 2008. **375**(1): p. 270-90.

41.   Marquay Markiewicz, J., et al., *A proteomic and cellular analysis of uropods in the pathogen Entamoeba histolytica.* PLoS Negl Trop Dis, 2011. **5**(4): p. e1002.

CHAPTER 6

STRUCTURAL DETERMINANTS OF UBIQUITIN CONJUGATION

IN Entamoeba histolytica[1]

6.1   OVERVIEW

Ubiquitination is important for numerous cellular processes in most eukaryotic organisms, including cellular proliferation, development, and protein turnover by the proteasome. The intestinal parasite *Entamoeba histolytica* harbors an extensive ubiquitin-proteasome system. Proteasome inhibitors are known to impair parasite proliferation and encystation, suggesting the ubiquitin-proteasome pathway as a viable therapeutic target. However, no functional studies of the *E. histolytica* ubiquitination enzymes have yet emerged. Here, we have cloned and characterized multiple *E. histolytica* ubiquitination components, spanning ubiquitin and its activating (E1), conjugating (E2), and ligating (E3) enzymes. Crystal structures of EhUbiquitin reveal a clustering of unique residues on the α1 helix surface, including an eighth surface lysine not found in other organisms, that may allow for a unique polyubiquitin linkage in *E. histolytica*. EhUbiquitin is activated by and forms a thioester bond with EhUba1 (E1) *in vitro*, in an ATP- and magnesium-dependent fashion. EhUba1 exhibits a greater maximal initial velocity of pyrophosphate:ATP exchange than its human homolog,

[1] Bosch, D.E. and Siderovski, D.P. (2012) Structural determinants of ubiquitin conjugation in *Entamoeba histolytica*. *J. Biol. Chem.* **288**(4):2290-302.

suggesting different kinetics of ubiquitin activation in *E. histolytica*. EhUba1 engages the E2 enzyme EhUbc5 through its ubiquitin-fold domain to transfer the EhUbiquitin thioester. However, EhUbc5 has a >10-fold preference for EhUba1~Ub compared to unconjugated EhUba1. A crystal structure of EhUbc5 allowed prediction of a non-covalent "backside" interaction with EhUbiquitin and with E3 enzymes. EhUbc5 selectively engages EhRING1 (E3) to the exclusion of two HECT family E3 ligases, and mutagenesis indicates a conserved mode of E2/RING-E3 interaction in *E. histolytica*.

6.2   INTRODUCTION

Post-translational modification by ubiquitin and ubiquitin-like modifiers (ULMs) plays critical roles in the cellular biology of many eukaryotes [1]. Isopeptide bond-mediated attachment of monoubiquitin or polyubiquitin chains to specific target proteins regulates processes as diverse as cell cycle progression, development, and the immune response [2]. Polyubiquitination also serves as a primary signal for protein turnover, targeting proteins for degradation by the 26S proteasome [3]. ULM attachments are catalyzed by conserved enzyme cascades, with ATP-dependent ULM activation catalyzed by E1, conjugation by E2, and targeting of specific protein substrates by E3 ligases [3]. Ubiquitin activating enzymes (E1s) in eukaryotes serve to adenylate ubiquitin, consuming ATP and releasing pyrophosphate (PP_i). Subsequently, a thioester bond is formed between an E1 active site cysteine and the dual-glycine C-terminus of the activated ubiquitin, releasing AMP, and a second ubiquitin molecule is also adenylated [4]. The resulting E1~Ub thioester complex engages a ULM conjugating enzyme (E2), with major contributions to the E1/E2 interface arising from the E1 ubiquitin-fold domain (UFD) [5]. The thioester bond between the

ubiquitin C-terminus and the E1 enzyme is transferred to a conserved cysteine residue on the E2, resulting in a covalent E2~Ub adduct [6]. E3 ubiquitin-ligating enzymes can be classified into two major groups, containing either homology to the E6AP C-terminus (HECT) or really interesting new gene related (RING) domains [3]. HECT family E3 enzymes possess a catalytic cysteine that accepts the ubiquitin thioester from bound E2~Ub prior to ubiquitin isopeptide bond formation with a lysine acceptor [7]. In the case of RING E3 enzymes, ubiquitin is directly transferred to the target protein from E2~Ub; the RING E3 enzyme serves primarily as an adaptor pairing ubiquitin-charged E2 enzymes with specific E3-bound substrates [3]. However, RING interactions with E2~Ub also prime the E2 for efficient ubiquitin conjugation to acceptor lysines [8]. Ubiquitin monomers can be added directly to surface lysines on substrate proteins, or can instead be linked into polyubiquitin chains, utilizing any of the seven lysines on the surface of ubiquitin, although some polyubiquitin linkages are more frequently utilized than others [9].

The single-celled intestinal parasite *Entamoeba histolytica* is the causative agent of amoebic colitis and systemic amoebiasis, with the latter characterized by cysts in liver, lungs, and brain [10]. Spread in an encysted form, *E. histolytica* infection is endemic in developing countries with poor barriers between sewage and drinking water [11]. Although *E. histolytica* has been the subject of research for more than 50 years, the relatively recent sequencing of its genome [12] affords the opportunity for further insight into cellular machinery that may be amenable to pharmacologic manipulation, such as the ubiquitin-proteasome pathway. The cloning and characterization of an *E. histolytica* ubiquitin gene (termed *ehubiquitin*) highlighted a surprising degree of sequence divergence from homologs in humans and lower organisms, given the remarkably high degree of conservation among other species [13].

Despite its relatively dissimilar sequence, EhUbiquitin complements deletion of the polyubiquitin gene *ubi4* in yeast, suggesting conserved functions in *E. histolytica* [14]. More recent bioinformatic analyses of the *E. histolytica* genome revealed an extensive family of putative ubiquitin activating, conjugating, and ligating enzymes, as well as parallel systems for other ubiquitin-like modifiers [15]. However, functional studies of this putative ubiquitination machinery have not yet emerged. Interestingly, treatment with proteasome inhibitors impairs growth of *E. histolytica* trophozoites and encystation in the related species *E. invadens*, suggesting that the ubiquitin-proteasome pathway may be a valuable therapeutic target [16]. Altered expression of ubiquitin-proteasomal genes has also been correlated with perturbations in virulence (*e.g.* [17]). Our own study of heterotrimeric G-protein signaling in *E. histolytica* demonstrated that markers of trophozoite virulence are enhanced or reduced upon overexpression of the Gα subunit, EhGα1, or a dominant negative EhGα1 mutant, respectively [18]. A transcriptome analysis by RNA-seq revealed differential expression of multiple ubiquitin-proteasome pathway-related genes upon expression of wild-type or mutant EhGα1, including the *ehubiquitin* gene itself (Table 6.1). In the present study, we sought to characterize, both structurally and biochemically, various components of the *E. histolytica* ubiquitination machinery, spanning ubiquitin and its interacting E1-E3 enzymes. We hypothesize that differences revealed between the *E. histolytica* components and well-studied mammalian homologs may elucidate potential means for specific targeting of ubiquitination within the parasitic amoeba.

6.3   EXPERIMENTAL PROCEDURES

6.3.1 Cloning and protein purification

Genomic DNA was isolated from the virulent HM-1:IMSS strain of *Entamoeba histolytica* using a DNeasy Blood and Tissue Kit (Qiagen). Open reading frames of *ehubiquitin* (AmoebaDB accession EHI_083410), *ehuba1* (EHI_020270), *ehubc5* (EHI_083560), *ehring1* (EHI_020100), *ehhect1* (EHI_011530), and *ehhect2* (EHI_124600) were PCR amplified from genomic DNA and subcloned as hexahistidine fusions into a pET vector-based ligation-independent cloning vector, pLIC-His, as described previously [19]. PCR primer sequences were: *ehubiquitin*, 5'-ATGCAAATATTTGTTAAGAC-3' and 5'-TTAATATCCTCCTCTTAATC-3'; *ehuba1*, 5'-ATGACACAACAAATTGATGAAGCCGTATTG-3' and 5'- TTAGAAATTCAAAAG AACATCTGGAAATTC-3'; *ehubc5*, 5'-ATGGCTATGCGTAGAATTCAAAAAG-3' and 5'-TTATGGTCGAGCATACATAC-3'; *ehring1*, 5'-ATGTCAAGAGAAGATTGTG-3' and 5'-TTAGTGATAAATAATTCGGTG-3'; *ehhect1*, 5'- ATGCGGAAACACCTCAAATA ACAAATAA-3' and 5'- TTAAATTAAACCAAATCCATTTGTATT-3'; *ehhect2*, 5'- ATGAGACCGGCTTGGAGACTTA-3' and 5'-TTAAGAAAATGCAAATCCTGAT TTTGATG-3'. Fragments subcloned and purified as recombinant proteins included the ubiquitin-fold domain of EhUba1 (a.a. 882-984), EhRING1 a.a. 1-246, EhHECT1 HECT domain (a.a. 277-660), and EhHECT2 HECT domain (a.a. 370-750). Point mutations to EhUbc5 were made using PCR and the overlap extension method [20].

Recombinant human Uba1, derived from insect cells, was purchased from Boston Biochem. Ubiquitin from bovine erythrocytes was purchased from Sigma. Since the sequences of bovine and human ubiquitin are identical, the protein is referred to as human

ubiquitin throughout the study, for clarity. For each of the *E. histolytica* components, BL21 *E. coli* were grown to an OD_{600nm} of 0.8 at 37°C and expression induced with 500 μM isopropyl-β-D-thiogalactopyranoside (IPTG) for 14-16 hours at 20°C. Pelleted bacterial cells were resuspended in N1 buffer containing 30 mM HEPES pH 8.0, 250 mM NaCl, and 30 mM imidazole and lysed by high-pressure homogenization with an Emulsiflex (Avestin; Ottawa, Canada). Cellular lysates were cleared by centrifugation at 100,000 x *g* for 1 hour at 4°C, and the supernatant applied to a nickel-nitrilotriacetic acid (NTA) FPLC column (GE Healthcare), washed extensively with N1, and eluted in N1 buffer with 300 mM imidzaole. For proteins used in biochemical experiments, eluted protein was pooled and resolved using a size exclusion column (HiLoad 16/60 Superdex 200, GE Healthcare) in S200 buffer containing 50 mM HEPES pH 7.5 and 100 mM NaCl (5 mM $ZnCl_2$ was included in the case of EhRING1 purification). For proteins used in crystallographic studies, protein eluted from the NTA column was pooled and dialyzed into imidazole-free N1 supplemented with 5 mM DTT overnight at 4°C in the presence of His_6-tobacco etch virus (TEV) protease to cleave the N-terminal affinity tag. The dialysate was then passed over a second NTA column to remove TEV protease and uncleaved protein, followed by resolution by size exclusion in S200 buffer. All proteins except EhUba1 were concentrated to 0.25 - 2 mM and snap frozen in a dry ice/ethanol bath for storage at -80°C. EhUba1 was found to precipitate upon freeze/thaw, but could be stably maintained at 4°C for at least two weeks. Protein concentration was determined by A_{280nm} measurements upon denaturation in 8 M guanidine hydrochloride, based on predicted extinction coefficients for each protein (http://us.expasy.org/tools/protparam.html).

6.3.2 Crystallization and structure determinations of EhUbiquitin and EhUbc5

Crystals of EhUbiquitin were obtained by vapor diffusion from hanging drops at 18°C in two different crystal forms. EhUbiquitin crystal form #1 was obtained by mixing EhUbiquitin at 17 mg/mL 1:1 with crystallization solution containing 25% (w/v) PEG 3350 and 100 mM citric acid pH 3.5. A single crystal grew to ~500 x 400 x 400 μm over 5 days, exhibiting the symmetry of space group P3$_2$21 (a = b = 49.8 Å, c = 63.8 Å, α = β = 90°, γ = 120°) and containing one monomer in the asymmetric unit. For the second crystal form, EhUbiquitin at 13 mg/mL in S200 buffer was mixed 1:1 with (and equilibrated against) crystallization solution containing 22% (w/v) PEG 3350, 200 mM LiSO$_4$, and 100 mM Bis-Tris pH 5.5. Crystals grew to ~200 x 100 x 100 μm over 3 days, exhibiting the symmetry of space group P2$_1$2$_1$2$_1$ (a = 38.6 Å, b = 49.9 Å, c = 76.8 Å, α = β = γ = 90°) and containing two monomers in the asymmetric unit. For data collection at 100K, crystals were serially transferred for ~1 minute into crystallization solution supplemented with 30% (v/v) glycerol in 10% increments and plunged into liquid nitrogen. Native data sets were collected at the GM/CA-CAT 23-ID-B beamline at the Advanced Photon Source (Argonne National Laboratory). Data were processed using HKL2000 [21]. The crystal structure model of human ubiquitin (PDB ID 1UBQ) was used as a molecular replacement search model using PHENIX AutoMR [22]. Refinement was carried out using phenix.refine [22], interspersed with manual revisions of the model using the program Coot [23]. Refinement consisted of conjugate-gradient minimization and calculation of individual anisotropic displacement and translation/libration/screw (TLS) parameters [24]. The current model of crystal form #1 contains one EhUbiquitin monomer; residues 7-10 and 73-77 could not be located in the electron density. Ramachandran plot analysis indicated 100% favored residues. The current

model of crystal form #2 contains two EhUbiquitin monomers; residues 75-77 of chain A and 74-77 of chain B could not be accurately modeled in the electron density. Ramachandran plot analysis indicated 98.6% favored, 1.4% allowed, and 0% disallowed residues.

Crystals of EhUbc5 were obtained by vapor diffusion from hanging drops at 18°C. EhUbc5 at 8 mg/mL in S200 buffer was mixed 1:1 with (and equilibrated against) crystallization solution containing 100 mM Tris pH 7.5, 14% (w/v) polyvinylpyrrolidone K15, and 500 μM $CoCl_2$. Hexagonal crystals grew to ~400 x 300 x 200 μm over 5 days and exhibited the symmetry of spacegroup $P2_12_12_1$ (a = 47.0 Å, b = 49.6 Å, c = 63.5 Å, α = β = γ =90°) with one monomer in the asymmetric unit. For data collection at 100K, crystals were transferred for ~1 minute into crystallization solution supplemented with 25% glycerol and plunged into liquid nitrogen. A native data set was collected at the SER-CAT 22-BM beamline at the Advanced Photon Source (Argonne National Laboratory). Data processing and refinement were carried out similarly to EhUbiquitin, as mentioned above. A molecular replacement solution was obtained using the crystal structure model of human UbcH5b (PDB ID 2ESK), modified to exclude water. The current model contains one EhUbc5 monomer with a cobalt ion coordinated by surface residues and by a monomer from the neighboring asymmetric unit. All EhUbc5 residues could be located in the electron density. Ramachandran plot analysis indicated 97.3% favored, 2.7% allowed, and 0% disallowed residues.

6.3.3 *In vitro* ubiquitin transfer assay

Ubiquitin transfer experiments were conducted essentially as previously described for NEDD8ylation [25], with adaptations to allow visualization of EhUba1~EhUb thioester

formation. Briefly, 1 μM EhUba1, 5 μM wild-type or mutant EhUbc5, and 15 μM EhUbiquitin were incubated 15 minutes at room temperature in ubiquitin transfer buffer containing 50 mM HEPES pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, and 5 mM ATP. The reactions were halted by denaturation in 5X non-reducing SDS sample buffer, and 50 mM DTT was added to specified reactions for 10 minutes prior to protein separation by denaturing SDS-PAGE and staining with Coomassie blue.

6.3.4 *In vitro* polyubiquitin chain formation assay

Polyubiquitin chain formation experiments were conducted essentially as previously described [26]. 50 nM EhUba1, 1 μM EhUbc5, 8 μM N-terminal FLAG epitope-tagged EhUbiquitin, and 10 μM EhRING1 were incubated at 37 °C for 45 minutes in reaction buffer containing 50 mM HEPES pH 7.5, 100 mM NaCl, 10 mM $MgCl_2$, and 5 mM ATP. Reactions were halted by denaturation in 5X reducing SDS sample buffer, and 50 mM DTT was added to all reactions 10 minutes prior to protein separation by SDS-PAGE. Proteins were transferred to a nitrocellulose membrane and probed with anti-FLAG (M2 monoclonal; Sigma).

6.3.5 PP_i:ATP radioisotope exchange assay

Isotope exchange assays were conducted as previously described [4]. All assays were performed at 37°C and contained 6 nM E1 ubiquitin-activating enzyme, 1 μM ubiquitin protein, 100 μM AMP, 10 mM $MgCl_2$, 100 μM non-radioactive pyrophosphate (PP_i), and variable ATP concentrations. Reactions were initiated by addition of the E1 enzyme. Concentrations of EhUba1 and human Uba1 were determined by the Bradford method.

Incubation times were <15 minutes and within the linear portion of the progress curve (prior to approaching equilibrium). Incorporation of $^{32}PP_i$ into ATP was determined by adsorbing the ATP to activated charcoal. Reactions were quenched with 5% (w/v) trichloroacetic acid (TCA) containing 4 mM non-radioactive PP_i, mixed with a 10% (w/v) slurry of charcoal in 2% TCA, pelleted by centrifugation, and washed three times with 2% TCA. Charcoal-bound radioactivity was quantified by Cherenkov radiation. Background radiation from control reactions lacking E1 was subtracted, and no significant exchange was observed in the absence of E1, ubiquitin, or ATP. All assays were performed in duplicate, and error bars represent standard error from three independent experiments. Data were fit by linear regression using GraphPad Prism v5.0. Approximate K_m and V_{max} values were derived from a Lineweaver-Burk plot, where the y-intercept = $1/V_{max}$ and the x-intercept = $-1/K_m$.

6.3.6 Surface plasmon resonance (SPR) assays

SPR-based measurements of protein-protein interactions were performed on a Biacore 3000 (GE Healthcare), essentially as described previously [19]. Briefly, purified His_6-EhUbc5 and His_6-EhUbc5(F62A) proteins were separately immobilized on an NTA biosensor chip using covalent capture coupling [27]. EhUbiquitin, EhUba1, or E3 ligase proteins were injected in 30 μL volumes at increasing concentrations. For binding experiments with EhUba1~EhUb, 20 μM EhUba1 and 50 μM EhUbiquitin were incubated 20 minutes at 30°C in buffer containing 50 mM HEPES pH 7.5, 100 mM NaCl, 1 mM ATP, and 10 mM $MgCl_2$ to allow for ubiquitin activation and thioester formation. EhUba1~EhUb was injected in 30 μL volumes using the KINJECT command with a 300 second dissociation time. Each EhUba1~EhUb injection was followed by a 10 μL injection of 50 mM DTT to detach

thioester-coupled EhUbiquitin from the EhUbc5 surface. Experiments were performed in running buffer containing 50 mM HEPES pH 7.4, 150 mM NaCl, 0.05% NP-40 alternative (Calbiochem), 50 μM EDTA, and 1 mM $MgCl_2$. Background changes in refractive index upon injection of samples was subtracted from all curves using BIAevaluation software v3.0 (GE Healthcare). Equilibrium binding analyses were conducted as previously described [28] using GraphPad Prism v5.0 to determine binding affinities.

6.4   RESULTS

6.4.1 Structural features of a divergent ubiquitin from *Entamoeba histolytica*

To give spatial context to the variant residues within *E. histolytica* ubiquitin [13] and their effects, if any, on the overall structure, we crystallized EhUbiquitin and determined its structure in two crystal forms. Under the first set of conditions, a single crystal was obtained that yielded diffraction data extending to 1.35 Å resolution. Although the diffraction data were of otherwise high quality (Table 6.2), detector overload resulted in exclusion of a significant fraction of low-resolution data during processing. Incomplete low-resolution data was manifested as unusually high R-factors and average *B*-factors during refinement (Table 6.2); however, the electron density was of good quality (Figure 6.1B), with clear electron density obtained for the divergent residues of EhUbiquitin upon phasing by molecular replacement with human ubiquitin (PDB ID 1UBQ). We also obtained EhUbiquitin crystals under a second condition, and a structure was determined using diffraction data to 2.15 Å (Table 6.2; Figure 6.1C). The data collection and refinement statistics corresponding to the second EhUbiquitin crystal form were near the mean of other deposited structures of similar resolution [22]. Although the structures derived from both diffraction data sets were

essentially identical, the second crystal form was utilized for further analyses due to the unusually high R-factors of crystal form #1.

The Cα trace of EhUbiquitin is highly similar to human ubiquitin (r.m.s.d. 0.78 Å; Figure 6.1A). This finding is consistent with the protein core residues being identical between the two homologs [13], with the exception of position 26, which exhibits a subtle variation of isoleucine in *E. histolytica* compared to the conserved valine in a broad diversity of other species (Figure 6.1D). Mapping the variant EhUbiquitin residues compared to its human homolog revealed clustering of divergent residues on a single surface, dominated by the α1 helix and including proximal portions of the β2-α1 and β3-β4 loops (Figure 6.1A). Notably, one of the residues unique to *E. histolytica* (Figure 6.1D) is an extra surface lysine (Lys-54). Since each surface lysine on mammalian ubiquitin can be utilized for polyubiquitin chain formation [9], it may be that *E. histolytica* possesses a unique K54 linkage polyubiquitination pattern.

To identify potential interactions that may involve the divergent EhUbiquitin surface, we superimposed the EhUbiquitin structure on a number of ubiquitin complex structures available in the RCSB database. Most of the surface of ubiquitin is utilized by one or more ubiquitin-binding proteins; however, the well-characterized interactions with E2 conjugating enzymes (*e.g.* UbcH5), HECT family E3 ligases, deubiquitinating enzymes (*e.g.* the SAGA complex) and ubiquitin-interacting motifs (*e.g.* RAP80) are not predicted to form significant interactions with the unique ubiquitin residues of *E. histolytica* (Figure 6.2A). Thus, at this time, we do not predict that the sequence variation seen in EhUbiquitin will significantly affect its interactions with similar ubiquitin-binding proteins encoded by the *E. histolytica* genome. The distinct EhUbiquitin surface may be utilized for other interactions unique to *E.*

histolytica. However, no specific function of the divergent EhUbiquitin surface can be ascribed at this time.

6.4.2 EhUbiquitin is activated by the E1 enzyme EhUba1

The *E. histolytica* genome encodes a single predicted ubiquitin-activating E1 enzyme, as well as predicted activating enzymes for Nedd8, SUMO, and other ubiquitin-like modifiers [15]. The putative ubiquitin-activating enzyme EhUba1 (AmoebaDB accession EHI_020270) was cloned from genomic DNA, expressed and purified from *E. coli*. EhUba1 possesses a predicted domain structure similar to *S. cerevisiae* Uba1 and other eukaryotic E1s (not shown). The sequence of EhUba1 was most similar to that of the slime mold *D. discoideum*, followed by *S. cerevisiae*, when compared to a broad spectrum of E1 enzyme sequences from different species (Figure 6.2C). To determine whether EhUba1 was capable of activating EhUbiquitin, we examined *in vitro* ubiquitin transfer with purified components. EhUba1 was seen to form a typical thioester bond with EhUbiquitin that was sensitive to reduction by DTT (Figure 6.3A). Magnesium and ATP were required for ubiquitin activation and thioester bond formation, suggesting a conserved mechanism of ubiquitin activation in *E. histolytica*. A number of mammalian E2 enzymes, together with an E1, are capable of catalyzing formation of isopeptide-linked polyubiquitin chains [29]. *In vitro* polyubiquitination experiments revealed that EhUbc5, together with EhUba1 could efficiently promote formation of EhUbiquitin chains up to four molecules in length (Figure 6.3B). Polyubiquitin chain formation required E2 enzyme, as well as the EhUba1 cofactors of magnesium and ATP.

We next sought to compare the activities of ubiquitin activating enzymes from *E. histolytica* and humans. To quantify the kinetics of enzyme-dependent incorporation of pyrophosphate (PP$_i$) into ATP, an *in vitro* radioactive ^{32}PP$_i$:ATP isotope exchange assay was employed as described previously [4]. In the presence of 1 μM ubiquitin protein, 100 μM AMP, 10 mM MgCl$_2$, and 100 μM non-radioactive PP$_i$, the initial rate of PP$_i$:ATP exchange catalyzed by human Uba1 was 132 picomoles/min (linear regression 95% confidence interval (C.I.) 88 - 176 picomoles/min) (Figure 6.3C), in good agreement with previous measurements under similar conditions [4]. EhUba1, in contrast, exhibited a significantly faster initial isotope exchange velocity of 460 picomoles/min (95% C.I. 428 - 493 picomoles/min), under identical conditions. The ~3.5-fold greater velocity of exchange for EhUba1 compared to human Uba1 is intrinsic to the E1 enzyme, given that the respective exchange rates in the presence of either human or *E. histolytica* ubiquitin substrates were indistinguishable (Figure 6.3C). A superimposition of EhUbiquitin with the structural model of *S.c.* Uba1~Ub suggests that the divergent ubiquitin surface in *E. histolytica* is not utilized in the E1/ubiquitin interface (Figure 6.2B), consistent with similar kinetics of the E1 enzymes in activating either the human or *E. histolytica* ubiquitin substrate. Residues that contact ATP in other E1 enzymes (*e.g.* [30]), such as the GxGxxGCE motif, are well-conserved in EhUba1 (not shown). A Lineweaver-Burk plot was constructed based upon PP$_i$:ATP exchange experiments with varying ATP concentrations, allowing estimation of K$_m$ and V$_{max}$ with respect to the ATP substrate for both human and *E. histolytica* E1 enzymes (Figure 6.3D). EhUba1 exhibited higher velocities of isotope exchange with V$_{max}$ = 459 picomoles/min (95% C.I. 392 - 552 picomoles/min), compared to its human homolog with V$_{max}$ = 115 picomoles/min (95% C.I. 93 - 152 picomoles/min) under these conditions.

However, the K_m values with respect to ATP were highly similar, being 45 μM and 50 μM, respectively (Figure 6.3D).

6.4.3 EhUba1 engages the E2 enzyme EhUbc5 and transfers activated EhUbiquitin

A number of predicted E2 ubiquitin-conjugating enzymes have been identified within the *E. histolytica* genome, although none has been functionally assessed [15]. We cloned a subset of candidate E2s from the *E. histolytica* genome and attempted to express and purify each of them from *E. coli*. One E2 protein (Amoeba DB accession EHI_083560; Figure 6.4A) with similarity to yeast Ubc4 and Ubc5 (termed EhUbc5) was highly expressed and thus selected for further study. Surface plasmon resonance was performed with immobilized EhUbc5, indicating a low affinity interaction with non-ubiquitin associated EhUba1 (Figure 6.4B). A K_D value for the EhUba1/EhUbc5 interaction could not be precisely quantified by equilibrium binding analyses due to the protein concentration limitations of our assay, but was greater than 150 μM. The isolated ubiquitin-fold domain (UFD) of EhUba1 (a.a. 882-984) exhibited a similar apparent affinity for EhUbc5, indicating its sufficiency for binding the E2 (Figure 6.4B). In similar SPR experiments, EhUba1 was first allowed to activate EhUbiquitin and form an EhUba1~Ub thioester complex (under conditions similar to Figure 6.3A) prior to injection over an EhUbc5 surface. Following dissociation of EhUba1~Ub, persistent residual resonance was observed (Figure 6.4C), suggesting that some activated ubiquitin may be transferred covalently to the EhUbc5-laden surface. Indeed, the residual resonance due to apparent thioester-coupled EhUbiquitin could be rapidly eliminated by injection of the reducing agent DTT (Figure 6.4C). An equilibrium binding analysis suggested at least a 10-fold greater affinity of EhUbc5 for EhUba1~Ub than for unconjugated

EhUba1 (Figure 6.4D). Notably, the apparent affinity of EhUbc5 for EhUba1~EhUb ($K_D \approx$ 12 μM) is likely underestimated by this approach, given that saturation could not be reached with each analyte injection (Figure 6.4C) and the reported EhUba1~EhUb concentrations (Figure 6.4D) assume that 100% of the EhUba1 enzyme injected was conjugated to ubiquitin. However, correction of either potential source of error would further lower the apparent EhUba1~Ub/EhUbc5 dissociation constant value, resulting in a >10-fold preference of EhUbc5 for EhUbiquitin-conjugated over unconjugated EhUba1. Transfer of the EhUbiquitin thioester from EhUba1 to EhUbc5 was also demonstrated by an *in vitro* assay (Figure 6.3A). Ubiquitin activation by EhUba1 and subsequent transfer to the E2 enzyme was found to be dependent on the presence of ATP and magnesium. The EhUbc5~EhUb thioester bond was also sensitive to the reducing agent DTT.

6.4.4 Structural features of the E2 ubiquitin conjugating enzyme EhUbc5 and its non-covalent interaction with EhUbiquitin

To gain further insight into ubiquitin conjugation in *E. histolytica,* EhUbc5 was crystallized and its structure determined by molecular replacement with diffraction data extending to 1.6 Å resolution (Table 6.2). Although EhUbc5 crystallized under a variety of conditions, the crystal form described here required a cobalt (II) salt. A cobalt ion could be identified in the electron density, contacting EhUbc5 molecules in adjacent asymmetric units (Figure 6.5C). The cobalt ion appears to be octahedrally coordinated with two histidine and one aspartate ligand from one EhUbc5 monomer, complemented by a glutamine from a symmetry-related EhUbc5 and by a water molecule. The overall structural model of EhUbc5 is remarkably similar to human UbcH5B (Cα r.m.s.d of 0.69 Å), given only 73% sequence identity and

83% similarity (Figures 6.4A and 6.5A). As seen in the case of EhUbiquitin, the majority of the EhUbc5 hydrophobic core is identical to human UbcH5B, while divergent residues predominantly reside on the protein surface (Figure 6.5A, B). The $\beta4$-$\alpha2$ and $\alpha2$-$\alpha3$ loops are highly conserved among E2 enzymes, including EhUbc5 (Figure 6.4A), suggesting a likely conserved mode of interaction with covalently attached ubiquitin (Figure 6.6A). In particular, EhUbc5 Cys-85 likely forms a thioester bond with the C-terminus of activated EhUbiquitin (Figure 6.3A). E2 enzymes are also known to bind ubiquitin-like modifiers in a non-covalent, "backside" interaction thought to be important for assembly of polyubiquitin chains [31, 32]. To assess a potential non-covalent interaction between EhUbc5 and EhUbiquitin, we predicted E2 enzyme residues likely to be involved based on the known interaction between human UbcH5A and human ubiquitin (Figure 6.4A) [33]. The analogous residues in EhUbc5 were 62% identical and 87% similar, suggesting a potentially conserved non-covalent interaction with EhUbiquitin. In support of this hypothesis, EhUbiquitin was found to bind EhUbc5, as measured by SPR (Figure 6.6B, C). The apparent low affinity of the EhUbc5/EhUbiquitin interaction (K_D = 410 $\pm$ 80 μM) is consistent with other monoubiquitin interactions, specifically homologous non-covalent interactions with other E2 enzymes (K_D values ~100-500 μM) [2, 31].

6.4.5 EhUbc5 engages a RING family E3 ubiquitin ligase

We next sought to identify E3 ligases in *E. histolytica* that may partner with EhUbc5 to ubiquitinate specific substrate proteins. The *E. histolytica* genome encodes a large number of putative E3 enzymes including HECT, RING, PHD, and U-box domain-containing proteins [15]. A subset of candidate E3s were cloned from *E. histolytica* genomic DNA and expressed

in *E. coli*. Of this subset, one RING family E3 and two HECT domains could be purified to near homogeneity in quantities suitable for biochemical experiments. The RING family protein (AmoebaDB accession EHI_020100), here termed EhRING1, contains predicted RING and zinc finger motifs according to SMART [34], although significant sequence divergence results in relatively high domain prediction E-values (Figure 6.7B). We were unable to identify a clear homolog for EhRING1 in either humans or yeast. The first 246 amino acids of EhRING1 were seen to bind EhUbc5 with a dissociation constant of 9.5 ± 0.5 μM, as determined by SPR (Figure 6.7B, C). In contrast, the HECT domains of two putative E3 ligases, termed EhHECT1 (EHI_011530) and EhHECT2 (EHI_124600), did not bind immobilized EhUbc5 (Figure 6.7B, C). A superimposition of EhUbc5 on crystal structure models of E2 enzymes in complex with a cIAP2 RING domain [35] and the CHIP U-box domain [36] suggests a well-conserved EhUbc5 surface at the predicted RING family E3 binding site (Figure 6.7A). To test whether EhUbc5 engages EhRING1 in a similar fashion, we mutated the conserved phenylalanine-62 to alanine, a residue predicted to contribute to RING binding. Indeed, EhUbc5(F62A) displayed a drastically reduced affinity for EhRING1 (Figure 6.7). However, EhUbc5(F62A) maintained its ability to form a thioester bond with activated EhUbiquitin (Figure 6.3A) and to bind EhUba1~EhUb (Figure 6.4D), indicating proper folding of the point-mutated EhUbc5 protein. EhRING1 had only modest effects, if any, on EhUbc5-catalyzed polyubiquitin chain formation, as assessed qualitatively by *in vitro* assays (Figure 6.3B).

6.5   DISCUSSION

Our experiments demonstrate the presence of a functional ubiquitin activation and conjugation pathway in *E. histolytica.* The substantial differences in the EhUbiquitin protein sequence compared to other species cluster on a single surface, constructed primarily of the α1 helix. Our analyses suggest that this particular surface is not central to the structurally-elucidated ubiquitin interfaces with E1s, E2s, HECT and RING E3s, ubiquitin interacting motifs (UIMs), and deubiquitinating enzymes (DUBs) in other species. However, the high degree of conservation at this surface in a diverse set of other organisms suggests a likely role(s) in ubiquitin functions. EhUbiquitin may have evolved to lack these functions, allowing sequence drift on the α1 helix and surrounding surface. Alternatively, EhUbiquitin may have evolved an as yet undetermined alternative use for this surface. The function of the α1 helix region has not yet been established in *E. histolytica*; accordingly, its potential value as a therapeutic target is unclear. Of particular interest is the presence of an eighth surface lysine (Lys-54) unique to EhUbiquitin (arginine in all other organisms examined) and included in the divergent surface. Complex polyubiquitination patterns utilizing all seven surface lysines and the N-terminus of ubiquitin exist in other species, corresponding to an array of interaction modes and affinities for various ubiquitin binding domains [37]. Thus, it is likely that additional unique polyubiquitination patterns and interactions arise in *E. histolytica,* involving the additional exposed lysine. Further studies are necessary to determine the prevalence of Lys-54 polyubiquitination in *E. histolytica* and its potential functions.

EhUba1 appears to activate ubiquitin in a similar fashion to its homologs in other species. However, the observed significant difference in maximal velocity of PP_i:ATP

exchange compared to human Uba1 suggests differences in ubiquitin activation kinetics. Additional work is needed to determine whether ubiquitin activation by this enzyme is necessary for parasitic virulence, and whether specific inhibition of EhUba1 is a viable therapeutic goal. Inhibitors of mammalian E1s have been used with some success, demonstrating the feasibility of this approach [38].

EhUbc5 exhibited a striking selectivity for EhUba1~Ub compared to unconjugated EhUba1. Since the EhUba1 ubiquitin-fold domain (UFD) bound EhUbc5 with an affinity similar to that of unconjugated EhUba1, it appears that some portion of the EhUba1~Ub complex, in addition to the UFD, contributes to E2 binding following ubiquitin activation. An altered EhUba1 conformation, or perhaps the EhUbiquitin molecule itself, may provide a higher affinity surface for EhUbc5. This functionality may help EhUbc5 recognize ubiquitin-bound E1 for efficient transfer and/or allow for rapid release of the E1/E2 complex once ubiquitin transfer has occurred. The non-covalent interaction of EhUbiquitin with EhUbc5 suggests a likely conserved mechanism for conjugating polyubiquitin chains [31]. EhUbc5 was also seen to engage a RING family E3 (EhRING1) through a conserved mode of interaction, to the exclusion of two HECT family E3s. This finding suggests a possible RING E3 specificity for EhUbc5; however, we cannot rule out the possibility that EhUbc5 interacts with other, untested HECT E3 ligases, like its yeast homologs [39]. It is unclear at this time which target proteins are ubiquitinated downstream of EhUbc5 and EhRING1.

The *E. histolytica* ubiquitin-proteasome pathway may provide therapeutic targets for potential treatment of amoebic colitis and amoebiasis. Of particular feasibility may be a proteasome inhibitor with selectivity for the *E. histolytica* protein target, given the previously demonstrated effects of proteasome inhibition on trophozoite proliferation and encystation

[16]. Alternatively, EhUba1-specific E1 inhibition may be expected to grossly perturb trophozoite function and viability, given the necessity of ubiquitin activation for multiple vital cellular processes in other eukaryotes [38].
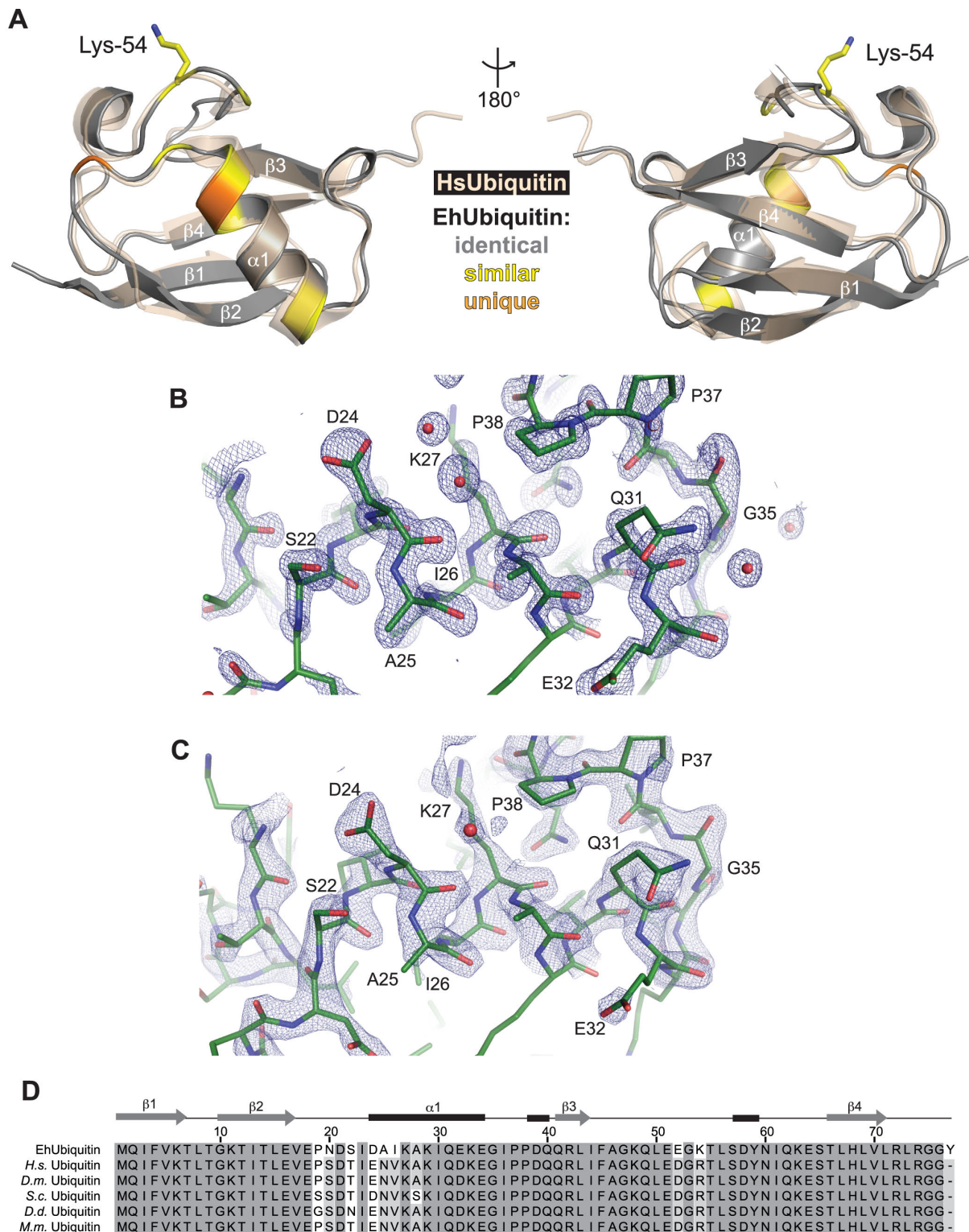
Figure 6.1. Sequence variations in EhUbiquitin cluster on a single surface. (A) The crystal structure of EhUbiquitin (*gray*) is superimposed with human ubiquitin (*wheat*; PDB ID 1UBQ), highlighting the similarities of their backbone structures. Sequence and structural

alignments were used to identify conservative (*yellow*) and non-similar (*orange*) amino acid differences between the *E. histolytica* and human homologs. The non-identical residues of EhUbiquitin cluster onto a single surface dominated by the first α-helix. EhUbiquitin possesses an additional surface lysine (Lys-54) that may allow for unique polyubiquitin linkage in *E. histolytica*. **(B)** Diffraction data for EhUbiquitin crystal form #1 extended to 1.35 Å resolution. Despite low-resolution incompleteness, the $2F_o$-F_c electron density map contoured to σ = 3.0 was of good quality. Red spheres represent ordered water molecules. **(C)** Diffraction data from the EhUbiquitin crystal form #2 were of high quality, given a lower resolution limit of 2.15 Å. The $2F_o$-F_c electron density map is contoured to σ = 3.0. **(D)** The protein sequence of EhUbiquitin was aligned to ubiquitin molecules from other species. EhUbiquitin shows the highest degree of sequence divergence, with 7 non-identical residues compared to human ubiquitin. The predicted sequence of EhUbiquitin also includes a C-terminal tyrosine that may be cleaved in *E. histolytica* trophozoites to expose a conserved C-terminal Gly-Gly motif. *D.m.* indicates *Drosophila melanogaster*, *S.c.* is *Saccharomyces cerevisiae*, *H.s.* is *Homo sapiens*, *D.d.* is *Dictyostelium discoideum*, and *M.m.* is *Mus musculus*.
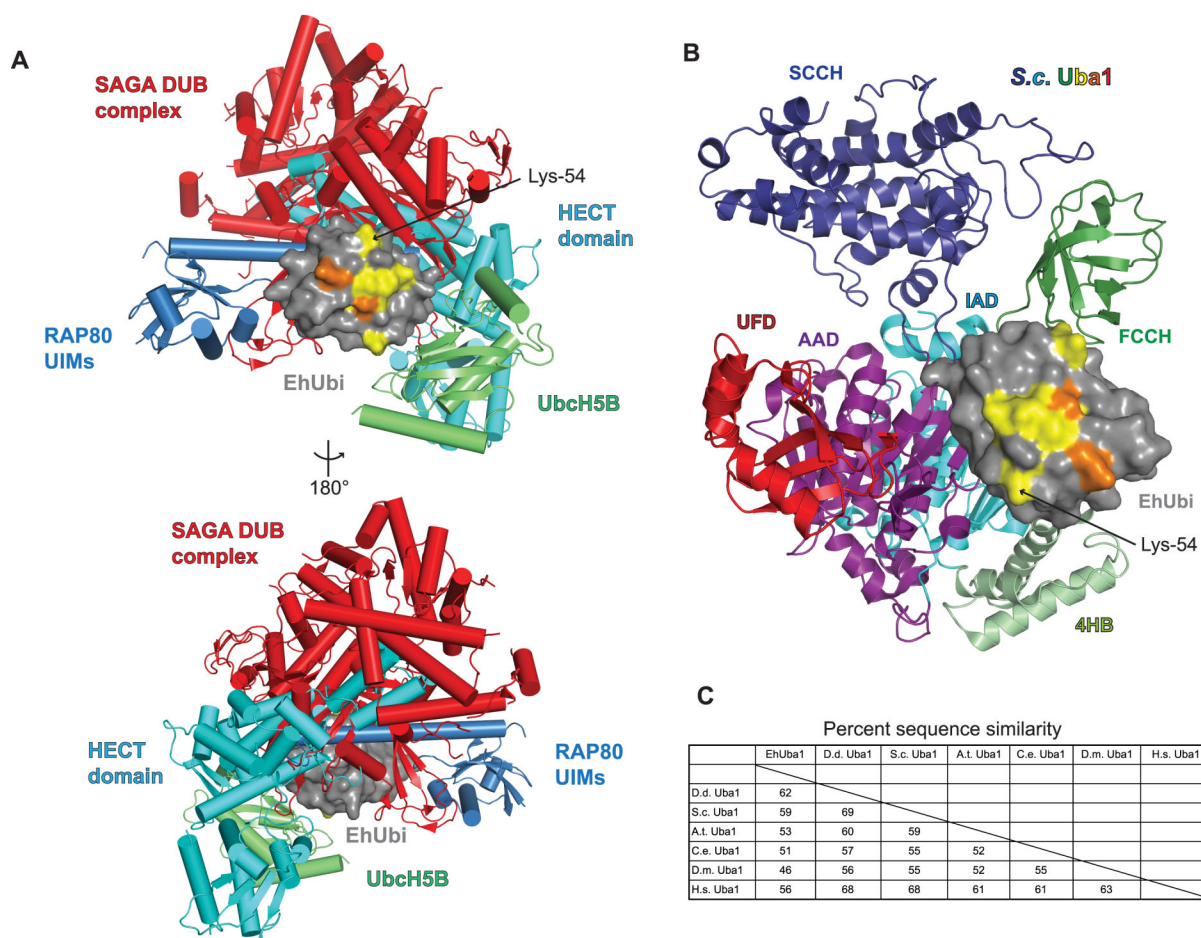
Figure 6.2. The divergent EhUbiquitin surface is not frequently utilized by structurally elucidated ubiquitin binding proteins. (A) The non-identical residues of EhUbiquitin cluster onto a single surface dominated by the first α-helix. This ubiquitin surface is not utilized by many of the best-studied and structurally characterized ubiquitin binding partners. EhUbiquitin (*gray*) was superimposed on structural models of ubiquitin-like modifier (ULM) homologs in complex with ULM-interacting proteins. Ubiquitin thioester formation with conjugating E2 enzymes and a HECT family E3 ligase are represented by UbcH5B (PDB ID 3A33) [6] and the Nedd4 HECT domain (PDB ID 2XBB) [40], respectively. Ubiquitin interaction with de-ubiquitinating (DUB) enzymes is illustrated by the SAGA DUB complex (PDB ID 3MHS) [41]. A ubiquitin interacting motif (UIM) protein is also exemplified by RAP80 (PDB ID 3A1Q) [42]. **(B)** The domain structure of yeast Uba1 is drawn and colored as in [30]. The interaction of EhUbiquitin with E1 enzymes was modeled by superposition of EhUbiquitin with yeast Uba1~Ub (PDB ID 3CMM). The divergent residues of EhUbiquitin (colored *yellow* and *orange* as in Figure 6.1) are not predicted to alter interaction with E1 enzymes. IAD indicates the inactive adenylation domain, FCCH is the first catalytic cysteine half-domain, 4HB is the four helix bundle domain, AAD is the active adenylation domain, SCCH is the second catalytic cysteine half-domain, and UFD is the ubiquitin-fold domain. **(C)** E1 enzyme sequence similarity and identity were compared across species. EhUba1 is most similar to that of the slime mold *D. discoideum*.

Percent sequence similarity

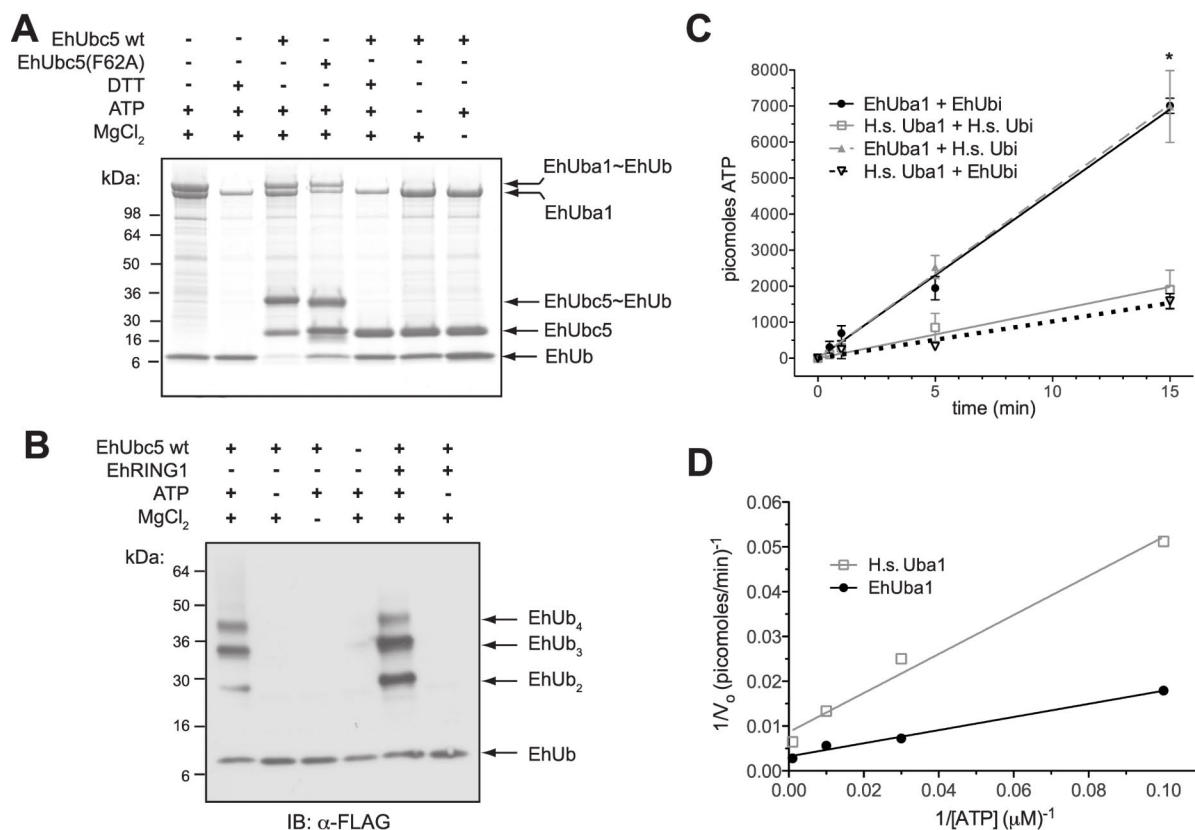|	EhUba1	D.d. Uba1	S.c. Uba1	A.t. Uba1	C.e. Uba1	D.m. Uba1	H.s. Uba1
D.d. Uba1	62						
S.c. Uba1	59	69					
A.t. Uba1	53	60	59				
C.e. Uba1	51	57	55	52			
D.m. Uba1	46	56	55	52	55		
H.s. Uba1	56	68	68	61	61	63	

Figure 6.3. The ubiquitin activating enzyme EhUba1 catalyzes EhUbiquitin thioester formation and transfer to the E2 enzyme EhUbc5. (A) Recombinant EhUba1 derived from *E. coli* forms a thioester bond with EhUbiquitin in an ATP- and magnesium-dependent fashion, as illustrated by SDS-PAGE and Coomassie blue staining under non-reducing conditions. Activated EhUbiquitin is seen to be transferred to the ubiquitin conjugating enzyme EhUbc5. Each of the covalent interactions detected was sensitive to 50 mM DTT reducing agent. The loss-of-E3-binding mutant EhUbc5(F62A) did not significantly affect ubiquitin thioester formation. **(B)** EhUba1 and EhUbc5 were sufficient to catalyze polyubiquitin chain formation in the presence of ATP and magnesium, as detected by SDS-PAGE and western blotting under reducing conditions. EhUba1 and EhUbc5 were seen to efficiently promote formation of chains up to four ubiquitin molecules under these conditions, and addition of EhRING1 had modest effects. **(C)** PP_i:ATP radioisotope exchange was utilized to compare *in vitro* ubiquitin activation by EhUba1 and human Uba1 in the presence of 6 nM E1 enzyme, 1 μM ubiquitin, 1 mM ATP, 100 μM AMP, and 100 μM non-radioactive PP_i. Under these conditions EhUba1 exhibited a statistically significant ~3.5-fold increase in isotope exchange rate compared to its human homolog (460 ± 33 (95% C.I.) versus 132 ± 46 (95% C.I.) picomoles/minute). The observed difference in ubiquitin activation kinetics is apparently intrinsic to the E1 enzyme, because isotope exchange rate was independent of the ubiquitin species utilized. These data represent average values of triplicate experiments with standard error of the mean. * indicates a statistically significant difference between EhUba1 and human Uba1 determined by Student's t-test. **(D)** The dependency of human Uba1- and EhUba1-mediated isotope exchange on ATP was assessed

under identical conditions as panel C, excepting ATP concentration. A Lineweaver-Burk plot allowed estimation of K_m and V_{max} values describing E1 activity with respect to ATP. Under these conditions, the apparent K_m and V_{max} values were 45 μM and 459 picomoles/min (95% C.I. 392 - 552 picomoles/min) for the *E. histolytica* components and 50 μM and 115 picomoles/min (95% C.I. 93 - 152 picomoles/min) for the human components, with respect to ATP. Each initial velocity measurement was conducted in duplicate, and the displayed results are representative of two independent experiments.
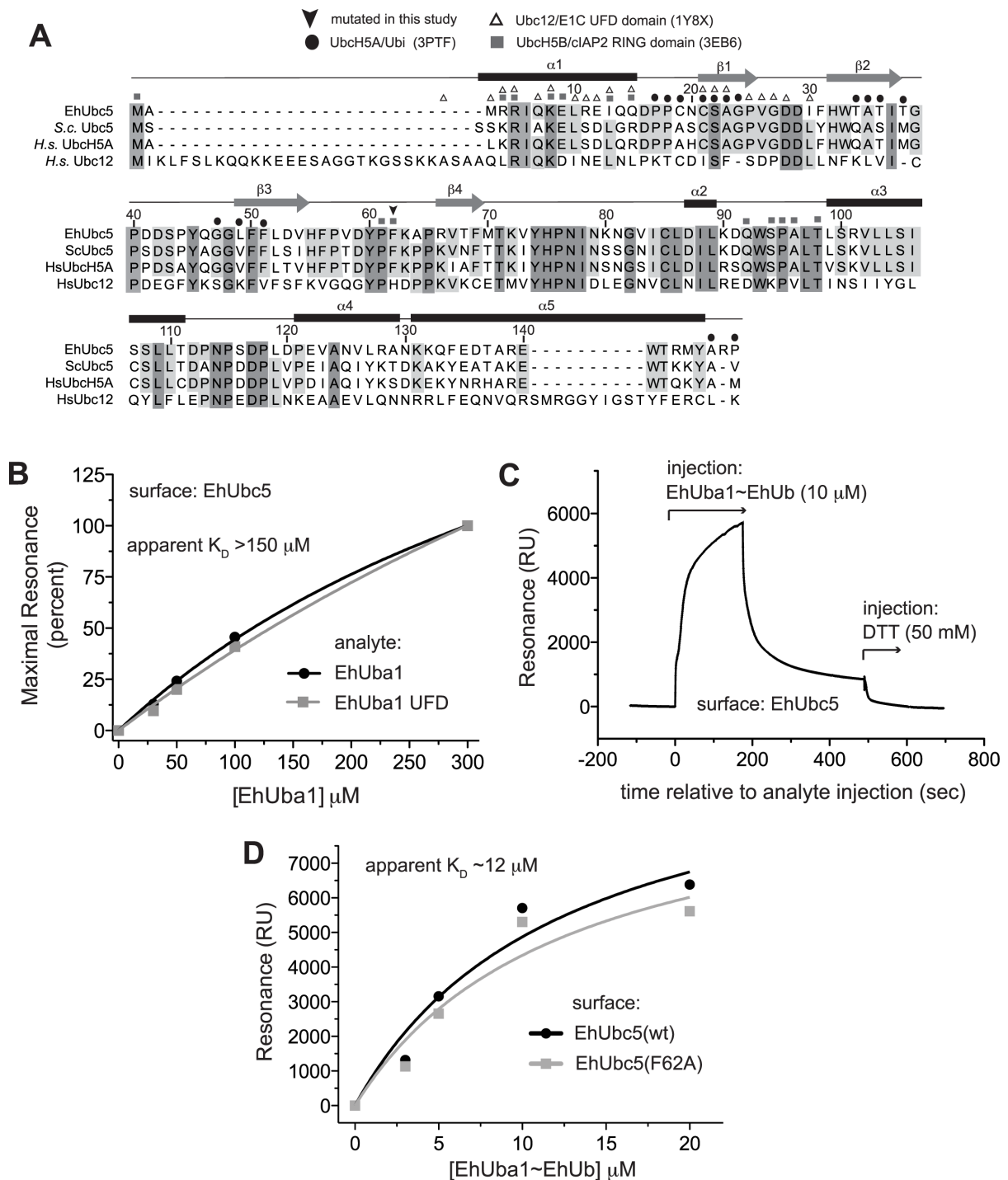
Figure 6.4. EhUba1 interacts directly with EhUbc5 through its ubiquitin-fold domain, and affinity is enhanced by the presence of activated EhUbiquitin. (A) The protein sequence of EhUbc5 was aligned to its closest homologs in yeast and humans, as well as human Ubc12. The secondary structure elements reflect the crystal structure of EhUbc5. Residues involved in interaction with E2 binding partners were derived from previously published structures (PDB IDs in parentheses). The majority of predicted interaction residues

are conserved in EhUbc5. **(B)** Surface plasmon resonance was utilized to measure direct binding of either unconjugated EhUba1 or its isolated ubiquitin-fold domain (UFD) to immobilized EhUbc5. EhUba1 binds EhUbc5 with low affinity in the absence of activated EhUbiquitin, and the UFD of EhUba1 is sufficient for binding. **(C)** EhUba1 was allowed to form covalent linkage with activated EhUbiquitin prior to injection over immobilized EhUbc5. Residual resonance following injection of EhUba1~EhUb was sensitive to the reducing agent DTT, likely indicating thioester transfer of EhUbiquitin to the immobilized EhUbc5. **(D)** The apparent affinity of EhUba1~EhUb for EhUbc5 was significantly higher than that of unconjugated EhUba1 (*panel B*). The apparent K_D value of ~12 μM is likely overestimated because less than 100% of the injected EhUba1 is expect to be conjugated to EhUbiquitin (Figure 6.3) and equilibrium of binding could not be reached under these experimental conditions (*panel C*). The loss-of-E3-binding mutant EhUbc5(F62A) did not significantly affect affinity for EhUba1~EhUb.
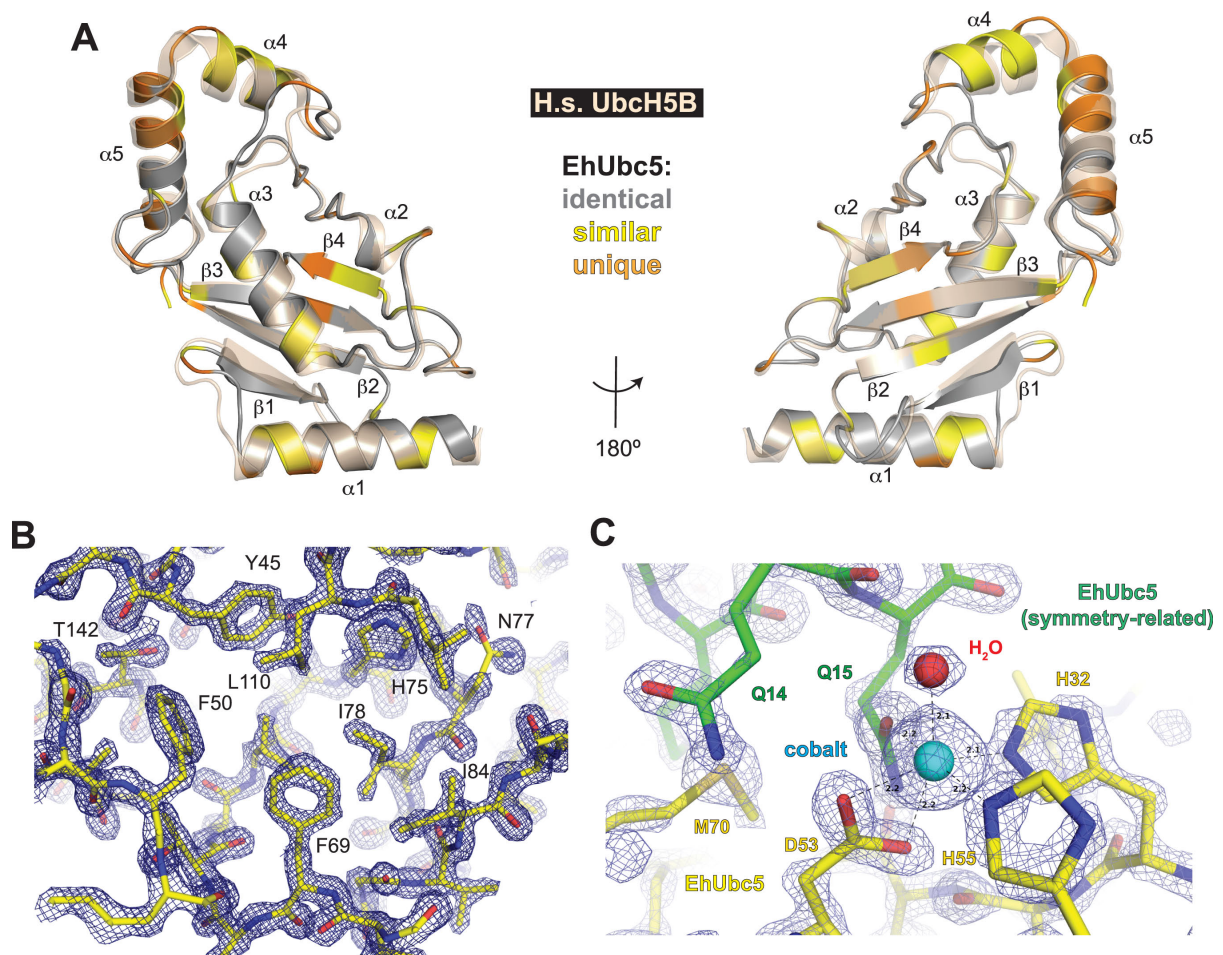
Figure 6.5. The crystal structure of EhUbc5 is highly similar to human UbcH5B despite sequence divergence. (A) The crystal structure of EhUbc5 (*gray*) is superimposed on its closest human homolog, UbcH5B (*wheat*, PDB ID 2ESK) [43]. Sequence and structural alignment was used to identify conservative (*yellow*) and non-similar (*orange*) amino acid differences between the *E. histolytica* and human homologs. **(B)** A representative $2F_o$-F_c electron density map of the EhUbc5 hydrophobic core, contoured to $\sigma = 2.5$, was rendered using PyMol. **(C)** Cobalt (*cyan sphere*) was required for crystallization and found incorporated in the crystal lattice at an interface between symmetry-related EhUbc5 molecules. Cobalt ion (*cyan sphere*) is coordinated by two histidines and an aspartic acid from one EhUbc5 molecule (*yellow*), a glutamine from the adjacent EhUbc5 molecule (*green*), and a water molecule. Each of the six apparent coordinating groups is 2.1-2.2 Å from the cobalt ion. The $2F_o$-F_c electron density map is contoured to $\sigma = 3.5$.
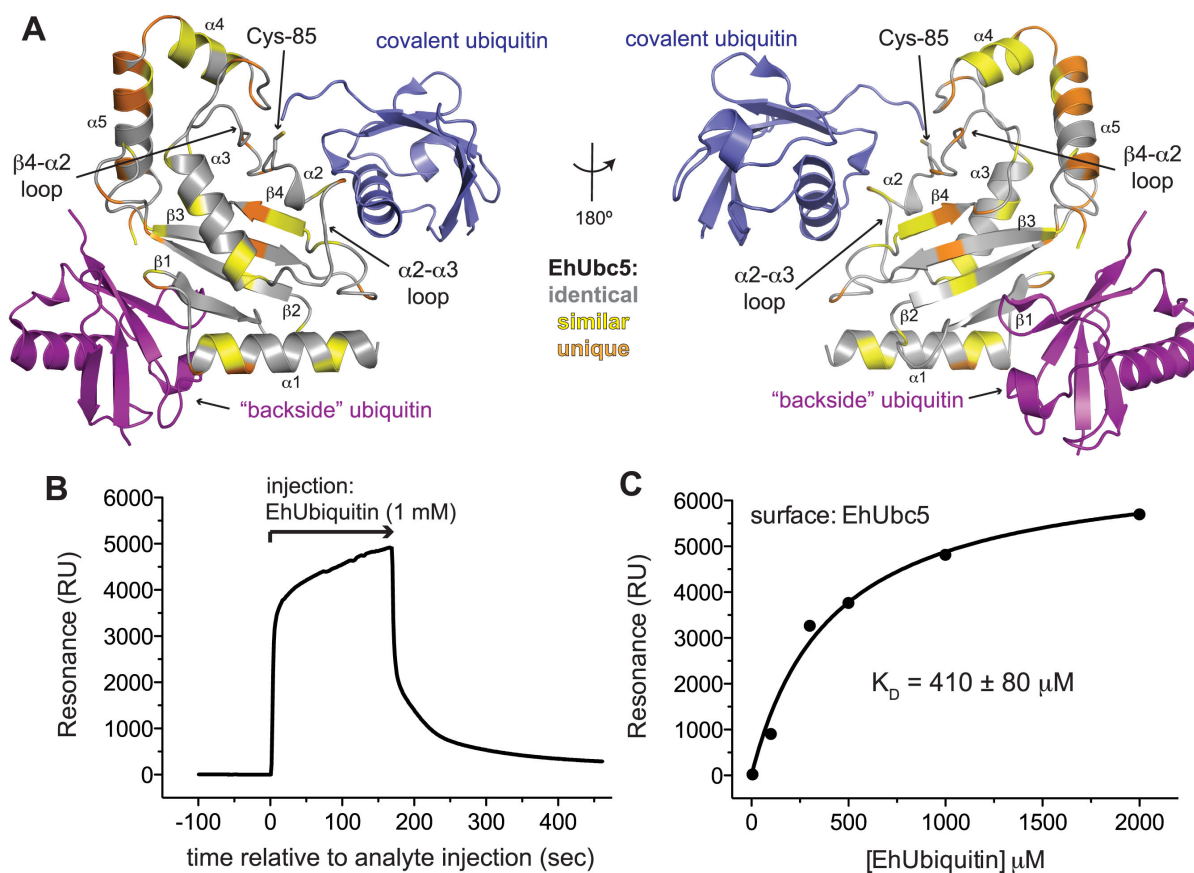
Figure 6.6. EhUbc5 engages EhUbiquitin through covalent thioester and non-covalent "backside" interactions. (A) The predicted interactions of EhUbc5 with thioester-conjugated ubiquitin (*blue*) and non-covalent "backside" bound ubiquitin (*purple*) was modeled by superimposition of EhUbc5 with human UbcH5B~Ub (PDB ID 3A33) [6] and non-covalent UbcH5A/Ub (PDB ID 3PTF) [33]. The two predicted ubiquitin interfaces of EhUbc5 are relatively well-conserved, suggesting similar modes of interaction in *E. histolytica*. **(B, C)** EhUbc5 interacts non-covalently with EhUbiquitin, as determined by SPR. The low affinity of interaction is consistent with the analogous interaction between human E2 domains and human ubiquitin [2].
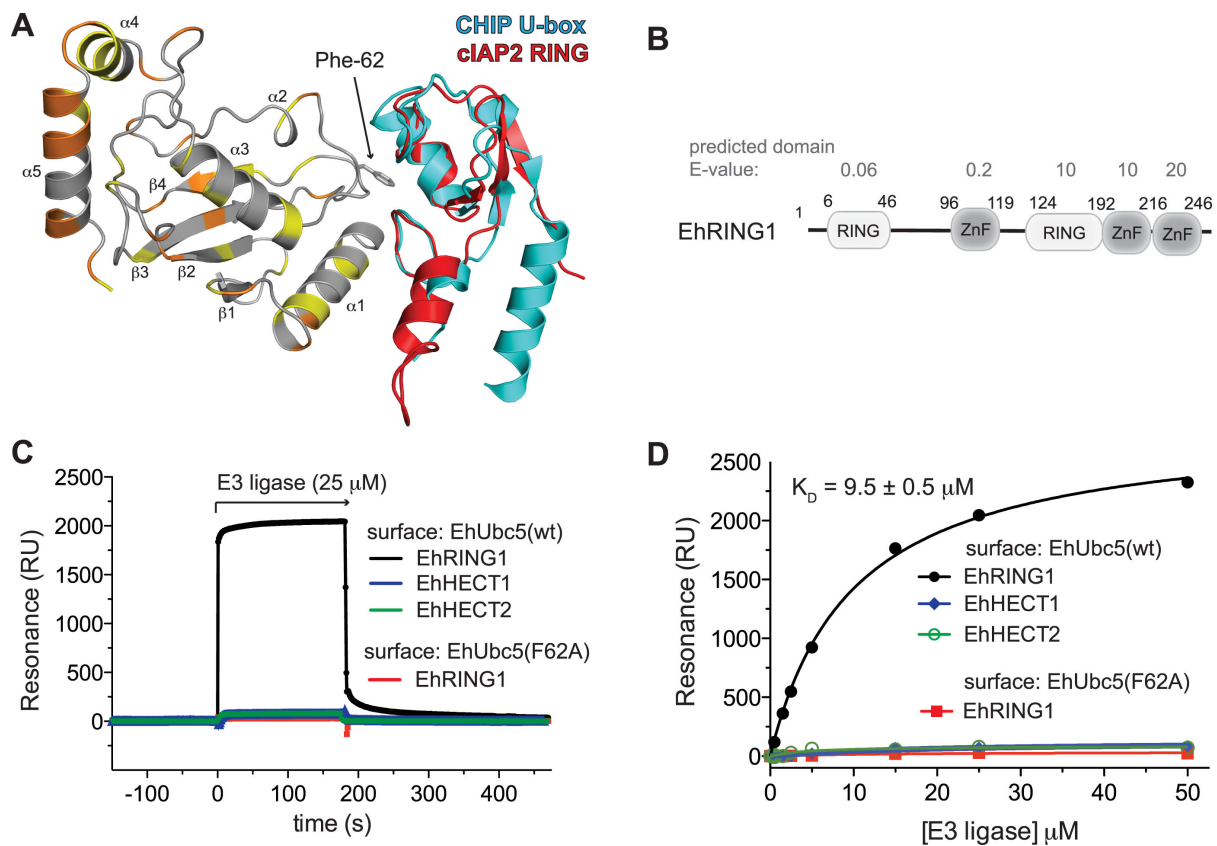
Figure 6.7. EhUbc5 exhibits a conserved mode of interaction with a RING-family E3 ligase. (A) The predicted interaction between EhUbc5 and RING family E3 ligases was modeled by superposition of EhUbc5 on the structural coordinates of the human UbcH5/CHIP U-box complex (PDB ID 2OXQ) [36] and the UbcH5/cIAP2 RING domain complex (PDB ID 3EB6) [35]. Phenylalanine 62 is highly conserved across species and contributes to the E2/RING E3 interface. **(B)** EhRING1 possesses an N-terminal predicted RING domain and zinc finger (ZnF) domain, as well as additional, weakly predicted RING and zinc finger domains. Domain prediction E-values were derived using SMART [34]. All biochemical experiments described in this study were conducted with EhRING1 residues 1-246. **(C, D)** EhUbc5 bound a putative RING family E3 ubiquitin ligase from *E. histolytica* with low micromolar affinity, but not two HECT family E3 ligases, as measured by SPR. Mutation of Phe-62 drastically reduced the affinity of EhUbc5 for EhRING1, indicating a likely conserved mode of E2/E3 interaction in *E. histolytica*.

Table 6.1. Ubiquitin and proteasome system genes differentially transcribed in *E. histolytica* trophozoites expressing EhGα1 or the dominant negative EhGα1^{S37C}.

Gene name	AmoebaDB accession no.	fold change upon EhGα1 expression (p value)	fold change upon EhGα1^{S37C} expression (p value)
ubiquitin system / proteasome			
hypothetical (RING finger domain)	EHI_104520	1.9 (0.004)	
ubiquitin-activating enzyme	EHI_004850	1.7 (0.013)	
zinc finger protein (RING-type)	EHI_165120	1.8 (0.004)	
Ulp1 protease family, catalytic	EHI_138530	2.2 (0.002)	
F-box domain containing protein	EHI_103710	-1.9 (0.007)	
26S protease regulatory subunit	EHI_185410	-1.8 (0.011)	
RING zinc finger protein	EHI_023310		-2.1 (0.008)
hypothetical (UBX domain)	EHI_027910		-2.5 (< 0.001)
proteasome subunit alpha type 3	EHI_098060		-2.0 (0.007)
zinc finger protein (RING-type)	EHI_130650		-1.9 (0.004)
RING zinc finger protein	EHI_161980		-2.1 (< 0.001)
zinc finger, RING-type	EHI_159840		-2.3 (< 0.001)
proteasome subunit alpha type 2-A	EHI_052140		1.7 (0.019)
WWE domain	EHI_069610		1.7 (< 0.001)
ubiquitin	EHI_083410		2.2 (< 0.001)
zinc finger, RING-type	EHI_110790		2.3 (< 0.001)
ubiquitin conjugating enzyme E2	EHI_131530		2.0 (< 0.001)
ubiquitin conjugating enzyme E2	EHI_135460		1.8 (0.009)
26S protease regulatory subunit	EHI_177320		1.9 (< 0.001)
19S cap proteasome S2 subunit	EHI_198010		3.6 (< 0.001)
26S protease regulatory subunit	EHI_053020		2.2 (0.020)

Table 6.2. Data collection and refinement statistics for EhUbiquitin and EhUbc5.

	EhUbiquitin (#1)	EhUbiquitin (#2)	EhUbc5
PDB access. code	4GU2	4GSW	4GPR
Data collection			
Space group	P3$_2$21	P2$_1$2$_1$2$_1$	P2$_1$2$_1$2$_1$
Cell dimensions			
a, b, c (Å)	49.80, 49.80, 63.83	38.61, 49.87, 76.82	46.97, 49.58, 63.46
α, β, γ (°)	90.0, 90.0, 120.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Wavelength (Å)	1.000	1.000	1.000
Resolution (Å)	25.7 - 1.35 (1.36-1.35)*	40.0 - 2.15 (2.17-2.15)*	50.0 - 1.60 (1.62-1.60)
No. unique reflect.	20493 (486)	8430 (210)	19998 (426)
R_{merge} (%)	4.3 (55.7)**	8.5 (27.9)**	4.4 (30.9) **
$I / \sigma I$	100 (6.9)	32.8 (3.1)	45.6 (2.8)
Completeness (%)	99.0 (100.0)	97.5 (96.8)	99.2 (87.3)
Redundancy	19.2 (17.9)	3.5 (3.3)	8.2 (3.7)
Wilson B-factor (Å^2)	18.9	31.9	17.1
Refinement			
Resolution (Å)	25.6 - 1.35 (1.42 - 1.35)	34.5 - 2.15 (2.28 - 2.15)	39.0 - 1.60 (1.64 -1.60)
No. reflections	20454 (2718)	8379 (1154)	19950 (1148)
R_{work} / R_{free} (%)	22.0 / 27.4 (22.8 / 31.3)	20.9 / 27.1 (26.6 / 35.0)	17.6 / 20.5 (20.6 / 26.0)
No. atoms			
Protein	1157	1171	1195
Ligand/ion	-	-	1
Water	66	44	179
Avg. B-factors (Å^2)			
Protein	55.2	41.5	15.0
Ligand/ion	-	-	8.1
Water	58.5	41.9	27.2
R.m.s deviations			
Bond lengths (Å)	0.015	0.014	0.014
Bond angles (°)	1.642	1.493	1.504

*Values in parentheses are for highest-resolution shell.
**All data were collected from a single crystal.

6.6   REFERENCES

1.   Burroughs, A.M., L.M. Iyer, and L. Aravind, *Structure and evolution of ubiquitin and ubiquitin-related domains.* Methods Mol Biol, 2012. **832**: p. 15-63.

2.   Hicke, L., H.L. Schubert, and C.P. Hill, *Ubiquitin-binding domains.* Nat Rev Mol Cell Biol, 2005. **6**(8): p. 610-21.

3.   Deshaies, R.J. and C.A. Joazeiro, *RING domain E3 ubiquitin ligases.* Annu Rev Biochem, 2009. **78**: p. 399-434.

4.   Haas, A.L. and I.A. Rose, *The mechanism of ubiquitin activating enzyme. A kinetic and equilibrium analysis.* J Biol Chem, 1982. **257**(17): p. 10329-37.

5.   Huang, D.T., et al., *Structural basis for recruitment of Ubc12 by an E2 binding domain in NEDD8's E1.* Mol Cell, 2005. **17**(3): p. 341-50.

6.   Sakata, E., et al., *Crystal structure of UbcH5b~ubiquitin intermediate: insight into the formation of the self-assembled E2~Ub conjugates.* Structure, 2010. **18**(1): p. 138-47.

7.   Huang, L., et al., *Structure of an E6AP-UbcH7 complex: insights into ubiquitination by the E2-E3 enzyme cascade.* Science, 1999. **286**(5443): p. 1321-6.

8.   Plechanovova, A., E.G. Jaffray, M.H. Tatham, J.H. Naismith, and R.T. Hay, *Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis.* Nature, 2012.

9.   Emmerich, C.H., A.C. Schmukle, and H. Walczak, *The emerging role of linear ubiquitination in cell signaling.* Sci Signal, 2011. **4**(204): p. re5.

10.   Haque, R., C.D. Huston, M. Hughes, E. Houpt, and W.A. Petri, Jr., *Amebiasis.* N Engl J Med, 2003. **348**(16): p. 1565-73.

11.   Stanley, S.L., Jr., *Amoebiasis.* Lancet, 2003. **361**(9362): p. 1025-34.

12.   Loftus, B., et al., *The genome of the protist parasite Entamoeba histolytica.* Nature, 2005. **433**(7028): p. 865-8.

13.   Wostmann, C., E. Tannich, and T. Bakker-Grunwald, *Ubiquitin of Entamoeba histolytica deviates in six amino acid residues from the consensus of all other known ubiquitins.* FEBS Lett, 1992. **308**(1): p. 54-8.

14.   Wostmann, C., D. Liakopoulos, A. Ciechanover, and T. Bakker-Grunwald, *Characterization of ubiquitin genes and -transcripts and demonstration of a ubiquitin-conjugating system in Entamoeba histolytica.* Mol Biochem Parasitol, 1996. **82**(1): p. 81-90.

15.   Arya, S., G. Sharma, P. Gupta, and S. Tiwari, *In silico analysis of ubiquitin/ubiquitin-like modifiers and their conjugating enzymes in Entamoeba species.* Parasitol Res, 2012. **111**(1): p. 37-51.

16.   Makioka, A., M. Kumagai, H. Ohtomo, S. Kobayashi, and T. Takeuchi, *Effect of proteasome inhibitors on the growth, encystation, and excystation of Entamoeba histolytica and Entamoeba invadens.* Parasitol Res, 2002. **88**(5): p. 454-9.

17.   Gilchrist, C.A., et al., *Impact of intestinal colonization and invasion on the Entamoeba histolytica transcriptome.* Mol Biochem Parasitol, 2006. **147**(2): p. 163-76.

18.   Bosch, D.E., et al., *Heterotrimeric G-protein Signaling is Critical to Pathogenic Processes in Entamoeba histolytica.* PLoS Pathogens, 2012(in press).

19.   Bosch, D.E., E.S. Wittchen, C. Qiu, K. Burridge, and D.P. Siderovski, *Unique structural and nucleotide exchange features of the Rho1 GTPase of Entamoeba histolytica.* J Biol Chem, 2011. **286**(45): p. 39236-46.

20.   Ho, S.N., H.D. Hunt, R.M. Horton, J.K. Pullen, and L.R. Pease, *Site-directed mutagenesis by overlap extension using the polymerase chain reaction.* Gene, 1989. **77**(1): p. 51-9.

21.   Otwinowski, Z.a.W.M., *Processing of X-ray Diffraction Data Collected in Oscillation Mode*, in *Methods in Enzymology*. 1997, Academic Press: New York. p. 307-326.

22.   Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.

23.   Emsley, P., B. Lohkamp, W.G. Scott, and K. Cowtan, *Features and development of Coot.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.

24.   Painter, J. and E.A. Merritt, *Optimal description of a protein structure in terms of multiple groups undergoing TLS motion.* Acta Crystallogr D Biol Crystallogr, 2006. **62**(Pt 4): p. 439-50.

25.   Calabrese, M.F., et al., *A RING E3-substrate complex poised for ubiquitin-like protein transfer: structural insights into cullin-RING ligases.* Nat Struct Mol Biol, 2011. **18**(8): p. 947-9.

26.   Yin, Q., et al., *E2 interaction and dimerization in the crystal structure of TRAF6.* Nat Struct Mol Biol, 2009. **16**(6): p. 658-66.

27.   Kimple, A.J., R.E. Muller, D.P. Siderovski, and F.S. Willard, *A capture coupling method for the covalent immobilization of hexahistidine tagged proteins for surface plasmon resonance.* Methods Mol Biol, 2010. **627**: p. 91-100.

28.   Kimple, R.J., et al., *Guanine nucleotide dissociation inhibitor activity of the triple GoLoco motif protein G18: alanine-to-aspartate mutation restores function to an inactive second GoLoco motif.* Biochem J, 2004. **378**(Pt 3): p. 801-8.

29.   Hochstrasser, M., *Lingering mysteries of ubiquitin-chain assembly.* Cell, 2006. **124**(1): p. 27-34.

30.   Lee, I. and H. Schindelin, *Structural insights into E1-catalyzed ubiquitin activation and transfer to conjugating enzymes.* Cell, 2008. **134**(2): p. 268-78.

31.   Brzovic, P.S., A. Lissounov, D.E. Christensen, D.W. Hoyt, and R.E. Klevit, *A UbcH5/ubiquitin noncovalent complex is required for processive BRCA1-directed ubiquitination.* Mol Cell, 2006. **21**(6): p. 873-80.

32.   Knipscheer, P., W.J. van Dijk, J.V. Olsen, M. Mann, and T.K. Sixma, *Noncovalent interaction between Ubc9 and SUMO promotes SUMO chain formation.* EMBO J, 2007. **26**(11): p. 2797-807.

33.   Bosanac, I., et al., *Modulation of K11-linkage formation by variable loop residues within UbcH5A.* J Mol Biol, 2011. **408**(3): p. 420-31.

34.   Letunic, I., T. Doerks, and P. Bork, *SMART 7: recent updates to the protein domain annotation resource.* Nucleic Acids Res, 2012. **40**(Database issue): p. D302-5.

35.   Mace, P.D., et al., *Structures of the cIAP2 RING domain reveal conformational changes associated with ubiquitin-conjugating enzyme (E2) recruitment.* J Biol Chem, 2008. **283**(46): p. 31633-40.

36.   Xu, Z., et al., *Interactions between the quality control ubiquitin ligase CHIP and ubiquitin conjugating enzymes.* BMC Struct Biol, 2008. **8**: p. 26.

37.   Husnjak, K. and I. Dikic, *Ubiquitin-binding proteins: decoders of ubiquitin-mediated cellular functions.* Annu Rev Biochem, 2012. **81**: p. 291-322.

38.   Edelmann, M.J., B. Nicholson, and B.M. Kessler, *Pharmacological targets in the ubiquitin system offer new ways of treating cancer, neurodegenerative disorders and infectious diseases.* Expert Rev Mol Med, 2011. **13**: p. e35.

39.   Stoll, K.E., P.S. Brzovic, T.N. Davis, and R.E. Klevit, *The essential Ubc4/Ubc5 function in yeast is HECT E3-dependent, and RING E3-dependent pathways require only monoubiquitin transfer by Ubc4.* J Biol Chem, 2011. **286**(17): p. 15165-70.

40.   Maspero, E., et al., *Structure of the HECT:ubiquitin complex and its role in ubiquitin chain elongation.* EMBO Rep, 2011. **12**(4): p. 342-9.

41.   Samara, N.L., et al., *Structural insights into the assembly and function of the SAGA deubiquitinating module.* Science, 2010. **328**(5981): p. 1025-9.

42.   Sato, Y., et al., *Structural basis for specific recognition of Lys 63-linked polyubiquitin chains by tandem UIMs of RAP80.* EMBO J, 2009. **28**(16): p. 2461-8.

43.   Ozkan, E., H. Yu, and J. Deisenhofer, *Mechanistic insight into the allosteric activation of a ubiquitin-conjugating enzyme by RING-type ubiquitin ligases.* Proc Natl Acad Sci U S A, 2005. **102**(52): p. 18890-5.

CHAPTER 7

CLINICAL IMPLICATIONS AND FUTURE DIRECTIONS

7.1 TARGETING HETEROTRIMERIC G PROTEIN SIGNALING

Although effective therapies for invasive amoebiasis, such as nitroimidazoles, have been available for many years, infection with *E. histolytica* remains prevalent in many parts of the world, primarily due to poor sanitation conditions and lack of access to health care resources [1]. Furthermore, standard treatments (typically metronidazole followed by paromycin) are not effective in all cases of symptomatic infection, with drug-related toxicities such as allergic reactions, neuropathies, and additional gastrointestinal symptoms contributing to treatment intolerance [1, 2]. Drug resistance has not yet emerged as a major problem for *E. histolytica* therapeutics; however, other causative agents of communicable infectious diseases frequently evolve to survive common therapies, particularly where inappropriate drug usage and failures to complete treatment abound [3]. Thus, pursuit of alternative therapeutics for invasive amoebiasis is warranted (*e.g.* [4-6]). One of the most attractive parasite-specific targets is the heterotrimeric G protein signaling pathway described in chapters 2 and 3, given that approximately one third of all currently FDA-approved drugs modulate G protein signaling [7, 8]. In particular, agonists, inverse agonists, and antagonists frequently target seven-transmembrane GPCRs [7]. As such, identification of a GPCR/ligand pair in *E. histolytica* would be valuable as a means for perturbing G protein signaling, both as a tool for

investigating contributions to amoebic biology and pathogenesis, and as a potential target for therapeutic intervention.

7.1.1 Identification of a GPCR and ligand in *E. histolytica*

No functional GPCR encoded by the *E. histolytica* genome has yet been established, although various hormones (typical mammalian GPCR ligands) modulate its pathogenicity as well as encystation of the related species *E. invadens* (reviewed in chapter 1) [9-13]. A bioinformatic analysis of the *E. histolytica* genome identified numerous predicted seven-transmembrane proteins (unpublished data). Among them was a single protein, termed Eh7TM1, that exhibited some sequence features held commonly among other known GPCRs as detected by 7TMRmine [14], as well as very limited sequence similarity to the cyclic AMP receptor cAR1 in *Dictyostelium discoideum* [15]. A codon-optimized Eh7TM1 gene was expressed in mammalian cells as a mCherry fusion protein and found to localize to the plasma membrane (Adam Kimple, unpublished data), consistent with GPCR-like behavior. The TANGO assay, based upon β-arrestin recruitment to activated GPCRs [16], was employed with mammalian cells stably expressing an Eh7TMR1 fusion protein in an attempt to identify a putative agonist for the candidate GPCR (Adam Kimple, unpublished data). Although no clear agonists were identified among thousands of screened compounds, expression of Eh7TM1 did cause a slightly elevated rise in apparent basal β-arrestin activity, suggesting that the membrane protein may adopt a recognizable conformation. Since the *E. histolytica* genome does not encode clear G protein coupled receptor kinases or β-arrestin homologs (unpublished data), Eh7TM1 may not assume a conformation upon ligand binding that is efficiently recognized by mammalian signaling machinery. Another approach utilizing

co-expression of Eh7TM1 and chimeric G proteins in yeast [17] likewise failed to identify clear ligands or establish Eh7TM1 unequivocally as a GPCR (Wes Kroeze, unpublished data). Putative *E. histolytica* GPCRs may not couple to yeast or mammalian G protein subunits, given the distinct sequence and structural features of EhGα1 (chapter 2) [18].

Given the limitations of leveraging mammalian GPCR de-orphaning systems to identify an *E. histolytica* receptor, *e.g.* dependency on coupling to mammalian or yeast proteins, an approach utilizing axenic trophozoite cultures may be preferable. *Entamoeba histolytica* prefers anaerobic growth conditions [19], a limitation for cell-based, medium or high throughput screening approaches. However, others have utilized mineral oil layering to minimize the oxygen exposure of trophozoites in traditional culture plate-formatted assays [20]. A number of genes are differentially transcribed in *E. histolytica* upon overexpression of either EhGα1 or EhGα1^{S37C} (chapter 2) [18]; accordingly, some may also be expected to undergo transcriptional changes upon trophozoite exposure to putative G protein-coupled receptor ligands. Promoters from differentially transcribed genes may be coupled with the open reading frames of firefly luciferase or other reporter enzymes, with preference given to those transcripts with large up- or down-regulation upon overexpression of EhGα1 or EhGα1^{S37C} (chapter 2) [18]. Trophozoites stably transfected with the resulting transcriptional reporters may be monitored for altered luciferase activity upon exposure to candidate GPCR ligands in a relatively high throughput fashion. Limitations of this approach include the current lack of positive controls for G protein signaling activation or deactivation, and the likelihood that transcriptional regulation of individual genes may not be specific to heterotrimeric G protein signaling. Simultaneously screening compounds against multiple transcriptional reporter strains and pursuing only compounds that alter transcription of

multiple target genes may lessen the latter limitation. One may also potentially counter-screen with trophozoites containing the transcriptional reporter and simultaneously overexpressing the dominant negative EhGα1^{S37C} in an attempt to isolate G protein signaling-dependent compounds. This transcriptional reporter approach would also not directly identify a specific *E. histolytica* GPCR, although further experiments such as compound-based affinity purification and radioligand binding with trophozoite membrane preparations may assist in receptor identification.

Although Eh7TM1 is the current best candidate GPCR by bioinformatic analysis (unpublished data), its roles in *E. histolytica* are currently unknown. Trophozoites could be engineered to overexpress, under tetracycline control, either an epitope-tagged wild type Eh7TM1 protein or a point mutant at the conserved "ionic lock" [21] of Eh7TM1, predicted to render the putative receptor constitutively active. Subsequent experiments could determine whether perturbed Eh7TM1 expression modulates processes related to EhGα1 signaling, such as trophozoite chemotactic migration, invasion, host cell killing, and cysteine protease secretion (chapters 2, 3) [18]. If the phenotypic effects of overexpressing Eh7TM1 were seen to parallel those of the EhGα1- or EhRGS-RhoGEF-overexpressing strains, there would be indirect, correlative evidence for Eh7TM1 being a GPCR. Limitations to this approach include its inability to demonstrate unequivocal GPCR (GEF) activity for Eh7TM1, the potential for mis-localization and non-physiological effects of Eh7TM1 overexpression, as well as the possibility that Eh7TM1 has little or no basal activity and that its overexpression may not significantly alter trophozoite behaviors in the absence of a stimulating ligand.

7.1.2 Investigating other heterotrimeric G protein signaling components

As discussed in chapter 1, the *E. histolytica* genome encodes a potential second Gα subunit, here termed EhGα2 (AmoebaDB accession EHI_186910), consisting of a N-terminal predicted Gα-like fold lacking conserved nucleotide binding motifs and a C-terminal PP2C-like phosphatase domain [22]. Full length EhGα2 and its isolated Gα-like domain (a.a. 1-338) could each be purified as recombinant proteins. EhGα2, but not the isolated Gα-like domain bound immobilized recombinant EhRGS-RhoGEF in SPR experiments (Figure 7.1) with an apparent affinity of 2 μM, comparable to that of EhGα1·AMF/EhRGS-RhoGEF (chapter 3) [23]. The unexpected lack of Gα-like domain binding to EhRGS-RhoGEF at concentrations up to 100 μM (Figure 7.1) could potentially be explained by the recombinant protein fragment being improperly folded or otherwise non-functional, or by a true requirement of the phosphatase domain for high-affinity interaction. These possibilities could potentially be addressed by similar binding experiments with a recombinant phosphatase domain fragment (*i.e.* EhGα2 lacking the Gα-like domain). To determine whether the Gα-like domain engages the RGS domain of EhRGS-RhoGEF in a manner akin to EhGα1 and like mammalian Gα/RGS interactions, the RGS domain point mutants E39K and N87A may be utilized (chapter 3) [23]. The *Drosophila* S2 cell spreading assay previously used to detect EhRGS-RhoGEF activation by both EhGα1 and EhRacC (chapter 3) [23] could also allow investigation of whether EhGα2 binding to EhRGS-RhoGEF can trigger downstream signaling in a cellular context. Interestingly, EhGα2/EhRGS-RhoGEF binding was insensitive to nucleotide state (Figure 7.1), suggesting together with absent nucleotide-binding motifs that EhGα2 may not cycle nucleotide like other known Gα subunits. GTPγ^{35}S

radionucleotide binding and single turnover GTPγ^{32}P hydrolysis assays can be employed to determine the nucleotide cycling, if any, of recombinant EhGα2. Initial attempts at detecting general phosphatase activity by recombinant EhGα2 using a broad specificity Ser/Thr phosphatase substrate have been unsuccessful (unpublished data), but other phosphatase substrates may be tried. An interesting possibility is that EhGα2 both binds and de-phosphorylates EhRGS-RhoGEF in trophozoites. Although no evidence currently exists for phosphorylation of EhRGS-RhoGEF *in vivo*, the *E. histolytica* strain overexpressing FLAG-EhRGS-RhoGEF [23] could be utilized to assess posttranslational modifications, for instance by anti-FLAG immunoprecipitation of EhRGS-RhoGEF in the presence of phosphatase inhibitors, and either phospho-antibody immunoblotting or mass spectrometry experiments. If EhRGS-RhoGEF is phosphorylated *in vivo*, the putative phosphatase activity of recombinant EhGα2 could potentially be assessed in trophozoite cell lysates or in EhRGS-RhoGEF enriched immunoprecipitates.

Potential signaling by EhG$\beta\gamma$, independent of EhGα1 and EhRGS-RhoGEF, also remains to be explored. G$\beta\gamma$ subunit signaling typically accompanies Gα signaling, or can be solely responsible for downstream effects, as seen in *A. thaliana* sugar sensing [24]. Candidate EhG$\beta\gamma$ effectors encoded by the *E. histolytica* genome include two phosducin-like proteins and multiple phosphoinositide 3-kinases. Although EhG$\beta\gamma$ could not be recombinantly produced in *E. coli* or insect cells (unpublished data), the dimer can be expressed in mammalian cells (chapter 2) [18]. Thus, potential interactions with co-expressed candidate effector proteins could be assessed with co-immunoprecipitation experiments. A more unbiased approach to identifying EhG$\beta\gamma$ binding partners could be accomplished with

yeast two hybrid assays utilizing EhGβ1 as bait, or yeast three hybrid using EhGβ1 and EhGγ1/2 as baits.

7.2   RHO FAMILY GTPase SIGNALING SPECIFICITY

7.2.1 Structure and function of the EhRacC/EhPAK4 complex

Three of seven encoded p21-activated kinases have been characterized in *E. histolytica*, as detailed in chapter 1, with overexpression studies suggesting roles for EhPAK1 and EhPAK2 in migration, surface receptor capping, and phagocytosis [25-27]. PAKs are Ser/Thr kinases that contain N-terminal regulatory regions that include a p21 binding domain (PBD) and a C-terminal kinase domain [28]; some PAK isoforms in lower organisms also possess a PH domain [29]. The regulatory regions of mammalian type I PAKs (PAK1-3) contain an autoinhibitory switch domain, and Rac and Cdc42 activate autoinhibited PAK dimers [30, 31]. In contrast, the type II PAKs (PAK4-6) exhibit a higher basal kinase activity that is weakly modulated through different mechanisms upon Rho family GTPase binding *in vitro*; cellular localization alterations induced by activated Rho GTPase binding are also thought to impact function [30]. The mammalian PAK isoforms are important players in the pathogenesis of multiple diseases, including specific cancers [32, 33], and have thus been the target of small molecule inhibitor development efforts. For instance, the compound IPA-3 (Figure 7.2A) was found to selectivity inhibit Cdc42-mediated PAK1 activation *in vitro* [33]. Despite the reactive and reduction-sensitive nature of IPA-3, as suggested by its internal disulfide bond (Figure 7.2A), the compound was found to inhibit PAK1 activation in mammalian cells at mid-micromolar concentrations, apparently through covalent

modification of the regulatory region including the PBD [33, 34]. Application of IPA-3 to *E. histolytica* trophozoites resulted in dose-dependent inhibition of chemotactic migration toward serum-containing media (IC_{50} ~22 μM; Figure 7.2B), consistent with its potential impairment of known PAK functions in amoebic motility [27, 35]. However, compound toxicity may contribute to the observed reduction in migration, as a 4-hour exposure to high concentrations of IPA-3 induced trophozoite death (LD_{50} ~200 μM), as detected by a membrane integrity assay (Figure 7.2C). An inactive structural isomer of IPA-3, PIR3.5 [33] can be employed in parallel experiments to further assess the potential non-specific and/or toxic effects of IPA-3 on *E. histolytica* trophozoites. Interpretation of these results are also limited by the current lack of a validated IPA-3 target in *E. histolytica*, an issue of particular concern given the reactive nature of IPA-3 (likely a covalent cysteine-modifier) [34]. The hypothesized binding of IPA-3 to *E. histolytica* PAK isoforms could be examined, for instance, with multiple recombinant PAK proteins and radioactive IPA-3, as described previously for mammalian PAK1 [34].

Together with three described diaphanous related formins and the GBD-containing actin binding protein EhNCABP166 (chapter 1), the seven PAK homologs in *E. histolytica* constitute the putative Rho family GTPase effectors with annotated Rho-specific binding motifs [36]. EhRacA has been observed to bind the regulatory region of EhPAK2 [35], and a crystal structure and biochemical analysis of the EhRho1/EhFormin1 GBD-FH3 domain tandem revealed some determinants of Rho family signaling specificity (chapter 5) [37]. However, the degree of signaling specificity or redundancy among the >20 *E. histolytica* Rho GTPases and their putative effectors, as well as their effector activation mechanisms remain largely unknown at this time.

The isolated p21-binding domains from two uncharacterized *E. histolytica* PAK homologs, here termed EhPAK4 (AmoebaDB access. EHI_152540) and EhPAK5 (EHI_043140) were seen to bind purified, activated EhRacC with a high degree of specificity and nucleotide state selectivity, as assessed by SPR experiments (Figure 7.3). The affinities of the GTPase deficient EhRacC(Q/L)·GTP for EhPAK4 and EhPAK5 were ~170 nM and ~1.9 μM, respectively. To further elucidate the determinants for the specific EhRacC·GTP/EhPAK4 PBD interaction, a stable complex was assembled and crystallized, and a structure determined to 2.85 Å resolution. A preliminary structural model of the complex (additional diffraction experiments and refinement are in progress) exhibits structural similarity to mammalian Rho family GTPase and PBD complexes, (*e.g.* Rac3/PAK4, PDB access. 2OV2; Figure 7.4). The overall structure of EhRacC(Q/L)·GTP is highly similar to other activated Rho family GTPases. The central β-sheet of the EhRacC G domain fold [38] is extended by a β-hairpin in the EhPAK4 PBD, a theme conserved among other Rho/PAK PBD complexes as well as Cdc42/WASP PBD [39]. However, the β-strands of the β-hairpin in the EhPAK4 PBD align slightly differently than mammalian Rho/PBDs of known structure, and the single EhPAK4 PBD α-helix is uniquely rotated ~90° relative to its homologs (Figure 7.4). Additional EhPAK4 contacts occur in the EhRacC switch regions (Figure 7.4), likely contributing to the nucleotide-state specificity observed in binding experiments (Figure 7.3). A detailed analysis of the protein/protein interface in this crystal structure will likely identify EhPAK4 and EhRacC residues important for binding, generating hypotheses that can be tested by mutagenesis and additional binding experiments. Furthermore, comparison of the EhPAK4-contacting regions of EhRacC with other *E. histolytica* Rho family GTPases may reveal variant residues responsible for specific

Rho/effector interactions. For instance, mutation of EhRacC residues to the corresponding EhRho1, EhRacD, or EhRacG residues would be expected to diminish binding to EhPAK4 and/or EhPAK5 (Figure 7.3). Conversely, it may be possible to introduce RacC-like residues into other Rho family GTPases and produce a gain-of-function binding to EhPAK4 and/or EhPAK5. Also of interest may be determining whether EhRacC binding activates EhPAK4 or EhPAK5 kinase activity. This may potentially be accomplished by quantifying EhPAK4/5 kinase activity *in vitro* with $ATP\gamma^{32}P$-dependent phosphorylation of myelin basic protein by recombinant PAK (*e.g.* [33]), or in mammalian cells with overexpressed EhPAK4/5 potentially phosphorylating itself or a typical PAK substrate, such as myosin light chain kinase or LIM kinase [28].

7.2.2 Characterization of FYVE domain-containing RhoGEFs

A family of eleven Dbl-related RhoGEFs, termed EhFP1-11, that also contain phospholipid-binding FYVE domains have been described in *E. histolytica*, as discussed in chapter 1 [40]. The FYVE domain of one such RhoGEF (EhFP4) was shown to mediate recruitment of the RhoGEF to trophozoite phagocytic cups and phagosomes through phospholipid binding, and when overexpressed in isolation, to impair phagocytosis [40]. Recombinant EhFP4 interacted with four *E. histolytica* Rho GTPases in pull-down assays, although nucleotide selectivity was not assessed and GEF activity could not be demonstrated in that report [40]. Thus, it was unknown whether the DH-PH tandems of these functionally important proteins serve as RhoGEFs and whether there is functional interplay between the FYVE and DH-PH domains. A number of FYVE domain-containing RhoGEFs were differentially regulated upon overexpression of either EhGα1 or EhGα1^{S37C} in *E. histolytica* trophozoites (chapter 2) [18],

276

suggesting a possible additional indirect link between heterotrimeric and Rho family G protein signaling. Among these transcriptionally regulated RhoGEFs, the DH-PH domain tandem from EhFP5 (AmoebaDB access. EHI_131540; a.a. 3-337) could be highly purified as recombinant protein (unpublished data). Although the EhFP5 DH-PH tandem did not accelerate fluorescent nucleotide exchange on a number of *E. histolytica* Rho GTPases tested (EhRho1, EhRacC, EhRacD, and EhRacG; unpublished data), it did promote nucleotide exchange on human RhoA (Figure 7.5). GEF activity has only been demonstrated at a single, relatively high concentration (10 μM EhFP5 DH-PH) thus far, but a GEF concentration-response curve generated with similar experiments would establish the potency of EhFP5 as an exchange factor for the non-physiological substrate human RhoA. Although an *E. histolytica* Rho substrate has not yet been identified, other recombinant family members can be assessed with this assay to determine the EhFP5/Rho GTPase specificity. A complementary approach would be to assess binding affinities (*e.g.* with SPR) for the EhFP5 DH-PH tandem toward *E. histolytica* Rho family members, as well as human RhoA, with the expectation that the GEF will have highest affinity for Rho GTPases in their nucleotide-free states [41]. Although human RhoA is a non-physiological substrate for EhFP5, this cross-species coupling allows utilization of mammalian cell-based assays for exploring the potential interplay between domains within EhFP5. For instance, full length EhFP5, the isolated DH-PH tandem, or the isolated FYVE domain may be exogenously expressed in mammalian cells, and human Rho signaling activation assessed by stress fiber formation or serum response element (SRE) transcriptional reporter induction [42]. The EhFP5 DH-PH tandem is hypothesized to activate human RhoA in cells, based on its *in vitro* GEF activity (Figure 7.5), while the isolated FYVE domain likely does not. The FYVE domain in the full

277

length EhFP5 may potentially modulate the GEF activity of the DH-PH tandem, perhaps in a membrane phospholipid-dependent fashion.

7.3   DE-UBIQUITINATING ENZYMES IN E. histolytica PATHOGENESIS

In addition to functional ubiquitin conjugation enzymes (chapter 6), the *E. histolytica* genome encodes a ~20-member family of de-ubiquitinating enzymes (DUBs) (unpublished data). DUBs are cysteine proteases that cleave isopeptide-linked ubiquitin molecules, allowing for reversibility of protein ubiquitination and recycling of ubiquitin monomers [43]. Development of specific DUB-inhibiting small molecules is ongoing, particularly for applications in the treatment of various cancers where inhibition of DUBs can induce selective cell death [44]. Since ubiquitination and de-ubiquitination are important for numerous critical cellular processes across species [45], DUB inhibition in *E. histolytica* may be expected to grossly impair or kill trophozoites. Three commercially available DUB inhibitors with reported broad specificity, PR-619, G5, and WP1130 [44], were selected and examined for effects on trophozoite proliferation. Two-day exposure to WP1130 and G5 impaired trophozoite proliferation in a typical concentration-dependent fashion (Hill slope ~1), with IC_{50} values of ~10 µM and ~25 µM, respectively (Figure 7.6B). PR-619 only impaired trophozoite proliferation at high micromolar concentrations, and exhibited an atypical concentration-response pattern (Hill slope $\neq$ 1), suggesting non-specific toxicity (Figure 7.6B). The observed effects of DUB inhibitors on *E. histolytica* proliferation could be due to either increased cell death or impaired cell division. Trophozoite cell death in response to DUB inhibitor application was investigated with a membrane integrity assay. At time points up to four hours, WP1130 and G5 did not induce marked amoebic cell death, while

PR-619 exhibited a dose-dependent cytotoxicity (Figure 7.6C, D). Although the toxicity of PR-619 could potentially be due to its on-target inhibition of DUB enzymes, the atypical concentration-response pattern with respect to proliferation and the dual azido groups in the PR-619 chemical structure suggest promiscuous reactivity (Figure 7.6A, B). WP1130 and G5 did not significantly induce cell death at times up to 4 hours; however, either compound might cause trophozoite death at later time points. Alternatively, these two compounds may exert their effects on proliferation by impairing amoebic cell division, a hypothesis that could be tested by cell division rate measurements, such as the degree of BrdU incorporation into newly synthesized DNA or dilution of a cytoplasmic fluorescent marker (CFSE). Also currently unknown are the molecular targets of WP1130 and G5. Ubiquitin derivatives that act as covalently-modifying "suicide substrates" for active DUBs have been leveraged, together with mass spectrometry, to yield profiles of DUB activity in cells on a proteomic scale [46]. The putative *E. histolytica* DUB targets of WP1130 and G5 could be suggested by profiling active DUBs in trophozoite lysates following treatment with either DUB inhibitors or DMSO, as done for mammalian cells in [47].

A single de-ubiquitinating enzyme from *E. histolytica*, here termed EhDUB1 was isolated for biochemical characterization and comparison to mammalian DUBs. Recombinant EhDUB1 hydrolyzed a human ubiquitin dye conjugate (ubiquitin-rhodamine) commonly used as a generic DUB substrate [48], a reaction impaired by pretreatment of EhDUB1 with the "suicide substrate" inhibitor ubiquitin-aldehyde [49] (Figure 7.7A-C). A Michaelis-Menten kinetic analysis of EhDUB1 activity revealed a K_m of ~500 nM with respect to the ubiquitin-rhodamine substrate and a V_{max} of ~3 s^{-1}. Some DUBs preferentially hydrolyze specific polyubiquitin linkages [50]. An internally quenched fluorescence (IQF) assay, with

Lys-48 or -63 linked di-ubiquitin substrates, demonstrated that EhDUB1 can cleave either linkage with apparently similar efficiencies (Figure 7.7D). WP1130, G5, and PR-619 did not inhibit EhDUB1 activity in similar *in vitro* experiments at concentrations up to 100 μM (unpublished data), suggesting that this specific enzyme is not a DUB inhibitor target contributing to impaired trophozoite proliferation. However, these experiments do establish a similarity of biochemical function among *E. histolytica* and mammalian DUBs and provide a viable method by which other DUBs may be profiled for small molecule inhibition.

Collectively, the work presented in this dissertation have elucidated structural and biochemical aspects of G protein signaling and ubiquitination in *E. histolytica*, as well as some of their relationships to the pathogenesis of invasive amoebiasis. However, much remains to be explored, and further efforts will likely add to our understanding of trophozoite biology and pathogenicity, as well as suggest potential targets for anti-amoebiasis drug development.
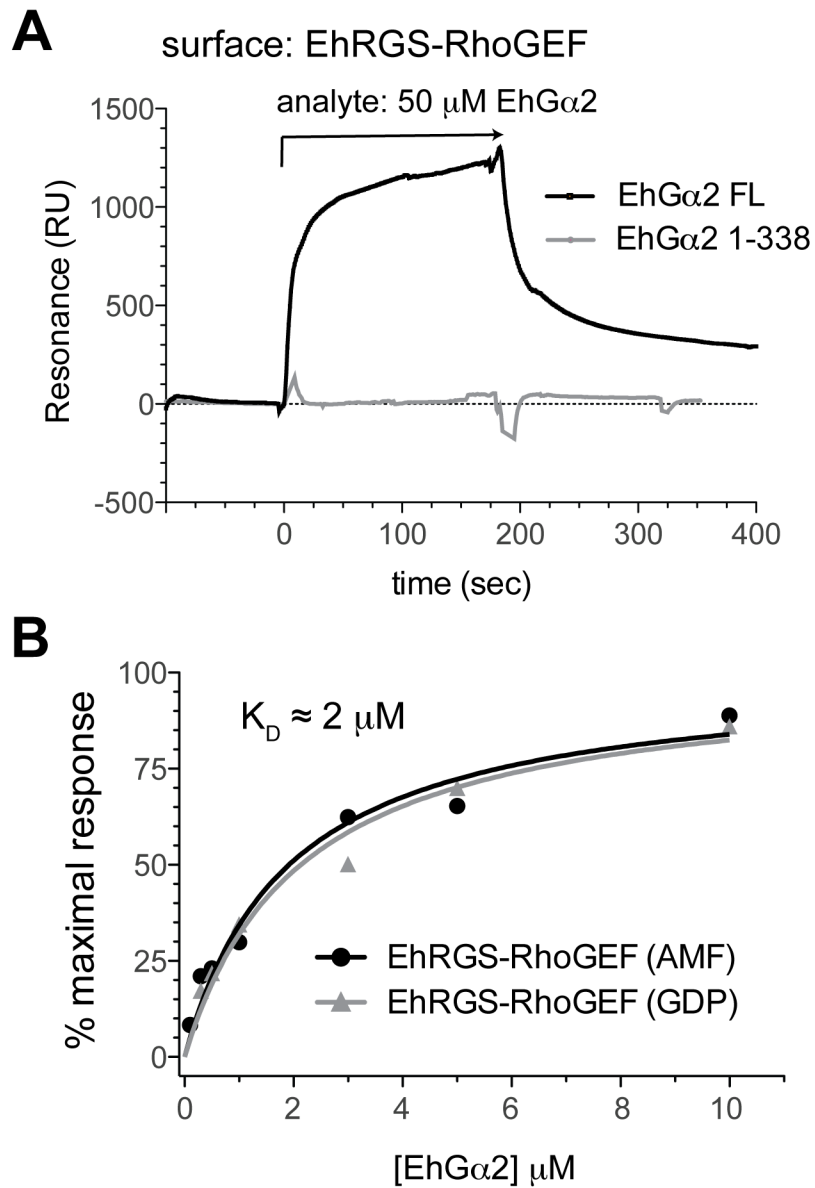
Figure 7.1. EhGα2 engages EhRGS-RhoGEF in a nucleotide state-independent fashion.
(A) Full length EhGα2, but not the isolated Gα-like domain (a.a. 1-338), bound immobilized
EhRGS-RhoGEF as measured by SPR. **(B)** The affinity of EhGα2 for EhRGS-RhoGEF was
similar to that of EhGα1 (chapter 2), but binding was independent of nucleotide state (AMF
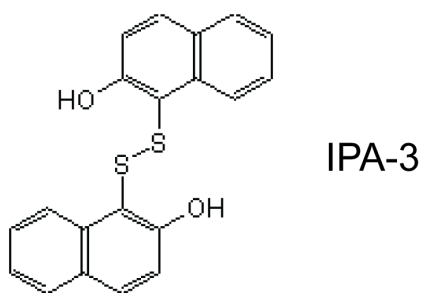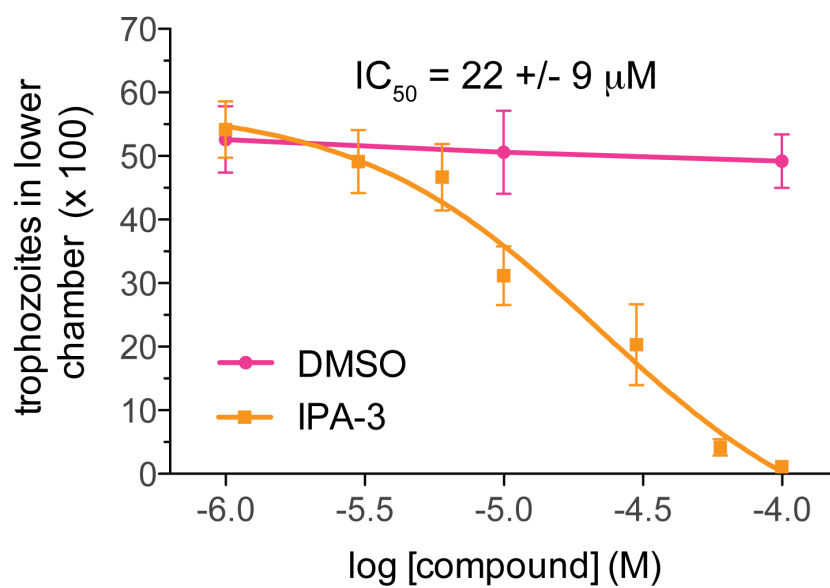indicates GDP·AlF₄⁻).

A

IPA-3

B

IC$_{50}$ = 22 +/- 9 μM

DMSO
IPA-3

trophozoites in lower chamber (x 100)

log [compound] (M)

C

100% lysis control

LD$_{50}$ ≈ 200 μM
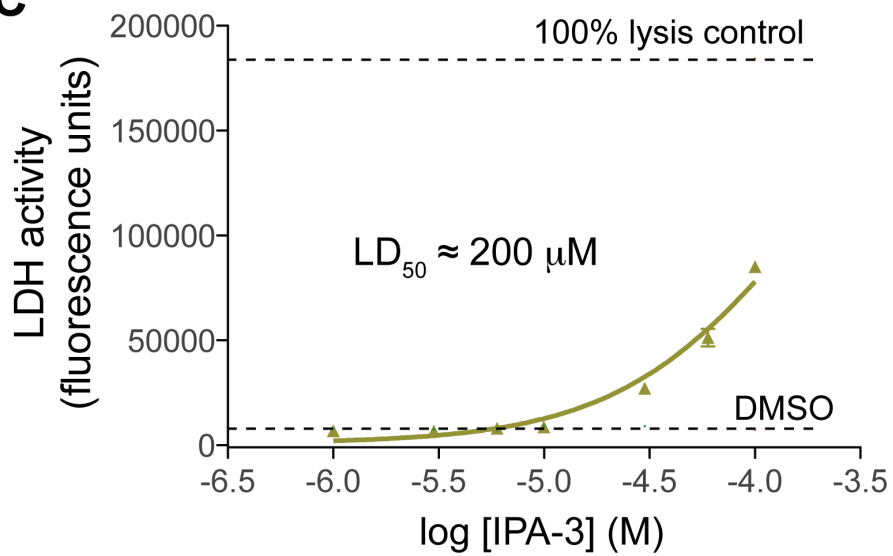
DMSO

LDH activity (fluorescence units)

log [IPA-3] (M)

Figure 7.2. The mammalian PAK inhibitor IPA-3 impairs *E. histolytica* chemotactic migration. (A) The chemical structure of IPA-3 includes a disulfide bond, likely implicated in covalent modification of PAK regulatory regions. **(B)** Exposure of trophozoites to IPA-3 or DMSO as a negative control for 2 hours resulted in a concentration-dependent impairment of chemotactic migration toward serum-containing TYI medium, as measured by a Transwell assay. **(C)** IPA-3 is cytotoxic to trophozoites (4-hr incubation) at high concentrations, as indicated by a membrane integrity assay.

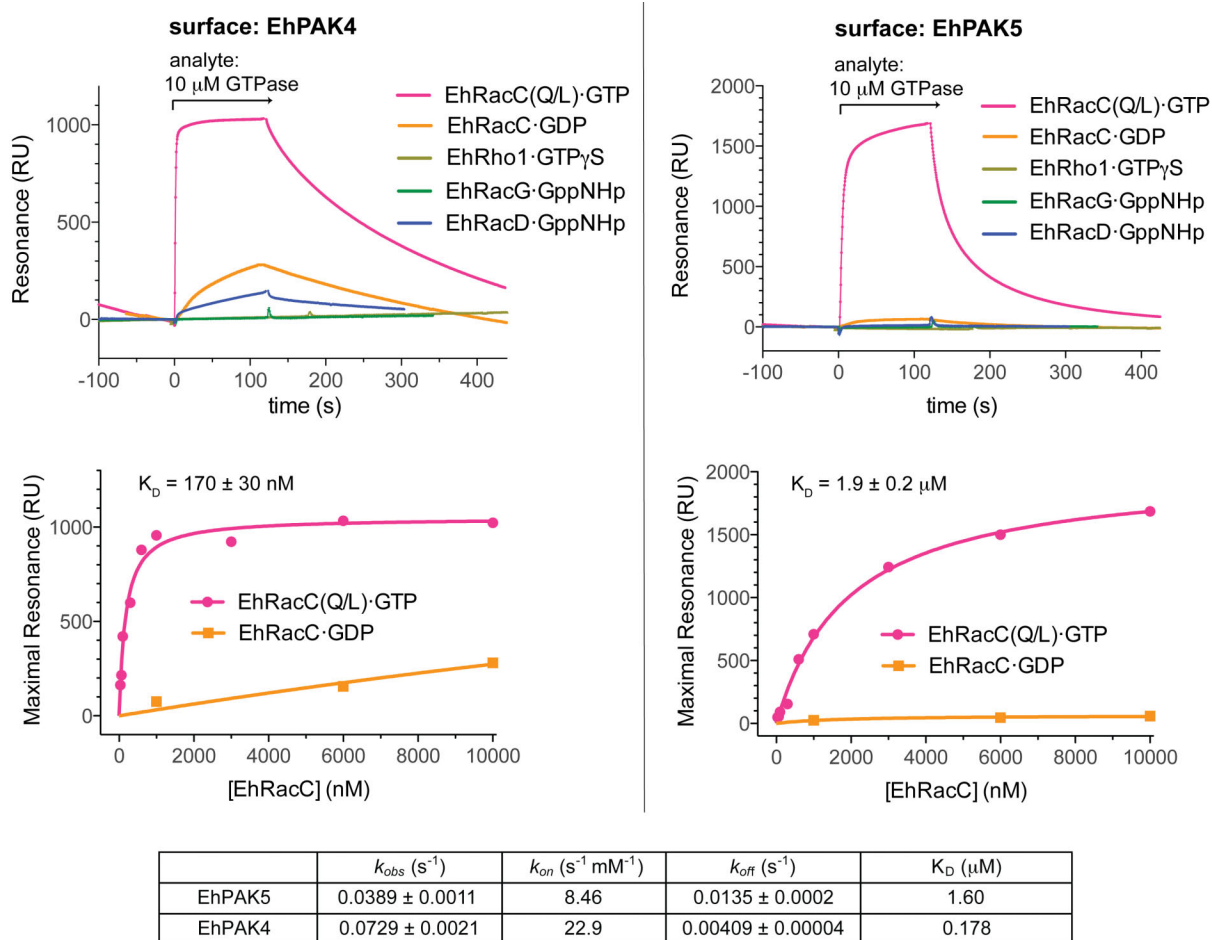|	k_{obs} (s^{-1})	k_{on} (s^{-1} mM^{-1})	k_{off} (s^{-1})	K_D (μM)
EhPAK5	0.0389 ± 0.0011	8.46	0.0135 ± 0.0002	1.60
EhPAK4	0.0729 ± 0.0021	22.9	0.00409 ± 0.00004	0.178

Figure 7.3. The EhPAK4 and EhPAK5 PBDs selectively bind activated EhRacC. The isolated p21-binding domains of two previously uncharacterized PAK isoforms were immobilized and found to bind the GTPase-deficient mutant EhRacC(Q/L)·GTP, to the exclusion of three other Rho family GTPases, as measured by SPR. Both PAKs exhibited high nucleotide state selectivity and affinity for activated EhRacC, suggestive of a Rho/effector relationship.
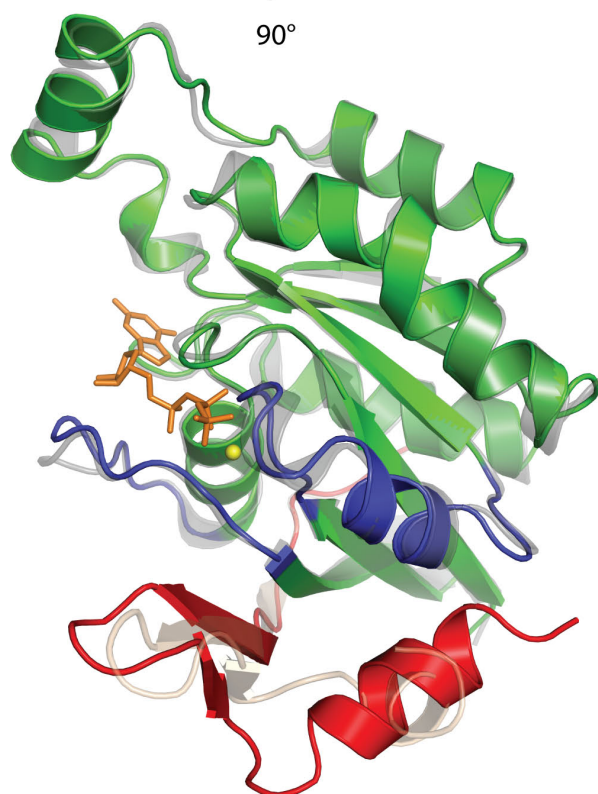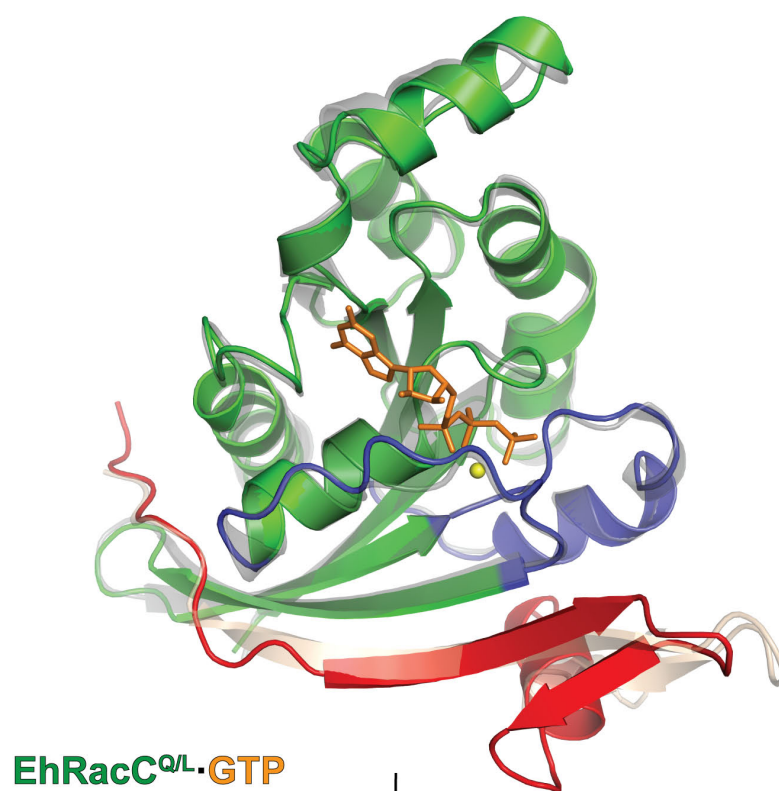
EhRacCQ/L·GTP
(switch regions)

EhPAK4 PBD

90°

285

Figure 7.4. The crystal structure of EhRacC·GTP/EhPAK4 PBD reveals similarity to mammalian Rho/PBD complexes. The crystal structure of EhRacC(Q/L)·GTP/EhPAK4 PBD was modeled using data extending to 2.85 Å resolution. Refinement is incomplete, but current R_{work} and R_{free} values are 19.3% and 27%, respectively. The *E. histolytica* complex is superimposed with the semi-transparent crystal structure of human Rac3/PAK4 PBD (PDB access. 2OV2). The signature β-harpin of the PBD is conserved, although the β-strands are aligned slightly differently. The C-terminal α-helix is rotated ~90° in the *E. histolytica* complex compared to other PBDs of known structure.
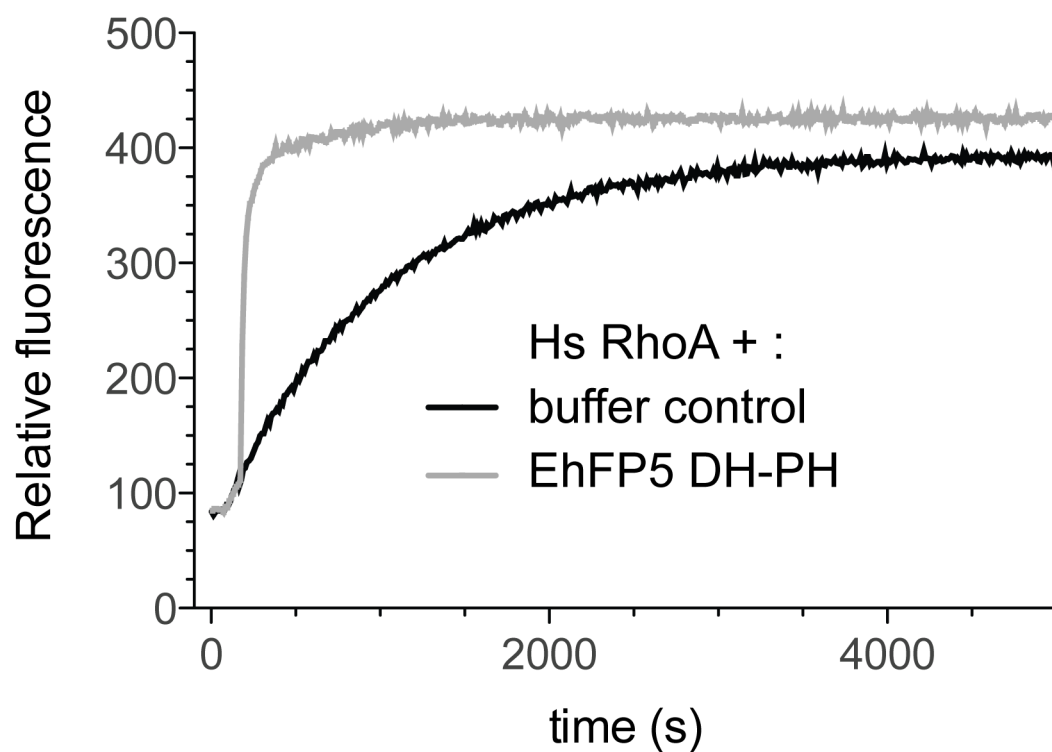
Figure 7.5. The DH-PH tandem of EhFP5 stimulates nucleotide exchange on human RhoA. Although EhFP5 GEF activity was not detected using multiple *E. histolytica* Rho GTPase substrates (unpublished data), nucleotide exchange on human RhoA was accelerated by 10 μM EhFP5 DH-PH, as measured by a fluorescent bodipy-GDP exchange assay.
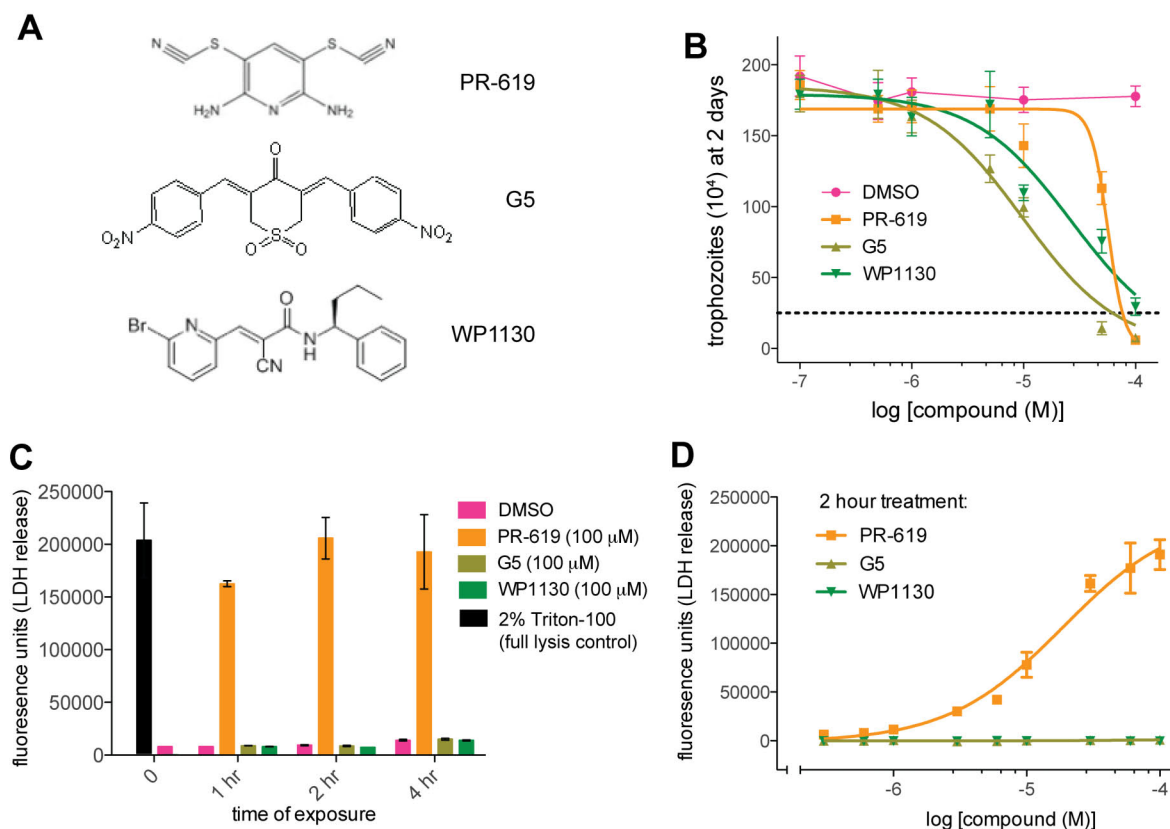
Figure 7.6. De-ubiquitinating enzyme inhibitors impair trophozoite proliferation. (A) Three broad-specificity DUB inhibitors were profiled for effects on *E. histolytica* proliferation and cell death. PR-619 contains two azido groups, suggestive of promiscuous covalent reactivity. **(B)** All three DUB inhibitors altered trophozoite proliferation following 2 days of exposure, although PR-619 was only effective at 50-100 μM and exhibited an atypical concentration-response curve. G5 and WP1130 exhibited IC_{50} values of 9.7 ± 3.0 μM and 26 ± 16 μM, respectively. A dotted line indicates initial seeding of cultures with 25 x 10^4 amoebae. **(C, D)** PR-619, but not G5 or WP1130, was cytotoxic to trophozoites as measured by a membrane integrity assay. Under these conditions PR-619 exhibited an LD_{50} of ~10 μM.
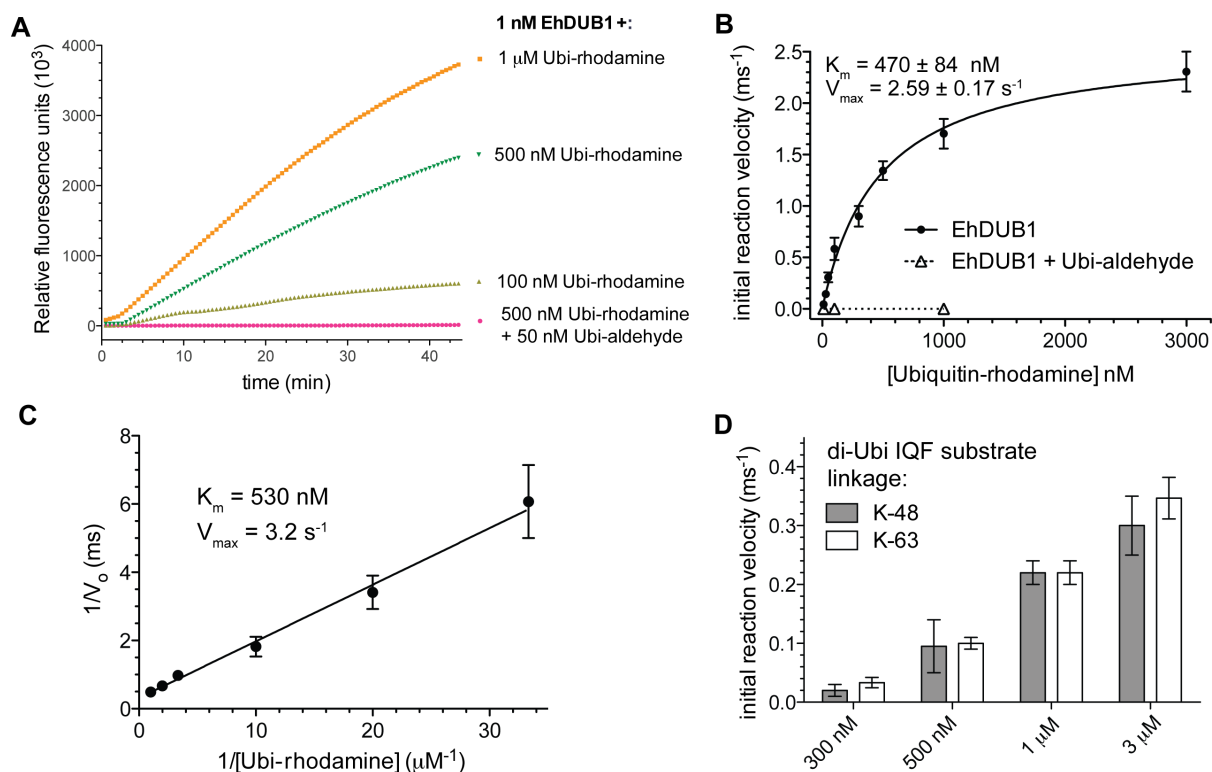
Figure 7.6. EhDUB1 hydrolyzes ubiquitin isopeptide bonds. (A) Recombinant purified EhDUB1 hydrolyzes a dye conjugated human ubiquitin substrate, and is inhibited by the "suicide substrate" ubiquitin-aldehyde. **(B)** A Michaelis-Menten kinetic analysis of EhDUB1 revealed a K_m value of ~500 nM with respect to the ubiquitin-rhodamine substrate. **(C)** An Eadie-Hofstee plot of the kinetic data gave consistent estimates of K_m and V_{max}. **(D)** EhDUB1 hydrolyzed K-48 and K-63 linked di-ubiquitin substrates with similar efficiency, as measured by an internally quenched fluorescence (IQF) assay.

7.4 REFERENCES

1. Haque, R., C.D. Huston, M. Hughes, E. Houpt, and W.A. Petri, Jr., *Amebiasis.* N Engl J Med, 2003. **348**(16): p. 1565-73.

2. Lofmark, S., C. Edlund, and C.E. Nord, *Metronidazole is still the drug of choice for treatment of anaerobic infections.* Clin Infect Dis, 2010. **50 Suppl 1**: p. S16-23.

3. Bansal, D., N. Malla, and R.C. Mahajan, *Drug resistance in amoebiasis.* Indian J Med Res, 2006. **123**(2): p. 115-8.

4. Debnath, A., et al., *A high-throughput drug screen for Entamoeba histolytica identifies a new lead and target.* Nat Med, 2012. **18**(6): p. 956-60.

5. Wani, M.Y., et al., *Synthesis and in vitro evaluation of novel tetrazole embedded 1,3,5-trisubstituted pyrazoline derivatives as Entamoeba histolytica growth inhibitors.* Eur J Med Chem, 2012. **54**: p. 845-54.

6. Mushtaque, M., F. Avecilla, and A. Azam, *Synthesis, characterization and structure optimization of a series of thiazolidinone derivatives as Entamoeba histolytica inhibitors.* Eur J Med Chem, 2012. **55**: p. 439-48.

7. Gilchrist, A., *Modulating G-protein-coupled receptors: from traditional pharmacology to allosterics.* Trends Pharmacol Sci, 2007. **28**(8): p. 431-7.

8. Rask-Andersen, M., M.S. Almen, and H.B. Schioth, *Trends in the exploitation of novel drug targets.* Nat Rev Drug Discov, 2011. **10**(8): p. 579-90.

9. Khan, M.A. and P.C. Sen, *Effect of histamine and serotonin on phagocytic activity of Entamoeba histolytica.* Indian J Exp Biol, 1988. **26**(9): p. 730-1.

10. Khan, M.A. and P.C. Sen, *Modulation of pathogenicity of Entamoeba histolytica by histamine.* Indian J Med Res, 1988. **88**: p. 225-7.

11. Khan, M.A. and P.C. Sen, *Demonstration of histamine receptors on the surface of Entamoeba histolytica.* Indian J Exp Biol, 1989. **27**(1): p. 96-7.

12. Khan, M.A., P.C. Sen, and B. Mishra, *Enhancement of virulence of Entamoeba histolytica by histamine in vitro.* Indian J Exp Biol, 1990. **28**(4): p. 376-7.

13. Coppi, A., S. Merali, and D. Eichinger, *The enteric parasite Entamoeba uses an autocrine catecholamine system during differentiation into the infectious cyst stage.* J Biol Chem, 2002. **277**(10): p. 8083-90.

14. Lu, G., Z. Wang, A.M. Jones, and E.N. Moriyama, *7TMRmine: a Web server for hierarchical mining of 7TMR proteins.* BMC Genomics, 2009. **10**: p. 275.

15.   Sun, T.J. and P.N. Devreotes, *Gene targeting of the aggregation stage cAMP receptor cAR1 in Dictyostelium.* Genes Dev, 1991. **5**(4): p. 572-82.

16.   Wetter, J.A., C. Revankar, and B.J. Hanson, *Utilization of the Tango beta-arrestin recruitment technology for cell-based EDG receptor assay development and interrogation.* J Biomol Screen, 2009. **14**(9): p. 1134-41.

17.   Minic, J., et al., *Functional expression of olfactory receptors in yeast and development of a bioassay for odorant screening.* FEBS J, 2005. **272**(2): p. 524-37.

18.   Bosch, D.E., et al., *Heterotrimeric G-protein signaling is critical to pathogenic processes in Entamoeba histolytica.* PLoS Pathog, 2012. **8**(11): p. e1003040.

19.   Diamond, L.S., *Axenic cultivation of Entamoeba hitolytica.* Science, 1961. **134**(3475): p. 336-7.

20.   Pires-Santos, G.M., K.G. Santana-Anjos, and M.A. Vannier-Santos, *Optimization of Entamoeba histolytica culturing in vitro.* Exp Parasitol, 2012. **132**(4): p. 561-5.

21.   Leschner, J., et al., *Interruption of the ionic lock in the bradykinin b2 receptor results in constitutive internalization and turns several antagonists into strong agonists.* J Pharmacol Exp Ther, 2013. **344**(1): p. 85-95.

22.   Anantharaman, V., S. Abhiman, R.F. de Souza, and L. Aravind, *Comparative genomics uncovers novel structural and functional features of the heterotrimeric GTPase signaling system.* Gene, 2011. **475**(2): p. 63-78.

23.   Bosch, D.E., et al., *Structural Determinants of RGS-RhoGEF Signaling Critical to Entamoeba histolytica Pathogenesis.* Structure, 2013. **21**(1): p. 65-75.

24.   Johnston, C.A., M.D. Willard, A.J. Kimple, D.P. Siderovski, and F.S. Willard, *A sweet cycle for Arabidopsis G-proteins: Recent discoveries and controversies in plant G-protein signal transduction.* Plant Signal Behav, 2008. **3**(12): p. 1067-76.

25.   Dutta, S., A. Sardar, D. Ray, and S. Raha, *Molecular and functional characterization of EhPAK3, a p21 activated kinase from Entamoeba histolytica.* Gene, 2007. **402**(1-2): p. 57-67.

26.   Labruyere, E., V. Galy, P. Sansonetti, and N. Guillen, *Distribution of a potential p21-activated serine/threonine kinase (PAK) in Entamoeba histolytica.* Arch Med Res, 2000. **31**(4 Suppl): p. S128-30.

27.   Labruyere, E., C. Zimmer, V. Galy, J.-C. Olivo-Marin, and N. Guillen, *EhPAK, a member of the p21-activated kinase family, is involved in the control of Entamoeba histolytica migration and phagocytosis.* Journal of Cell Science, 2003. **116**: p. 61-71.

28. Eswaran, J., M. Soundararajan, R. Kumar, and S. Knapp, *UnPAKing the class differences among p21-activated kinases.* Trends Biochem Sci, 2008. **33**(8): p. 394-403.

29. Kumar, A., et al., *PAK thread from amoeba to mammals.* J Cell Biochem, 2009. **107**(4): p. 579-85.

30. Baskaran, Y., Y.W. Ng, W. Selamat, F.T. Ling, and E. Manser, *Group I and II mammalian PAKs have different modes of activation by Cdc42.* EMBO Rep, 2012. **13**(7): p. 653-9.

31. Lei, M., et al., *Structure of PAK1 in an autoinhibited conformation reveals a multistage activation switch.* Cell, 2000. **102**(3): p. 387-97.

32. Ong, C.C., et al., *p21-activated kinase 1: PAK'ed with potential.* Oncotarget, 2011. **2**(6): p. 491-6.

33. Deacon, S.W., et al., *An isoform-selective, small-molecule inhibitor targets the autoregulatory mechanism of p21-activated kinase.* Chem Biol, 2008. **15**(4): p. 322-31.

34. Viaud, J. and J.R. Peterson, *An allosteric kinase inhibitor binds the p21-activated kinase autoregulatory domain covalently.* Mol Cancer Ther, 2009. **8**(9): p. 2559-65.

35. Arias-Romero, L.E., et al., *EhPAK2, a novel p21-activated kinase, is required for collagen invasion and capping in Entamoeba histolytica.* Mol Biochem Parasitol, 2006. **149**(1): p. 17-26.

36. Aurrecoechea, C., et al., *AmoebaDB and MicrosporidiaDB: functional genomic resources for Amoebozoa and Microsporidia species.* Nucleic Acids Res, 2011. **39**(Database issue): p. D612-9.

37. Bosch, D.E., B. Yang, and D.P. Siderovski, *Entamoeba histolytica Rho1 regulates actin polymerization through a divergent, diaphanous-related formin.* Biochemistry, 2012. **51**(44): p. 8791-801.

38. Wittinghofer, A. and I.R. Vetter, *Structure-function relationships of the G domain, a canonical switch motif.* Annu Rev Biochem, 2011. **80**: p. 943-71.

39. Abdul-Manan, N., et al., *Structure of Cdc42 in complex with the GTPase-binding domain of the 'Wiskott-Aldrich syndrome' protein.* Nature, 1999. **399**(6734): p. 379-83.

40. Nakada-Tsukui, K., H. Okada, B.N. Mitra, and T. Nozaki, *Phosphatidylinositol-phosphates mediate cytoskeletal reorganization during phagocytosis via a unique modular protein consisting of RhoGEF/DH and FYVE domains in the parasitic protozoon Entamoeba histolytica.* Cell Microbiol, 2009. **11**(10): p. 1471-91.

41.   Worthylake, D.K., K.L. Rossman, and J. Sondek, *Crystal structure of Rac1 in complex with the guanine nucleotide exchange region of Tiam1.* Nature, 2000. **408**(6813): p. 682-8.

42.   Chihara, K., et al., *Cytoskeletal rearrangements and transcriptional activation of c-fos serum response element by Rho-kinase.* J Biol Chem, 1997. **272**(40): p. 25121-7.

43.   Lee, M.J., B.H. Lee, J. Hanna, R.W. King, and D. Finley, *Trimming of ubiquitin chains by proteasome-associated deubiquitinating enzymes.* Mol Cell Proteomics, 2011. **10**(5): p. R110 003871.

44.   Colland, F., *The therapeutic potential of deubiquitinating enzyme inhibitors.* Biochem Soc Trans, 2010. **38**(Pt 1): p. 137-43.

45.   Burroughs, A.M., L.M. Iyer, and L. Aravind, *Structure and evolution of ubiquitin and ubiquitin-related domains.* Methods Mol Biol, 2012. **832**: p. 15-63.

46.   Borodovsky, A., et al., *Chemistry-based functional proteomics reveals novel members of the deubiquitinating enzyme family.* Chem Biol, 2002. **9**(10): p. 1149-59.

47.   Altun, M., et al., *Activity-based chemical proteomics accelerates inhibitor development for deubiquitylating enzymes.* Chem Biol, 2011. **18**(11): p. 1401-12.

48.   Hassiepen, U., et al., *A sensitive fluorescence intensity assay for deubiquitinating proteases using ubiquitin-rhodamine110-glycine as substrate.* Anal Biochem, 2007. **371**(2): p. 201-7.

49.   Hu, M., et al., *Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde.* Cell, 2002. **111**(7): p. 1041-54.

50.   Bozza, W.P., Q. Liang, P. Gong, and Z. Zhuang, *Transient Kinetic Analysis of USP2-Catalyzed Deubiquitination Reveals a Conformational Rearrangement in the K48-Linked Diubiquitin Substrate.* Biochemistry, 2012. **51**(50): p. 10075-86.