2007

A Structural Basis for Host Cytoskeletal Disruption and Virulence by Yersinia Protein Kinase A

Gerd Prehna

A STRUCTURAL BASIS FOR HOST CYTOSKELETAL DISRUPTION AND VIRULENCE BY *YERSINIA* PROTEIN KINASE A

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Gerd Prehna

June 2007

A STRUCTURAL BASIS FOR HOST CYTOSKELETAL DISRUPTION AND VIRULENCE BY *YERSINIA* PROTEIN KINASE A

Gerd Prehna, Ph.D.

The Rockefeller University 2007

Yersinia spp. cause gastroenteritis and the plague, representing historically devastating pathogens that are currently an important biodefense and antibiotic resistance concern. Although several antibiotic therapies exist, the emergence of strains that have garnered multiple drug resistances in combination with the weaponization of *Yersinia*, make understanding the biology of this pathogen a high priority. *Yersinia*, along with other pathogenic bacteria such as *Salmonella*, utilize a macromolecular complex, called a type III secretion apparatus, to deliver virulence proteins directly into cells. These factors commandeer several signaling pathways, often targeting the Rho family of small GTPases which regulate actin cytoskeletal dynamics. A critical virulence determinant in *Yersinia* species is the *Yersinia* protein kinase A, or YpkA, a multi-domain protein that disrupts the eukaryotic actin cytoskeleton. YpkA contains a Ser/Thr kinase domain whose activity modulates pathogenicity and a domain that binds to both Rac1 and RhoA of the Rho family of small GTPases. The crystal structure of a YpkA-Rac1 complex reveals that YpkA possesses a novel Rac1-binding domain that mimics the interactions of host guanine nucleotide dissociation inhibitors (GDIs) of the Rho GTPases. YpkA

inhibits the exchange of nucleotide in Rac1 and RhoA, and mutations that disrupt the YpkA-GTPase interface abolish this activity *in vitro* and significantly impair *in vivo* YpkA-induced cytoskeletal disruption. A *Yersinia pseudotuberculosis* mutant lacking the GDI activity of YpkA was significantly attenuated for virulence in a mouse infection assay as compared to wild type bacteria. We conclude that virulence in *Yersinia* depends strongly upon a novel mimicry of host GDI proteins by YpkA. Finally, the YpkA kinase domain has homology to known eukaryotic Ser/Thr kinases and thus could be targeted for small molecule inhibitor design. An efficient approach integrating a machine learning method, homology modeling, and multiple conformational high throughput docking was used for the discovery of YpkA inhibitors. The resultant small molecule compounds, which are the first reported inhibitors for YpkA, not only provide a useful means in probing the function and mechanism of YpkA in bacterial pathogenesis, but also are potential candidates for further development of novel anti-plague drugs.

BIOGRAPHICAL SKETCH

Gerd Prehna was born on April 10, 1978 in Stamford, New York. His father is Gerd John Prehna, a machinist and Harley Davidson motorcycle restorationist. His mother is Donna Jane Prehna, a teacher's assistant in special education. In 1996 he graduated from Stamford Central School in Stamford, New York, shortly after which he studied abroad in Denia, Spain as an exchange student with the Rotary Club. Upon his return, he attended college at the University of Rochester, in Rochester, New York. There he pursued studies in biology and chemistry, while spending several semesters and summers as a research assistant in the laboratory of biochemistry and biophysics of Barry M. Goldstein. He received his B.S. in biochemistry from the University of Rochester in 2001. That same year he enrolled as a graduate student at the Rockefeller University in New York City. There he completed the work described in this thesis under the guidance of crystallographer and microbiologist, C. Erec Stebbins.

PUBLICATIONS

1. Pankiewicz KW, Lesiak-Watanabe KB, Watanabe KA, Patterson SE, Jayaram HN, Yalowitz JA, Miller MD, Seidman M, Majumdar A, <u>Prehna G</u>, Goldstein BM. (2002). "Novel mycophenolic adenine bis(phosphonate) analogues as potential differentiation agents against human leukemia." *J Med Chem.* Jan 31;45(3):703-12.

2. Ceccarelli C, Liang ZX, Strickler M, <u>Prehna G</u>, Goldstein BM, Klinman JP, Bahnson BJ. (2004) "Crystal structure and amide H/D exchange of binary complexes of alcohol dehydrogenase from Bacillus stearothermophilus: insight into thermostability and cofactor binding." *Biochemistry.* May 11;43(18):5266-77.

3. <u>Prehna G</u>, Ivanov M, Bliska JB, Stebbins CE (2006). "*Yersinia* Virulence Depends on Mimicry of Host Rho-family Nucleotide Dissociation Inhibitors." *Cell.* 2006 126: 869-880.

4. <u>Prehna G</u>, Stebbins CE (2007). "A Rac1-GDP Trimer Complex Binds Zinc with Tetrahedral and Octahedral Coordination Displacing Magnesium." *Acta Crystallographica D.* 2007 D63, 628-635.

To my parents; my father, Gerd John Prehna and my mother, Donna Jane Prehna for

always supporting my creativity.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Å	Angstrom
ATP	adenosine triphosphate
DTT	dithiothreitol
EDTA	ethylenediaminotetraacetic acid
FPLC	fast performance liquid chromatography
GAP	GTPase activating protein
GDI	guanine nucleotide dissociation inhibitor
GEF	guanine nucleotide exchange factor
GST	glutathion-S-transferase
GTP	guanine triphosphate
HEPES	4-(2-hydroxyethyl)-piperazine-1-ethanesulfonic acid
IPTG	isopropylthiogalacopyranoside
kD	kilodalton
LB	Luria broth
M	molar
MES	2-(N-Morpholino)ethanesulfonic acid
mM	millimolar
μM	micromolar
NaCl	sodium chloride
PCR	polymerase chain reaction
PEG	polyethylene glycol
PMSF	phenylmethylsulfonyl fluoride
SDS	sodium dodecyl sulfate
Syc	specific Yop chaperone
T3SS/A	type III secretion system/apparatus
Tris	tris-(hydroxymethyl)-aminomethane
YpkA	*Yersinia* protein kinase A
Yop	*Yersinia* outer protein
Ysc	*Yersinia* secretion

CHAPTER ONE:

INTRODUCTION

1.1 Host-Bacteria Relationships

Prokaryotic life on earth is estimated to have appeared more than 3.5 billion years ago, and not only is the progenitor of all other organisms, but created the environmental conditions that made eukaryotic life possible (Salyers 2002). Even today, no section of the earth is free of prokaryotic life, and in fact, more than 50% of the entire biomass of the planet is attributed to bacteria. These organisms occupy varying and diverse niches and have evolved an astonishing array of metabolic processes that allow them to flourish in every conceivable environmental condition. Bacterial species have been found to culture in arctic ice (cryophiles), in the heat vents on the ocean floor (thermophiles), and to persist in otherwise toxic conditions, such as extreme pH levels, that would undoubtedly kill eukaryotic life forms. Although bacteria have adapted to life in exotic environmental conditions, as demonstrated by the omnipresence of these organisms, they have also evolved to take advantage of the environments provided by eukaryotic life. In particular, bacteria species have co-evolved with animal species to create a complex and delicate interplay of symbiotic and disease relationships within the bodies and on the surface of their hosts.

Eukaryotic symbiosis with prokaryotic life can be first traced back to the absorption of prokaryotes by eukaryotes to form endosymbiotes, such as the mitochondria and the chloroplast (Salyers 2002). In contrast, on the scale of a

multicellular organism we find that entire bacterial colonies persist on all surfaces of the body, such as the skin and eyes, occupy orifices, and the endothelium of the intestinal tract. In the case of humans, most bacterial species are commensals, whose populations vary depending upon physiological factors, and are not essential for life nor very often cause disease, though they can play a role in maintaining good health (Brooks 2001). The most prominent example of such flora, are the 300-500 bacterial species that inhabit the human gut, which can serve metabolic, trophic, and immunological functions (Guarner and Malagelada 2003). The metabolic role of the bacterial flora is to ferment normally non-digestible carbohydrates such as cellulose, to salvage energy in the production of short-chain fatty acids (SCFAs), to synthesize vitamin K, and to aid in the absorption of ions such as calcium, magnesium, and iron (Brooks 2001; Guarner and Malagelada 2003; Dale and Moran 2006). The resulting SCFAs have a trophic effect on the endothelial lining which serves to regulate cell proliferation in the colon. These bacterial colonies in the gut serve to condition the mucosal immune architecture of the intestine and also create a "barrier effect" protecting the host from infection (Guarner and Malagelada 2003). Simply put, the already established symbiotic colonies within the gut act to out-compete the ability of many potential pathogenic bacteria to proliferate. Finally, although mostly benign, if not helpful, some disease conditions can arise with bacterial symbiotes. This often results in an inflammatory bowel disease, such as Crohn's disease, where the immune system has become hypersensitive to the antigens within the intestinal lumen. Additionally, the over use of some antibiotics can drastically change the make-up of the flora weakening the "barrier effect", allowing opportunistic infections to occur (Brooks 2001; Guarner and Malagelada 2003).

While such the bacterial species have evolved in symbiosis with the host, others have evolved disease relationships with animal species. Such pathogenic bacteria have played an "arms race" with their host, evolving counter-strategies against immunological defenses to establish an infection. These bacteria often have evolved pathogenicity islands consisting of a collection of genes that encode virulence factors that interfere with specific host cell functions and act to help evade the immune system. These genes can be part of the normal bacterial genome, or can be encoded plasmids that can be transmitted between bacterial species. Disease causing bacteria can be found in the normal flora, such as *Streptococcus pneumoniae* and *Salmonella typhi,* ingested from contaminated food or water, such as *Yersinia entercolitica* and *Salmonella typhimurium*, and transmitted from animal to animal, even if not the same species, such as *Bacillus anthracis* and *Yersinia pestis* (Brooks 2001; Salyers 2002; Dale and Moran 2006). *B. anthracis* can be contracted due to contact with spores associated with farm animal waste, and is thus prevalent in agriculture, whereas in the case of *Y. pestis*, infection is introduced from the bite of a flea infected from the indigenous rodent population. These pathogenic bacteria have evolved a number of intricate molecular mechanisms of infection to both persist in the host and to effectively deliver their virulence factors. For the research presented here, the pathogenic mechanisms of *Yersinia spp*. will be explored in detail.

1.2 *Yersinia* Pathogenesis and Antibiotic Resistance

The *Yersinia* genus, named for its characterization by Alexandre Yersin, is a gram-negative bacteria that consists of 11 species that are found on every continent, apart from Australia (Perry and Fetherston 1997). Out of these eleven species, three cause disease in animal species, namely *Y. entercolitica*, *Y. pseudotuberculosis*, and *Y. pestis* (Perry and Fetherston 1997; Wren 2003). The former two species are the agents of gastroenteritis, whereas the later, *Y. pestis* is the causative agent of the plague. The enteropathogenic species of *Yersinia* can be found in soil samples, and infect several mammalian and animal species, including humans. Specifically, gastroenteritis is caused after exposure to contaminated food or water. The pathogenic bacteria migrate across Peyer's patches in the small intestine, then subsequently to lymph nodes, the liver, and the spleen (Wren 2003). Normally such an infection is localized, cleared within a few days, and is non-lethal in humans. In contrast, *Y. pestis*, is likely the most devastating bacterial pathogen in human history.

Nearly 200 million people are estimated to have died in the plague epidemics that devastated the ancient world (Perry and Fetherston 1997). This was witnessed through two pandemics: the first pandemic, or the Justinian plague from 541AD to 544AD, and the second pandemic, consisting of both the "Black Death" and subsequent epidemics during the 8th to 14th centuries, with an overall death toll representing as much as 60% of the population of the ancient world (Perry and Fetherston 1997). Finally, the third pandemic, starting from 1855 to the modern day, has witnessed additional deaths totaling in the millions. *Y. pestis*, unlike its enteropathogenic counterparts, has evolved a life

cycle between a mammalian and insect host, which is necessary for the propagation of the plague. To transmit the bacterium, a flea first feeds on an infected host, which is in most cases a rodent, where it contracts *Y. pestis* from the blood. The bacteria persist in the flea mid-gut where they form a biofilm within a few days that prevents proper digestion. The starving flea then transmits the bacteria to additional mammalian hosts by attempting to feed as often as possible, effectively injecting *Y. pestis* with each bite. From the sites of feeding, the bacteria migrate to lymph nodes, where they target and are initially taken up by polymorphonuclear leukocytes (PMNs) and monocytes (Marketon, DePaolo et al. 2005). This localized infection gives rise to the "Bubonic Plague" which is characterized by oversized, swollen, and black lymph nodes, or "bubos." From these sites, or from initial injection by the fleabite, *Y. pestis* can migrate to the blood stream and localizes in the various tissues of the body causing septicemic plague. This event is highly lethal, ranging in 30% to 50% death rates. Finally, *Y. pestis* can become pneumonic if it localizes to the lungs, which occurs in approximately 2% of cases. This form of the plague is not only close to 100% lethal, but results in a highly infectious airborne transmission of the bacterium through respiratory droplets (Perry and Fetherston 1997; Wren 2003).

Genetically, the three pathogenic species of *Yersinia* are quite diverse, with *Y. pseudotuberculosis* and *Y. pestis* being closely related, and *Y. entercolitica* distantly related. Although dissimilar at many levels, *Y. enterocolitica*, *Y. pseudotuberculosis*, and *Y. pestis* all share the highly conserved *Yersinia* virulence plasmid, or pYV, which is necessary to cause disease (Cornelis, Boland et al. 1998). This 70kb plasmid harbors numerous genes, a large number of which are associated with a type III protein secretion

system (T3SS) that confers the ability of diverse pathogens to deliver virulence factors into host cells (Galan and Collmer 1999). This plasmid not only includes the structural and functional genes for the T3SS, but additional virulence factors known as *Yersinia* Outer Proteins (Yops), their chaperones (Syc) and other members that are characteristics of the low calcium response (LCRS) (see Section 1.3 and Section 1.4).

Y. pestis is evolutionarily the newest version of *Yersinia*, having branched from y. pseudotuberculosis 1,500 to 20,000 years ago(Wren 2003). As discussed above, *Y. pestis* is far more virulent than its counterparts, which in part can be attributed to the addition of two plasmids to its genome. *Y. pestis* also contains the 96.3kb plasmid pMT1 and the 9.6kB plasmid pPla (Perry and Fetherston 1997; Wren 2003; Chromy, Choi et al. 2005). The pMT1 plasmid contains proteins that make up and regulate the expression of the F1 capsule, which helps to block adhesion by macrophages, and the gene Yet (murine toxin), which is a member of the phospholipase D family and is essential for the colonization of the flea mid-gut and thus the mammal and insect plague life cycle (Perry and Fetherston 1997; Wren 2003; Chromy, Choi et al. 2005). pPla encodes the plasminogen activator protein, Pla, which is a multifunctional protease. It shares high homology to known plasminogen activators, is known to cleave plasminogen and other matrix proteins, and in addition cleaves the complement component C3 (Chromy, Choi et al. 2005). This activity is thought to be necessary for the dissemination of *Y. pestis* after initial infection, as the removal of pla causes bacteria that are introduced at the subcutaneous level to be rendered non-virulent, but still be able to cause disease if introduced directly into the blood stream (Wren 2003; Chromy, Choi et al. 2005). In addition to its three virulence plasmids, *Y. pestis* also harbors genes on its 4.65mb chromosome that are essential for

virulence. Many of these genes are putative and are suspected to be involved in the insect life cycle. One such example is the *hmn* locus (haemin storage), which is necessary for altering the feeding habits of infected fleas and blockage of their digestion (Perry and Fetherston 1997). Finally, in addition to the acquisition of various genes, the virulence of *Y. pestis* relative to *Y. pseudotuberculosis* can be attributed to the loss of gene function. Specifically, *Y. pseudotuberculosis* contains active genes that express insecticidal toxin, baculovirus enhancin, and other putative insect toxins for the parasitism of insects, whereas many such genes are only present as pseudogenes in *Y. pestis* (Wren 2003). This would allow *Y. pestis* to adapt to an insect host as part of its life style as infection would not be as lethal initially.

Although plague is extremely virulent, modern antiseptic and hygienic practices in combination with antibiotics make it treatable and controllable. According to the World Health Organization, most plague cases occur in areas with a high endemic rodent population, such as the American southwest and Africa. On average, a couple of thousand cases are reported each year with an approximate death rate of 10%. Despite the lack of an effective vaccine, a number of antibiotics exist that can treat the plague. Upon diagnosis of such cases, the common practice is to administer streptomycin for the first couple days followed by tetracycline (Perry and Fetherston 1997). Even with such efforts, the overuse of antibiotics has given rise to medical concerns over the evolution of multi-drug resistant strains of the plague bacterium that have been reported in several locations from patient isolates (Galimand, Guiyoule et al. 1997; McCormick 1998). Of even more concern, are reports of these strains gaining their resistances from horizontal transfer interactions with other known resistant pathogens such as *E. coli* (Galimand,

Guiyoule et al. 1997; Galimand, Carniel et al. 2006). One reported *Y. pestis* strain became resistant to ampicillin, chloramphenicol, kanamycin, sulfonamide, streptomycin, and tetracycline through the addition of a 150kb plasmid designated pIP1202 (Galimand, Carniel et al. 2006). Additionally, with the advent of drug resistant *Y. pestis* stains and the successful weaponization of plague in the former Soviet Union bioweapons program, this pathogen has become a primary biodefense concern (Henderson 1999; Inglesby, Dennis et al. 2000). The emergence of antibiotic resistance strains and the threat of such bacteria as biological weapons, suggests the need for the immediate design of new plague antibiotics. The most attractive and effective approach for these new drugs would most likely be to target the various specific *Yersinia* virulence proteins (see section 1.3 and section 1.4). By targeting the critical components of the *Yersinia* pathogen new plague specific antibiotics are unlikely to evolve resistances to existing therapies, gain resistances to the target-specific designed drugs, or induce cross-resistance with other species (Frean, Arntzen et al. 1996; Henderson 1999; Frean, Klugman et al. 2003; Marra 2006).

1.3 The Type III Secretion System

As mentioned previously, pathogenic *Yersinia* contain an indispensable and highly conserved virulence plasmid, or pYV, which is necessary to cause disease (Cornelis, Boland et al. 1998). This plasmid harbors numerous genes, a large number of which are associated with a type III protein secretion system (T3SS) which confers the
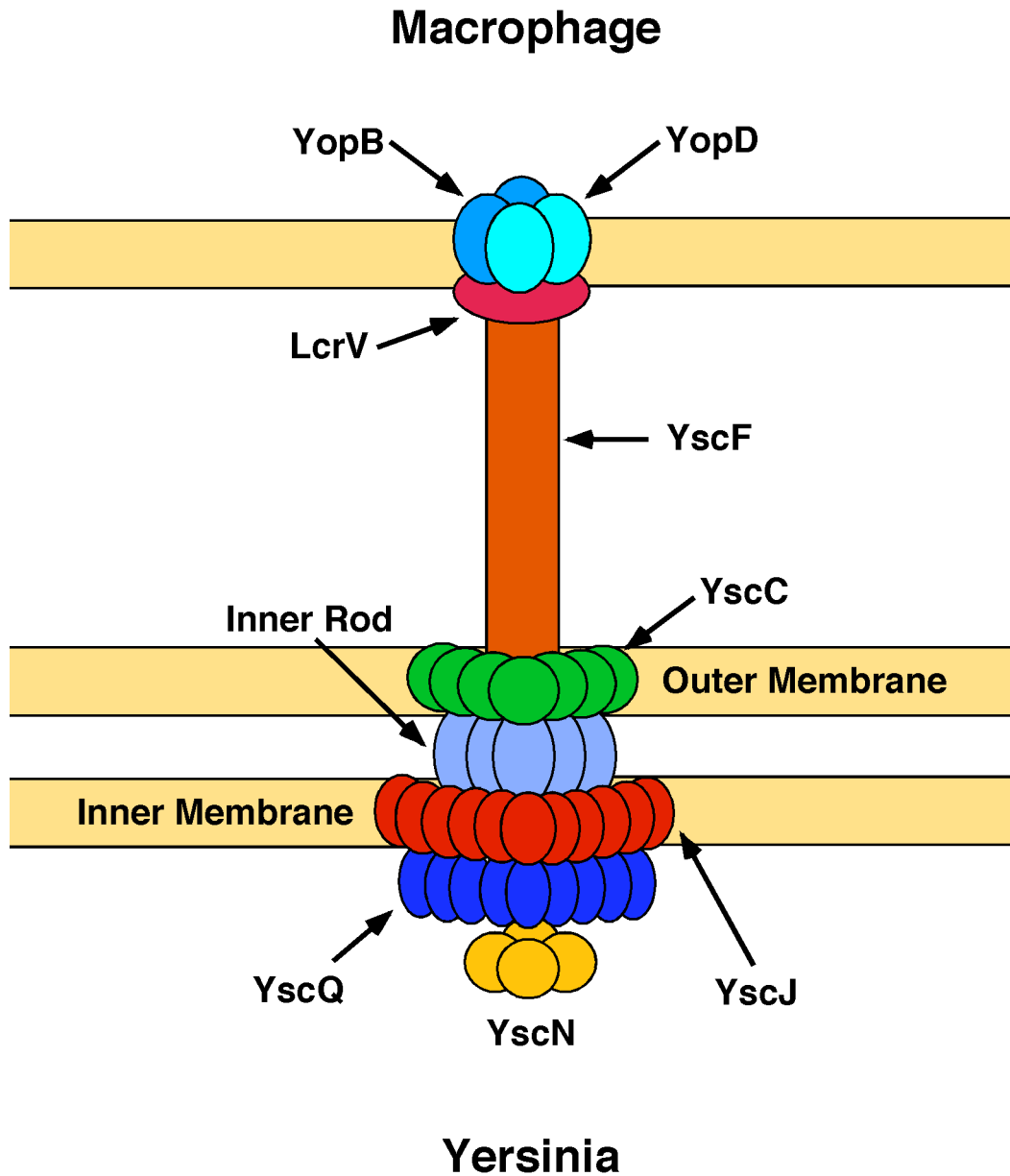
Figure 1.1: The Type III Secretion System

A cartoon representation of the various components of the T3SS in *Yersinia* is shown. See text for details on the various components.

ability of diverse pathogens to deliver virulence factors into host cells (Galan and Collmer 1999). These genes, termed *Yersinia* Secretion proteins (Ysc), consist of approximately 27 different members which assemble into a macromolecular complex that is literally a syringe-like structure that injects the virulence factors, *Yersinia* Outer Proteins (Yops) into a host cell (Cornelis 2002; Cornelis 2002; Yip and Strynadka 2006) (Figure 1.1). This structure is highly homologous and evolutionarily related to the flagellar system. During an infection with T3SS containing organisms, the bacteria construct syringe-like structures that are readily visible by electron microscopy (Galan and Collmer 1999; Yip, Kimbrough et al. 2005). As shown in Figure 1.1, the structure consists of a base that sits in the inner membrane, another circular section crossing the outer membrane, a needle complex, and finally the secretion of a pore-like structure to penetrate the host cell membrane. Overall, the needle length is estimated to be close to 600Å long, with an inner channel of 20 to 30Å in diameter. As such, this narrow width implies the caveat that any globular protein passing through the system would require at least some degree of unfolding, for which there is indirect evidence (Cornelis 2002; Stebbins and Galan 2003; Yip and Strynadka 2006). The base structure consists of the proteins YscJ, which in analogy to EscJ crystal structure from E. coli, oligomerizes into a ring consisting of 24 subunits at the inner membrane (Yip, Kimbrough et al. 2005), YscQ which also by analogy forms antiparallel β-sheets in a channel like manner (Fadouloglou, Tampakaki et al. 2004): and YscN, which has homology to F_0F_1 ATPases, and has measured ATPase activity (Cornelis 2002). It is this last protein that recognizes effectors bound to their chaperones (Syc proteins) and provides the energy for Yop secretion in an ATP dependent manner (Akeda and Galan 2005). Additionally, the platform of the type

III secretion apparatus contains the membrane anchoring proteins, YscR, YscS, YscT, YscU, and YscV, although their exact architecture is currently unknown (Yip and Strynadka 2006). The outermembrane ring of the complex is formed by YscC, which oligomerizes into 12-14 subunits with a diameter of 14 to 20nm, as judged by electron microscopy, which serves to anchor and stabilize the outgoing needle complex (Burghout, van Boxtel et al. 2004). The needle itself is formed by the protein YscF that extends from the base of the T3SS to the host cell. The inner diameter measures ~25Å and the overall length of the needle is regulated by the protein YscP (Cornelis 1998; Agrain, Sorg et al. 2005; Yip and Strynadka 2006). Finally, the needle is capped by LcrV, which serves as a bell-shaped extension to the needle which is necessary for pore formation in the host membrane along with the formation of the translocation pore, YopB and YopD (Cornelis 2002; Derewenda, Mateja et al. 2004).

Due to the size restrictions associated by the needle complex as referenced above, another critical component of the T3SS, are the molecular chaperones, or Syc proteins. These proteins recognize motifs on the various Yops, and prime them for unfolding and delivery by the T3SS (Cornelis 2002; Stebbins and Galan 2003; Ghosh 2004). Each Yop is associated with a chaperone, which binds to a recognition region in the N-terminus of the protein (Cornelis 2002; Lilic, Vujanac et al. 2006). Interestingly, some Yops have their own chaperone, for example SycE which recognizes YopE, whereas other chaperones can bind multiple Yops individually or at the same time, such as SycD for both YopB and YopD (Cornelis 2000; Cornelis 2002). At the structural level, the binding of virulence factors to their chaperones in any T3SS, maintains a portion of the virulence factor in a non-globular conformation, and wraps around the Syc proteins (Stebbins and

Galan 2001) (Birtalan, Phillips et al. 2002). Additionally, it is this region of the Yop that encodes a recognition motif, or β-motif, that in part directs the virulence factor to the chaperone (Lilic, Vujanac et al. 2006). Once delivered to the T3SS, YscN then releases the virulence protein from its chaperone, again in an ATP dependent manner, (Akeda and Galan 2005) and provides the additional energy for the unfolding and secretion of the virulence protein so that it may pass through the narrow needle structure into the host cell.

1.4 The *Yersinia* Outer Proteins

The *Yersinia* Outer Proteins, or Yops, are named for their initial characterization as outermembrane proteins (Cornelis, Boland et al. 1998). It was later shown that they are in fact secreted, which can be induced in growth conditions of low calcium (the low calcium response). In a normal infection, these virulence determinants are directly injected into a host cell by the T3SS where they interfere with normal cell signaling with the overall effect of paralyzing macrophages, dendritic cells, and neutrophils, thus preventing a primary immune response (Marketon, DePaolo et al. 2005). This mode of infection allows *Y. pestis* to persist in the extracellular medium, largely unchallenged by host defenses. Not only are the bacteria resistant to phagocytosis by immune cells due to the action of its Yops, but they also suppress interleukin secretion and can eventually induce apoptosis (Aepfelbacher, Zumbihl et al. 1999; Cornelis 2002; Viboud and Bliska 2005)

There are nine well characterized Yops: YopB, YopD, and LcrV, which are a critical part of the T3SS, forming a pore in the host cell membrane to allow the passage of the Yops, and six others, YopH, YopM, YopJ/P, YopT, YopE, and YopO/YpkA which specifically interfere with normal host cell processes (Cornelis 2002; Viboud and Bliska 2005). Although the structures of YopB and YopD are unknown, they have been known to form ion-conduction channels and thus are predicted to oligomerize at the tip of the T3SS during a *Y.pestis* infection (Tardy, Homble et al. 1999). LcrV has also been shown to be able to form channels in lipid bilayers, but is a soluble dimer in solution as demonstrated by the work on its crystal structure (Tardy, Homble et al. 1999; Derewenda, Mateja et al. 2004). YopH is currently the most well characterized Yop, as its mode of action is known in addition to thirteen crystal structures. YopH is a 51kD protein that is an active protein tyrosine phosphatase (PTP) whose catalytic domain is highly related to known eukaryotic PTPs of its kind, sharing the conserved $C(X)_5R(S/T)$ active site motif (Viboud and Bliska 2005). YopH acts by dephosphorylating a variety of proteins in the host cell. This interferes with the attachment of *Yersinia* to $\beta 1$ integrins of the macrophage in addition to the suppression of the oxidative burst in phagocytes (Persson, Sjoblom et al. 2004). Specifically, YopH localizes to focal adhesions and dephosphorlyates p130Cas and Fak (focal addition kinase) which acts to inhibit phagocytosis (Weidow, Black et al. 2000). Additionally, the interference of YopH is thought to prevent the recruitment of GEFs (guanine nucleotide exchange factors) to the membrane (Viboud and Bliska 2005).

In contrast to YopH, YopM is probably the least understood of the Yops. Although the crystal structure is known, which reveals a LRR motif (Evdokimov,

Anderson et al. 2001), the exact function of YopM in *Yersinia* virulence is unknown. YopM contains 15 leucine-rich repeats and potentially binds as an adaptor protein for the kinases RSK1 and PRK2, and can traffic to the nucleus, though the purpose of which is not established (Viboud and Bliska 2005). What is known, is that without YopM *Yersinia* strains cannot establish a systemic infection, as YopM prevents clearing of the bacteria by the immune system (Trulzsch, Sporleder et al. 2004). Additionally, in *Y. pestis*, YopM acts by an unknown mechanism to deplete the Natural Killer T-cells at the sites of infection (Kerschen, Cohen et al. 2004).

The virulence factor YopJ, also known as YopP in *Y. entercolitica*, is a 288 amino acid protein that has been shown to induce apoptosis in macrophages (Cornelis 2002; Viboud and Bliska 2005). YopJ acts to inhibit both the MAPK and the NF-κβ signaling pathways resulting in the loss of cytokine production and eventual apoptosis (Zhang, Ting et al. 2005). At the biochemical level this is accomplished by the cysteine protease activity of YopJ that is proposed to act to deubiquinate signaling molecules upstream of NF-κβ, that bind to, and keep NF-κβ inactive in the cytosol. Thus, NF-κβ is prevented from entering the nucleus where it can stimulate the transcription of proinflammatory factors (Viboud and Bliska 2005). The inhibition of NF-κβ is also the preliminary step in the apoptotic response of the macrophage; thus it is thought that this inactivation triggers the macrophage apoptosis. Additionally, YopJ acts upon the MAPK signaling pathway by inactivating the MAPK kinases through the acetylation of its normally phosphorylated serine residues. This additional activity is also thought to contribute to the loss of cytokine production and the induction of apoptosis (Bliska 2006).

YopT is a 322 amino acid protein that acts as a cysteine protease that removes the Rho family of small GTPases from the cell membrane (Viboud and Bliska 2005). YopT has a core catalytic triad of residues required for activity which are conserved among a variety of pathogens and eukaryotic proteases (Shao, Merritt et al. 2002). Within the host cell, YopT cleaves RhoA, Rac1, and Cdc42 at the C-terminus removing the isoprenyl modification that anchors the small GTPase in the cell membrane, removing it from efficient signaling (Shao, Vacratsis et al. 2003). Although its overall role in pathogenesis is not fully elucidated, in cultured cells the enzymatic activity of YopT contributes to the loss of actin stress fiber formation and it has been shown to have antiphagocytic activity (Viboud and Bliska 2005).

YopE and YpkA (YopO in *Y. entercolitica*) both target the Rho family of small GTPases (Aepfelbacher and Heesemann 2001). YopE is a critical virulence factor in *Yersinia spp*, that acts as a potent GTPase Activating Protein (GAP) on the Rho family of small GTPases. Necessary for virulence, YopE is a 219 amino acid protein that contains a GAP domain that associates with RhoA, Rac1, and Cdc42 to stimulate the intrinsic GTP hydrolysis of the molecule, and thus inactivates the signaling properties of the small GTPases (Viboud and Bliska 2005). The crystal structure of YopE is known, which reveals that is shares the ability of known GAPs to introduce an "arginine finger" into the active site of the small GTPase enhancing the catalytic site of the molecule (Evdokimov, Tropea et al. 2002). In cultured cells YopE isolates to the cell membrane and causes the loss of actin stress fiber formation, cell rounding, and inhibition of phagocytosis (Aepfelbacher and Heesemann 2001). YpkA also binds to RhoA and Rac1, but not Cdc42, although the function of which was unknown. Additionally, YpkA contains a

Ser/Thr kinase domain with high homology to eukaryotic kinases whose exact function remains unclear (Viboud and Bliska 2005). YpkA is discussed in further detail in section 1.5.

From an examination of the Yops, and other virulence factors, it is apparent that they have evolved such a close relationship with their host that they have employed strategies of functional and molecular mimicry. Specifically, the various virulence factors associated with the T3SS interfere with host pathways by replacing the functions of the host regulators of those systems to drive the pathway toward the needs of the pathogen (Stebbins and Galan 2001). This is not only true of animal pathogens, but plant pathogens as well (Janjusevic, Abramovitch et al. 2006). These observations represent the elegant and close evolutionary relationships between microbial organisms and their hosts.

1.5 *Yersinia* Protein Kinase A

An important virulence factor of *Yersinia* spp. is the *Yersinia* protein kinase A, or YpkA (also called YopO in *Y. enterocolitica*), a substrate of the T3SS which was first identified through its important contribution to disease progression and a region of sequence homology to eukaryotic serine/threonine kinases (Galyov, Hakansson et al. 1993). *Y. pseudotuberculosis* mutants with disruptions in YpkA have shown severely attenuated virulence in mouse infection models (Galyov, Hakansson et al. 1993). YpkA is an 82kD multi-domain protein, which, in addition to the protein kinase, possesses a

1	**89-151**	**399-450**	**710 732**
N terminal Domain	Kinase Domain	GTPase Binding Domain	
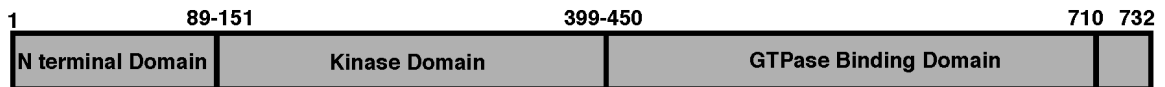
Figure 1.2: Domain Map of YpkA

The boundaries demarcating the various domains of YpkA from work prior to this thesis. The amino acid cut offs are based on the cited literature. Residues 710 to 732 represent a segment necessary for actin binding and kinase activation.

domain that binds to the small GTPases RhoA and Rac1, a domain required for activation by actin, and an N-terminal domain that targets YpkA to the cell membrane in addition to binding its chaperone, SycO (Galyov, Hakansson et al. 1993; Galyov, Hakansson et al. 1994; Roggenkamp, Schubert et al. 1995; Hakansson, Galyov et al. 1996; Barz, Abahji et al. 2000; Dukuzumuremyi, Rosqvist et al. 2000; Juris, Rudolph et al. 2000; Letzelter, Sorg et al. 2006). A domain map of YpkA is shown in Figure 1.2.

In cultured cells, YpkA appears to function primarily in disrupting the host actin cytoskeleton (Hakansson, Galyov et al. 1996; Dukuzumuremyi, Rosqvist et al. 2000; Juris, Rudolph et al. 2000).   Cells transfected with YpkA, or exposed to strains of *Yersinia* preferentially translocating YpkA over other virulence factors, lose their actin stress fibers and tend to round-up and detach upon prolonged exposure (Hakansson, Galyov et al. 1996; Juris, Rudolph et al. 2000; Nejedlik, Pierfelice et al. 2004). Interestingly, this effect appears to be only partially associated with the kinase activity, as mutations in the active site attenuate, but do not abolish, cytoskeletal alterations, and large internal deletions in the protein that are C-terminal to the kinase domain are unable to cause cytoskeletal disruptions (Dukuzumuremyi, Rosqvist et al. 2000; Juris, Rudolph et al. 2000). Additionally, experiments with yeast strains harboring inducible expression

of YpkA constructs, show that kinase-dead point mutants rapidly kill yeast just as well as wild type YpkA, whereas C-terminal deletions leave the yeast intact (Nejedlik, Pierfelice et al. 2004). These results suggesting that the C-terminal domain may play a large role in pathogenesis, which is supported by the fact that in infected Hela cells the amount of RhoA-GTP, or activated RhoA, is reduced (Dukuzumuremyi, Rosqvist et al. 2000). YpkA localizes to the plasma membrane in cultured cells (Hakansson, Galyov et al. 1996; Dukuzumuremyi, Rosqvist et al. 2000; Nejedlik, Pierfelice et al. 2004), and mutations in the protein that abolish binding to the Rho GTPases do not affect membrane localization (Dukuzumuremyi, Rosqvist et al. 2000). More specifically, the N-terminal domain region of YpkA (residues 20-90) contains all elements necessary for membrane localization (Letzelter, Sorg et al. 2006).

Biochemically, the kinase activity of YpkA is regulated in a host dependent fashion(Barz, Abahji et al. 2000; Dukuzumuremyi, Rosqvist et al. 2000; Juris, Rudolph et al. 2000; Trasak, Zenner et al. 2007). Recombinant and *Yersinia* secreted YpkA is inactive, but is activated in the presence of heat-sensitive host cell extract(Barz, Abahji et al. 2000; Dukuzumuremyi, Rosqvist et al. 2000; Juris, Rudolph et al. 2000), and in particular in the presence of actin alone (Juris, Rudolph et al. 2000). A twenty amino acid peptide at the COOH-terminus of YpkA was shown by deletion mutagenesis to be critical for YpkA binding to actin and for kinase auto-phosphorylation *in vitro*(Juris, Rudolph et al. 2000). Additionally, N-terminal elements are necessary for achieving the full activity of the YpkA kinase domain(Trasak, Zenner et al. 2007). Therefore, like many eukaryotic protein kinases, the *Yersinia* kinase appears to interact with and be dependent on several regulatory elements in order to achieve full activity (Huse and

Kuriyan 2002). Finally, although the exact role in cell signaling of the kinase domain is unknown, a potential substrate has been identified for YpkA, otubain 1 (Juris, Shah et al. 2006). Otubain 1 is phosporylated in Hela cells by an uncharacterized kinase and is shown to associated with E3-ligases that are involved in the regulation of the T-cell life cycle. As such it is speculated that the kinase domain of YpkA may regulate Otubain by phosphorylation and interfere in T-cell regulation (Juris, Shah et al. 2006).

1.6 The Rho-Family of Small GTPases

When examining the biological function of the Yops, a very apparent pattern emerges. Out of the six Yops that interfere with normal cell processes, four of the Yops directly or indirectly affect the Rho-family of small GTPases, namely YopH, YopE, YopT, and YpkA. In fact, this is a common theme seen throughout bacterial pathogenesis, especially when involving the T3SS (Aktories and Barbieri 2005). Many pathogens, including *Yersinia* spp. target this family of signaling molecules as the small GTPases are central to the regulation of the actin cytoskeleton, and thus phagocytosis and innate immunity (Gruenheid and Finlay 2003). As such, to properly understand and examine the biochemistry of the Yops, a comprehensive understanding of the biology of the Rho-family of GTPases is necessary.

The Rho-family of small GTPases consist of a group of well-characterized signaling molecules that regulate a variety of cellular functions. This family of genes is a subset of the ubiquitously expressed Ras-related small GTPases, and act as the signal regulators of actin cytoskeletal rearrangement, microtubule dynamics, vesicle trafficking,
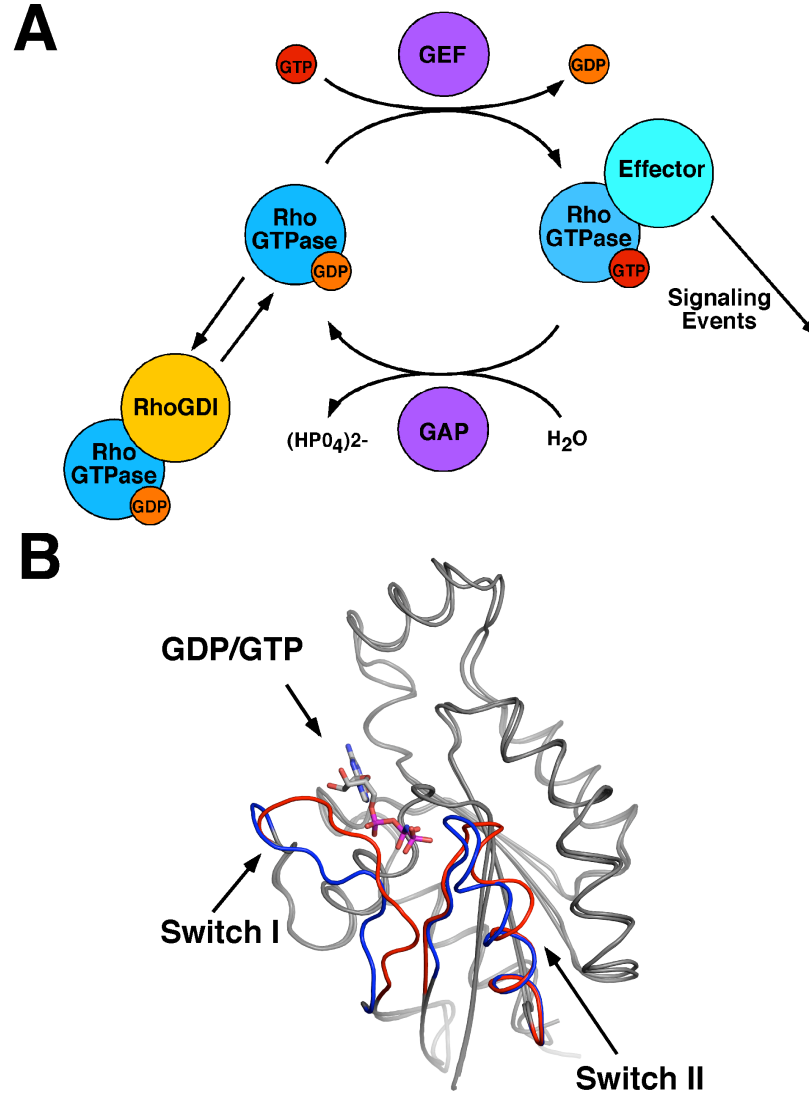
Figure 1.3: Biochemistry of the Rho-Family of small GTPases

(A) A schematic of the small GTPase signalling cycle. The small GTPases cycle between "off" and "on" signaling states based on interactions with GEFs (which turn the system on), GAPs (which turn the system off) and RhoGDIs which prevent the small GTPase from associating with the other regulators.

(B) The conformational changes induced by nucleotided binding are shown. Rac1-GMPPNP (1MH1 pdb identifier) and Rho-GDP (1FTN pdb identifier) are shown. The switch regions are colored in red for the GTP bound conformation (Rac1-GMPPNP) and in blue for the GDP conformation (RhoA-GDP).

cell cycle control, and cell motility, among other functions (Symons and Settleman 2000; Jaffe and Hall 2005). These membrane localized molecules regulate the aforementioned cellular processes by acting as molecular switches, cycling between an "off" and "on" state in response to extracelluar stimuli such as cytokines, cell adhesion molecules, and bacterial invasion (Symons and Settleman 2000; Aktories and Barbieri 2005).

Rho-GTPases localize to the cell membrane due to an isoprenyl modification at a C-terminal cysteine residue and achieve an "off" and "on" biological state by cycling between GDP (guanosine-5'-diphosphate) and GTP (guanosine-5'-triphsophate) bound forms, respectively (Figure 1.3A). The Rho-GTPase GDP/GTP cycle is controlled in the cell by three known families of regulatory molecules: GEFs, GAPs, and GDIs. GEFs (guanine nucleotide exchange factors) promote the exchange of GTP for GDP thus catalyzing activation and turning the system "on," whereas GAPs (GTPase activating proteins), stimulate the intrinsic rate of GTP hydrolysis in the Rho-GTPases, turning the system "off." Finally RhoGDIs act as a master-control switch for the system, removing the small GTPase from the membrane and inhibiting nucleotide dissociation (Van Aelst and D'Souza-Schorey 1997; Jaffe and Hall 2005).

At the molecular level, these signaling events are achieved through conformational changes in the Rho-GTPases at the so-called "switch" regions. Switch I (residues 29-40 Rac1 numbering) and Switch II (residues 55-74 Rac1 numbering) adopt discrete canonical conformations based on the presence of GDP or GTP as shown in Figure 1.3B (Hirshberg, Stockley et al. 1997; Wei, Zhang et al. 1997). In the GTP bound form, the Rho-GTPases are recognized by downstream effectors, many of which are kinases, that are activated due to a conformational change upon Rho-GTPase binding,

21

releasing a functional domain from an inactivation domain (Maesaki, Ihara et al. 1999; Bishop and Hall 2000). The cellular regulators of the small GTPase signaling cycle (GEFs, GAPs, and GDIs), modulate the Rho-GTPases by also interacting with the switch regions, adjusting the conformations of the switch regions to promote the relevant biochemistry of nucleotide exchange, GTP hydrolysis, and inhibition of nucleotide dissociation, respectively (Vetter and Wittinghofer 2001; Hakoshima, Shimizu et al. 2003; Dvorsky and Ahmadian 2004).

As the Rho-GTPases are the control switches for many important cellular processes (Jaffe and Hall 2005) and are the targets of several bacterial pathogens (Aktories and Barbieri 2005), these signaling molecules have been extensively studied. In fact, much work has been devoted to revealing their molecular structures alone and in complex with their regulators and effectors to gain a better understanding of their normal cellular function and how their function is subverted by bacterial virulence factors. To date, the crystal structures of the most extensively characterized Rho-GTPases (RhoA, Rac1, and Cdc42), have been solved in both the GDP and GTP bound state, and in complex with a variety of eukaryotic effectors and bacterial virulence factors, save for a Rac1-GDP complex without a bound effector (Hakoshima, Shimizu et al. 2003; Dvorsky and Ahmadian 2004).

MATERIALS AND METHODS

2.1 Cloning and Mutagenesis

The *Y. pseudotuburculosis* YpkA gene was cloned from the *Yersinia* virulence plasmid (generously provided by J. Bliska, State University of New York at Stony Brook) as an N-terminal fusion with glutathione-S-transferase (GST). This or the original virulence plasmid was then used as the template for the creation of additional YpkA clones. All N-terminal GST fusion constructs were created using the restriction sites SalI/NotI and placed into the in-lab modified pGEX4T3 vector (GE Healthcare). This laboratory stock was modified to have an additional thrombin site and the cleavage site for rhinovirus 3C protease. Clones (115-465) and (115-532) were also cloned using the restriction sites NdeI/NotI into a lab-modified version of pET28b (Novagen) that contained a 3C protease cleavage site. Clone (115-428) was cloned using the restriction sites SalI/NotI into a lab-modified version of the DUET (Novagen) vector, which has a 3C protease cleavage site. GST-Rac1 (1-184) was obtained in the standard commercial vector pGEX4T3 from C. Erec Stebbins. The Cdc42 (1-184) was also obtained from C. Erec Stebbins in pET3d but was subcloned into the in lab modified pGEX4T3 vector using the restriction sties, Sa1I/NotI. Both the genes for RhoA and SopE were obtained from RZPD (German Resource Center for Genomic Research) and then cloned into the in lab modified pGEX4T3 using restriction sites Sal1/Not1. RhoA was cloned as 1-181 (F25N) and SopE as (78-240). The mutations and constructs created were used to optimize solubility and expression. To create the YpkA and RhoA mutants, the amino

acid substitutions were introduced by PCR using primers containing the appropriate base changes and subsequent removal of the template plasmid by digestion with Dpn I prior to transformation, and verified by DNA sequencing. YpkA clones were transformed into DH5α cells or BL21(DE3) cells the latter for expression. Primers were ordered from IDT DNA technologies, molecular biology enzymes and reagents were obtained from New England Biolabs, and additional materials used for DNA purification were purchased from Quiagen.

2.2 Protein Expression and Solubility

Optimal expression of protein constructs was found to be at 20° C with induction using 1mM IPTG in an overnight culture. The solubility of each YpkA construct was determined by growing a 5mL prep and taking 1mL samples for time points before induction and after induction. Each sample was resuspended in lysis buffer (20mM Tris pH7.5 200mM NaCl) and lyzed via sonication. After centrifugation for 10 minutes the supernanent was removed and the pellet resuspended in denaturing buffer (50mM Tris pH7.5 5.7M urea). Each sample was normalized for protein concentration by Bradford assay and run on a 12% or 15% SDS-page gel to visualize. To further delineate smaller constructs, limited proteolysis was performed on soluble YpkA constructs after protein purification (see Protein Purification). This was achieved with subtilisin on ice for 20 minutes. Subtilisin was added at concentrations under 1% (wt:wt) of YpkA and the reaction initiated with the addition of 5mM $CaCl_2$. The digestions were visualized by

SDS-PAGE, transferred to a nitrocellulose membrane and sequenced by Edman degradation at the Rockefeller University Proteomics Resource Center.

2.3 Protein Purification

Proteins were expressed in BL21(DE3) cells by overnight induction with 1mM IPTG at 20°C using LB media (EMD Chemicals Inc). Cells were resuspended in lysis buffer (20mM Tris pH7.5, 200mM NaCl, and 1mM DTT) plus 1mM PMSF at a volume of 100mL per 9L of media and frozen at -80°C before use. Pellets were thawed then the cells were then lysed with an Emusiflux-C5 (Avestin) and the debris removed by centrifugation. DNA was digested by the addition of 1mM $MgCl_2$ and 1mg/mL DNase I. The lysate was cleared by centrifugation at 16000rpm for 30minutes before subsequent purification steps. After final purification all samples were aliquoted and flash frozen in liquid nitrogen.

2.3a Purification of YpkA (1-732) and (55-732)

Each of these N-terminal GST fusion constructs was purified "in batch." A 1.5L culture was grown and lysed as described previously. The resulting supernanent was mixed with 2 to 4mL of pre-prepared glutathione sepharose beads (GE Healthcare) pre-washed in lysis buffer plus 1mM DTT. YpkA was then bound to the beads by rotation at 4°C for 2hrs. The supernanent was decanted from the beads after centrifugation at 3400rpm for 5 minutes and the beads washed 3 times in wash buffer (20mM Tris pH7.5

500mM NaCl 1mM DTT) by rotation for 30minutes at 4°C. The beads were re-washed in lysis buffer plus 1mM DTT before digestion. To remove YpkA from GST and the affinity resin, YpkA was cleaved using an N-terminally tagged GST-3C site specific protease construct, liberating YpkA into the supernanent buffer (20mM Tris pH7.5 200mM NaCl 1mM DTT 1mM EDTA). The resulting material was decanted from the affinity resin, concentrated, and flash frozen for storage.

2.3b Purification of YpkA (434-732)

YpkA (434-732) was expressed as an N-terminal GST fusion construct in 6 1.5L cultures and was grown and lysed as described previously. The resulting lysate was then passed once over a 30mL Q-sepharose gravity column preincubated in lysis buffer (20mM Tris pH7.5 200mM NaCl) plus 1mM DTT. The resulting material was then passed over a 50mL gluthione-sepharose column to bind YpkA and then washed in 5 column volumes of wash buffer (20mM Tris pH7.5 500mM NaCl 1mM DTT). GST-YpkA was eluted from the resin by passing three column volumes of lysis buffer plus 10mM glutathione and 1mM DTT. This was then repeated with the initial flow through and the two elutions pooled together. YpkA was liberated from GST by an overnight digestion with GST-3C in a 1 to 50 ratio. During digestion the material as dialyzed overnight into lysis buffer plus 1mM DTT and 1mM EDTA to remove glutathione. After digestion, GST was removed by two passes over a 50mL glutathione-sepharose column. This was followed by dilution of the material in 20mM Tris pH7.5 to lower NaCl concentration to 50mM or less for ion exchange chromatography. YpkA was further

purified by anion exchange chromatography with a 25mL Q15 Source column (GE Healthcare) by FPLC (Åkta) removing GST as YpkA did not bind appreciably to the column. The flow throw was then reconcentrated to 40mg/mL and passed over an SD200 (120mL volumes) gel filtration column for final purification and buffer exchange into crystallization buffer (20mM Tris pH7.5 50mM NaCl 1mM DTT).

2.3c Purification of YpkA (115-465) and (115-732)

YpkA was expressed as an N-terminal GST fusion construct in six 1.5L cultures and was grown and lysed as described previously. The resulting lysate was then based once over a 30mL Q-sepharose gravity column preincubated in lysis buffer (20mM Tris pH7.5 200mM NaCl) plus 1mM DTT. The resulting material was then passed over a 50mL gluthione-sepharose column to bind YpkA and then washed in 5 column volumes of wash buffer (20mM Tris pH7.5 500mM NaCl 1mM DTT). GST-YpkA was eluted from the resin by passing three column volumes of lysis buffer plus 10mM glutathione and 1mM DTT. This was then repeated with the initial flow through and the two elutions pooled together. YpkA was liberated from GST by an overnight digestion with GST-3C in a 1 to 50 ratio. During digestion the material as dialyzed overnight into lysis buffer plus 1mM DTT and 1mM EDTA to remove glutathione. After digestion, GST was removed by two passes over a 50mL glutathione sepharose column. This was followed by dilution of the material in 20mM Tris pH7.5 to lower NaCl concentration to 50mM or less for ion exchange chromatography. YpkA was then further purified by anion exchange chromatography by a 25mL Q15 Source column (GE Healthcare) by FPLC

(Åkta). YpkA was eluted from the column by washing in a salt gradient of 20 column volumes to final concentration of 400mM NaCl. The eluted material was separated from the residual GST peak and concentrated to 25mg/mL and passed over an SD200 (120mL volume) gel filtration column for additional purification and buffer exchange into crystallization buffer (20mM Tris pH7.5 50mM NaCl 1mM DTT). The final material was reconcentrated and passed over an SD200 gel filtration column of 25mL to removed any aggregation.

2.3d Purification of YpkA (115-532)

YpkA was expressed as an N-terminal his-tag fusion construct in six 1.5L cultures and was grown and lysed as described previously, except that DTT was omitted from the buffer in all steps before and during the use of a nickel sepharose column. The lysate was first passed over a 30mL nickel-sepharose column by FPLC (Åkta). The column was washed with five column volumes of lysis buffer plus 30mM Imidazole then eluted by a 20 column volume imidazole gradient for a final concentration of 500mM. After elution the material was digested overnight with GST-3C while being dialyzed into 20mM Tris pH7.5 100mM NaCl. The digested material was repassed over a nickel-sepharose column to removed non-digested material. The flow through was then supplemented with 1mM DTT and 1mM EDTA, diluted 5x and passed over a Q15 source column for further purification by anion exchange chromatography. YpkA was eluted from the column by a 20 column volume salt gradient, reconcentrated to 23mg/mL and purified by

gel filtration using a 25mL SD200 column. The final buffer conditions were 20mM Tris pH7.5 100mM NaCl 1mM DTT.

2.3e Purification of YpkA (115-428) and (115-465) Refolding

YpkA was expressed as an N-terminal his-tag fusion construct in six 1.5L cultures and was grown as described previously. YpkA (115-428) was refolded for optimal purification. The initial cell pellet was resuspended in denaturation buffer (10mM Tris pH8.0 100mM NaPhosphate 8M Urea). The chemically lysed material was then sonicated three times for 10minutes to shear DNA. The lysate was cleared by centrifugation as with non-denatured preps as described previously, and passed over 15mL nickel-sepharose columns pre-equilibrated in denaturation buffer plus 30mM imidizole. The columns were washed in five column volumes of denaturation buffer plus 30mM imidizole followed by elution from the resin with denaturation buffer plus 500mM imidazole. Purified YpkA was first diluted to 0.5mg/mL and then was refolded by dialysis by four buffer exchange steps. Each step was done at 4°C and lasted at least 4 hours. The first three buffer exchange steps were into lysis buffer (20mM Tris pH7.5 200mM NaCl) plus 5% glycerol 2mM EDTA and 2mM DTT. The fourth step omitted glycerol and DTT. After refolding, YpkA was digested overnight with GST-3C during another dialysis step repeating the fourth buffer exchange. The digested material was passed over a 15mL passed over a glutathione-sepharose column to removed GST-3C then over a 15mL nickel sepharose column to remove undigested material. YpkA was then further purified by anion exchange chromatography by a 25mL Q15 Source column

(GE Healthcare) by FPLC (Åkta). YpkA was eluted from the column by washing in a salt gradient of 20 column volumes to final concentration of 400mM NaCl. The eluted material concentrated to 30mg/mL and passed over an SD200 (120mL volume) gel filtration column for additional purification and buffer exchange into crystallization buffer (20mM HEPES pH7.5 50mM NaCl 1mM DTT). The final material was re-concentrated and passed over an SD200 gel filtration column of 25mL to remove any aggregation.

2.3f Purification of Rac1 (1-184) GDP

Rac1 (1-184) (F78S) was cloned as an N-terminal GST fusion in the vector pGEXT4T3 (Stebbins and Galan 2000). The GST-Rac1 fusion protein was expressed and lysed as described previously using a total of six 1.5L cultures. The supernatant was passed over a Q-sepharose gravity column followed by a glutathione-sepharose gravity column. The glutathione-sepharose column was washed with 5x volume wash buffer (20mM Tris pH7.5, 500mM NaCl, and 2mM DTT) followed by elution of GST-Rac1 in lysis buffer plus 10mM glutathione. GST was removed by overnight cleavage with thrombin during dialysis into lysis buffer plus 2.5mM $CaCl_2$. The digest was then passed by gravity flow over a heparin sepharose column to remove thrombin followed by dilution in 20mM Tris pH7.5 1mM DTT to lower salt concentration for ion exchange chromatography. GST was separated from Rac1 by passing the digest over a Q sepharose column followed by an SP FF Sepharose column by FPLC for crystallization (GE Healthcare). Purified Rac1 not used in crystallization was purified instead with a

Q15 source column (GE Healthcare). Finally Rac1 (1-184) GDP was concentrated to 60mg/mL and run over an SD200 gel filtration column (GE Healthcare) to isolate non-aggregated material and to exchange the buffer into 20mM Tris pH7.5, 50mM NaCl, and 1mM DTT. All purification steps were performed at 4°C.

2.3g Purification of RhoA GDP, Cdc42 GDP, and SopE

RhoA (1-181 F25N) GDP, Cdc42 (1-184) GDP, and SopE (78-240) were expressed as N-terminal GST fusion construcst in 6 1.5L cultures and was grown and lysed as described previously. The resulting lysate was then based once over a 30mL Q-sepharose gravity column preincubated in lysis buffer (20mM Tris pH7.5 200mM NaCl) plus 1mM DTT. The resulting material was then passed over a 50mL gluthione-sepharose column to bind the protein and then washed in 5 column volumes of wash buffer (20mM Tris pH7.5 500mM NaCl 1mM DTT). The construct was eluted from the resin by passing three column volumes of lysis buffer plus 10mM glutathione and 1mM DTT. Each protein was liberated from GST by an overnight digestion with GST-3C in a 1 to 50 ratio. During digestion the material as dialyzed overnight into lysis buffer plus 1mM DTT to remove glutathione. The dialysis buffer for both RhoA and Cdc42 was supplemented with 1mM $MgCl_2$ and 0.2mM EDTA, whereas the buffer for SopE was supplemented with 1mM EDTA. After digestion, GST was removed by two passes over a 15mL glutathione-sepharose column. The flow through was concentrated to 40 to 60mg/mL then further purified by gel filtration using an SD200 (120mL volume) column. The final buffer conditions were 20mM Tris pH7.5 50mM NaCl 1mM DTT.

2.3h Expression and Purification of Selenomethionine YpkA (434-732)

To incorporate selenomethionine into YpkA for single anomalous dispersion experiments (see below), YpkA GST-(434-732) was first transformed into B834 (DE3) which is a methionine auxotroph. Minimal media was made containing 32mg/L thymine, 2mM MgSO4, 1% glucose, 20mg/L thiamine, 20mg/L pyridoxine, and 20mg/L biotin. This was supplemented with all amino acids, except methionine, for a total amount of 50μg each. Selenomethionine was added at 10mg/L. A 200mL YpkA culture in LB was first grown overnight at 37°C, the cells harvested by centrifugation, then resuspended in added to 1.5L of the selenomethionine media. This was and grown to OD 0.6 at 37°C and 100mL of each was used to inoculate one of six 1.5L cultures. The new culture was regrown to OD0.6 and then induced as described for GST-YpkA (434-732). The same purification procedure for native GST-YpkA (434-732) was followed to purify the selenomethionine substituted material.

2.4. Protein Crystallization, Data Collection, and Structure Refinement

2.4a YpkA (434-732)

YpkA (434-732) was reductively methylated following published protocols (Rypniewski, Holden et al. 1993) Diffracting crystals were grown at 4°C by hanging drop vapor diffusion using a 1:1 mixture of protein and well solution, which consisted of 100mM CAPs pH10.5, 140mM to 180mM NaCl, and 16% to 18% PEG1500. Crystals

formed in the hexagonal space group P6$_5$22 with unit cell dimensions of a=b=60.0Å, and

c=402.0Å. For phasing, selenomethionine substituted YpkA (434-732) was reductively

methylated, purified, and crystallized in the same manner as native YpkA (434-732).

Data were collected at beamline X29 at the National Synchrotron Light Source at

Brookhaven National Laboratories. For data collection, YpkA (434-732) crystals were

frozen by a step-wise exposure to cryobuffer (crystallization mother liquor plus 25%,

30%, or 35% PEG1500) and then flash frozen in liquid nitrogen. The frozen crystals

were then transferred directly into the cold stream at -180°C. Data for a single native

crystal were collected to 2.0Å. Data for the selenomethionine substituted protein were

collected at the selenium absorption edge, at a wavelength maximizing the anomalous

signal. The derivative data was collected to 2.9Å. Data were processed using the HKL

software package (Otwinowski and Minor 1997). Phases for the monomer structure were

determined using the anomalous signal using the SOLVE/RESOLVE(Terwilliger 2004)

package. The partial model built by RESOLVE was then rebuilt and initially refined

using ARP/wARP(Perrakis, Morris et al. 1999) with the native data set. The model

generated by ARP/wARP was then refined using REFMAC5 (Murshudov, Vagin et al.

1997) from the CCP4 suite of programs (1994; Potterton, Briggs et al. 2003). The final

model has an R/Rfree of 20.9 / 23.8, with 96.4% of the residues in the most favored

regions of the Ramchandran plot with no outliers. The solution was then refined using

REFMAC5 (Murshudov, Vagin et al. 1997) from the CCP4 suite of programs (1994;

Potterton, Briggs et al. 2003).

2.4b YpkA (434-732) Rac1 (1-184) GDP Complex

The complex between methylated YpkA (434-732) and Rac1 (1-184) bound to GDP was formed by first combining purified YpkA with an excess of Rac1 and then by isolation by gel filtration chromatography on a 120mL SD200 column. The two proteins were incubated at 4°C for 30minutes with the addition of 4mM $MgCl_2$. The final buffer conditions were 20mM Tris pH7.5 50mM NaCl 1mM DTT 0.2mM $MgCl_2$. Diffracting crystals were grown at 22°C by hanging drop vapor diffusion using a 1:1 or 3:2 mixture of complex and well solution, which consisted of 100mM HEPES pH6.5-7.5, 5% to 8% PEGMME2000, and 0.2μL of a 1/10 dilution of seed stock created using Seed Bead (Hampton Research). Crystals formed in the space group P1 with unit cell dimensions of a=66.4Å b=75.5Å, c=99.7Å, α=92.1°, β=103.4°, and γ=115.8°. Data were collected at beamline X29 at the National Synchrotron Light Source at Brookhaven National Laboratories. For data collection, complex crystals were frozen by a step-wise exposure to cryobuffer (crystallization mother liquor plus 25%, 30%, or 35% PEGMME2000) and then flash frozen in liquid nitrogen. The frozen crystals were then transferred directly into the cold stream at -180°C. Data for a two native crystals were collected to 2.6Å. The data for two crystals were merged and processed using the HKL software package (Otwinowski and Minor 1997). The YpkA (434-732) and Rac1 (1-184) GDP complex phases were determined by molecular replacement (search model 1MH1). The solution was then refined using REFMAC5 (Murshudov, Vagin et al. 1997) from the CCP4 suite of programs (1994; Potterton, Briggs et al. 2003). The final model has an R/Rfree of 22.4 / 25.9.

2.4c Rac1 (1-184) GDP Zinc

Rac1 was combined with reductively methylated YpkA (434-732) as previously reported and screened for crystallization (Prehna, Ivanov et al. 2006). The complex was concentrated to 39mg/mL and Rac1 crystals were found to grow after initial YpkA precipitation at 4°C by hanging drop vapor diffusion using a 1:1 mixture of protein to well solution which consisted of 100mM MES pH6.0, 15mM $ZnSO_4$, and 10% PEGMME 550. Rac1 crystals were further optimized by micro-seeding using the "seed-bead" from Hampton Research. Diffraction quality crystals could be grown in pH conditions ranging from 5.5 to 7.5, 5% to 20% PEGMME550, and 1mM to 40mM $ZnSO_4$. $ZnSO_4$ was necessary for crystallization and could not be replaced with $MgSO_4$. Crystals formed in the trigonal space group $P3_221$ with unit cell dimensions a=b=89.7Å c=191.6Å. Data were collected on a Marccd detector at beamline X3A at Brookhaven national laboratories. For data collection Rac1 crystals were frozen by a step-wise exposure to cryo-buffer (15%, 25%, 28% PEGMME550 plus mother liquor) and then flash frozen in liquid nitrogen. The frozen crystals were transferred directly to the cold stream at -180°C. The data on one crystal was obtained to 1.9Å resolution at a wavelength of 1.284Å. The data was processed using the HKL software package (Otwinowski 1997). Phases for the Rac1 GDP structure were determined by molecular replacement using the program Phaser as part of the CCP4 suite of programs (Potterton, Briggs et al. 2003) using the search model Rac1 search model (1MH1 pdb accession code) and YpkA (1H7O pdb accession code). The molecular replacement solution revealed that only Rac1 was present. The model was initial refined with REFMAC5

(Murshudov, Vagin et al. 1997) from the CCP4 suite of programs (Winn, Isupov et al. 2001; Potterton, Briggs et al. 2003; Winn 2003; Winn, Murshudov et al. 2003) then rebuilt using ARP/wARP (Perrakis, Morris et al. 1999). The rebuilt model was further refined with REFMAC5 for a final Rwork/Rfree of 17.4/20.8% with 92.0% of the residues in the most favored regions of the Ramachandran plot with no outliers. Zinc sites were found using the program SOLVE(Terwilliger 2004) then placed into the molecular replacement solution.

2.5 Structure Analysis and Molecular Graphics

To analyze the molecular contacts of the crystal structures, the programs Areamol and Contact were employed from the CCP4 suite of programs (Winn, Isupov et al. 2001; Potterton, Briggs et al. 2003; Winn 2003; Winn, Murshudov et al. 2003) in addition calculation tools within CNS (Brunger, Adams et al. 1998). The computer identified contacts were then verified by visual inspection. All molecular graphics were made using the program Pymol (DeLano 2002) and Figures constructed with Adobe Photoshop 7.0 and Canvas 8.

2.6 YpkA Binding and Nucleotide Exchange Assays

The complex between YpkA (434-732), and Rac1((1-184)) was formed by first combining purified YpkA with an excess of Rac1, examined by gel filtration chromatography, and finally visualized by SDS-PAGE as described previously. YpkA

Contact A mutant (N595A, Y591A, E599A) and YpkA Contact B mutant (N627A, R628A, S631A) were purified as described for YpkA (434-732) and the complex with Rac1 formed and analyzed as described for YpkA (434-732). Binding experiments with RhoA(F25N 1-181) and Cdc42 ((1-184)) were carried out in the same manner as Rac1.

For nucleotide exchange assays, purified Rac1 or RhoA was incubated alone with buffer (20mM Tris pH7.5 50mM NaCl 5mM MgCl$_2$) or with purified YpkA (434-732), YpkA (434-732) Contact A mutant, for 10 minutes on ice. Both YpkA (434-732) and YpkA (434-732) contact A mutant were added in either a 1.5 fold or in a 3 fold molar ratio access of Rac1. Purified Cdc42 was incubated alone with buffer or with purified YpkA (434-732) in a 3 fold molar ratio access. To initiate the intrinsic exchange reaction the stock incubations were then added to reaction buffer (20mM Tris pH7.5 50mM NaCl 5mM MgCl$_2$ and mant-GTP obtained from Invitrogen) resulting in a final concentration of Rac1 at 40μM and mant-GTP at 100μM. Reactions were then transferred to a 96 well plate and left for 5 minutes to bring to room temperature. The fluorescence of each reaction was measured every 5 minutes for a total of one hour using a SpectraMax GeminiXS fluorimeter from Molecular devices. An excitation wavelength of 355nm and an emission wavelength of 448nm with a cutoff of 435nm was used to detect the presence of bound mant-GTP. Competition assays with SopE were performed as described for the intrinsic nucleotide exchange. SopE was added to each reaction after the five minute incubation to room temperature at a final concentration of 0.22μM for Rac1 or RhoA, and 0.04μM for Cdc42.

2.7 YpkA Transfection Assays

2.7a Henle Cell Transfection and Visualization

YpkA constructs 1-732, (434-732), and (434-732) Contact A mutant (Tyr 591A, Asn 595A, Glu 599A), were cloned into the pFLAG-CMV-4 mammalian expression vector using the restriction sites NotI and XbaI resulting in N-terminally tagged constructs. Henle407 cells were grown to confluence in DMEM media containing 10% Fetal Bovine Serum (FBS) and penicillin/streptomycin. The cells were then transferred to six-well tissue culture dishes containing a glass coverslip. Approximately 200,000 cells were added to each well with a total volume of 2ml and grown overnight. The transfections were carried out using the Geneporter2 reagent and protocol (Genlantis). A total of 6µg of DNA was added to each reaction and reactions were performed in triplicate. Reactions were allowed to proceed for 24hrs before being fixed and stained. Cells were fixed with 3% Formaldyde, permeablized with 0.5% Triton, and blocked with 3% BSA PBS before exposure to antibodies. The YpkA constructs were visualized by primary antibody staining with mouse α-FLAG antibody (Sigma) and then by the secondary Alexa-Fluor goat α-mouse antibody (Molecular Probes).   The actin cytoskeleton was visualized by staining with Rhodamine Phalloidin (Molecular Probes). To quantify the effect of each construct on the host cytoskelon each reaction was counted blind and in triplicate. For each reaction, between 75 and 200 transfected cells were counted for each of the triplicate reactions (a total of 225 to 450 cells for each construct) and the state of their cytoskeleton scored as "wild type" or "intermediate" effect. The

cells were visualized by the use of an Axioplan2 upright microscope with Attoattic fluorescent filters and a Hammatsu Digital Camera or by the use of a Zeiss LSM510 confocal microscope.

2.7b Relative Stability of YpkA Transfection Constructs

Transfected and untransfected Henle407 cells expressing N-terminal FLAG-tagged YpkA full length constructs were lysed in buffer (50mM Tris pH 7.4 150mM NaCl 1% SDS 1% Sodium Deoxycholate) and run on a 7.5% Poly-acrylamide gel (total protein normalized using by Bradford assay). The gel material was transferred to a nitrocellulose membrane and blotted with α-FLAG M2 peroxidase conjugated antibody (Sigma). The blot was visualized using Western Detection Agents 1 and 2 (Amersham) followed by a 15 minute exposure to Hyper film ECL (Amersham). Transfected and untransfected Henle407 cells expressing either N-terminally FLAG-tagged YpkA (434-732) or YpkA (434-732) Contact A mutant were lysed as above and run on a 15% Poly-acrylamide gel. The gel material was transferred to a nitrocellulose membrane and blotted with α-FLAG M2 peroxidase conjugated antibody (Sigma). The blot was visualized usingWestern Detection Agents 1 and 2 (Amersham) followed by a 2 minute exposure to Hyper film ECL (Amersham).

2.8 *Yersinia* Strains and Mouse Infection Assays

This section was performed in collaboration by Maya I Ivanov and James B. Bliska at SUNY Stonybrook. IP2777 is a virulent serogroup O1 strain of *Y. pseudotuberculosis* obtained from Michel Simonet (Simonet et al., 1992). The LD50 of IP2777 in C57BL/6 mice challenged intragastrically is 5×10^8 CFU. YpkA null and contact A mutants were constructed as described in Supplementary Experimental Procedures. Two independent mouse infection experiments were carried out and the results presented are compiled from the two experiments. Eight week-old female C57BL/6 mice (Taconic) were challenged by the intragastric route using a 20 gauge feeding needle. Bacterial inocula were prepared from cultures grown in Luria Broth (LB) with shaking at 26°C. Bacteria were inoculated into LB, grown overnight, and subcultured in LB to an OD600 of 0.1. After a second overnight growth, the cultures were centrifuged, the bacterial pellets were washed once in Hank's balanced salt solution (HBSS), and suspended HBSS to yield 5×10^9 CFU per 0.2 ml. C57BL/6 mice were fasted for 18 h prior to infection. Groups of 4 animals were infected with 0.2 ml of suspended bacteria (a dose equivalent to 10 LD50s). Infected mice were provided food and water and carefully observed three times a day over a 14 day period. Mice exhibiting severe signs of disease (ruffled fur, hunched posture, and immobility) were humanely euthanized by CO_2 inhalation. Survival curves were plotted using Prism (GraphPad) and analyzed for significant differences using the Mantel-Haenszel logrank test. These experiments were carried out in compliance with protocols approved by the IACUC at Stony Brook University.

2.9 *Yersinia* Strain Constructions and HeLa Cell Infection Assays

This section was performed in collaboration with Maya I. Ivanov and James B. Bliska at SUNY Stonybrook. A *ypkA* null mutant of IP2777 in which nucleotides 1 to 2199 of the *ypkA* open reading frame were deleted was constructed by allelic replacement using the plasmid pYopOΔ1-2199 as described (Viboud et al., 2003). A *ypkA* mutant of IP2777 deficient in GDI activity was constructed using allelic replacement to introduce three codon changes (Tyr 591Ala, Asn595Ala, and Glu599Ala) into the *ypkA* open reading frame on the virulence plasmid. For this purpose, a fragment of *ypkA* containing the three codon substitutions was removed from pFLAG-CMV-4 (434-732) using *Xma*I and *Not*I restriction sites, and inserted into the corresponding sites in pSB890 (Viboud et al., 2003). The resulting plasmid was conjugated into IP2777, and following selection steps for allelic replacement carried out as described (Viboud et al., 2003), colonies of recombinants were screened for the presence of the codon changes. Following PCR amplification of *ypkA* sequences corresponding to codons 434 to 732, PCR products were subjected to digestion with *Nae*I, which only cuts within PCR products containing the Tyr 591Ala codon change. A recombinant with the Tyr 591Ala codon change based on *Nae*I digestion was isolated, confirmed by sequencing to contain all three codon changes, and designated the *ypkA* contact A mutant. A Yop secretion assay (Galyov et al. 1993) was used to confirm that the null mutant did not express YpkA protein, while the contact A mutant expressed full length YpkA protein. Immunoblotting of secreted Yops with antibodies specific for YopE and YopT was carried out as described [Viboud et al. 2006]. A HeLa cell rounding assay was performed as described [Hakansson et al. 1996] except

that the bacteria were grown in LB for 2 h at 37°C prior to infection, and the multiplicity of infection was 50. A *Y. pseudotuberuclosis* IP2777 *yopEHTJ* mutant that expresses catalytically inactive forms of YopE, YopH, YopT and YopJ was provided by Yue Zhang at Stony Brook University.

2.10 Radiological Assays

2.10a Rac1 GAP Activity

Purified Rac1 (1-184) was preloaded with γ-^{32}P GTP by incubation in preload buffer (20mMTris pH7.5 25mM NaCl 4mM EDTA 0.1mM DTT) at a ratio of 40μg Rac1 to 2μCi γ-^{32}P GTP at room temperature. The exchange was stopped after the addition of 20mM MgCl$_2$. Rac1 was either incubated alone or with YpkA (434-732) on ice for 20 minutes in stock buffer (20mM Tris pH7.5 0.1mM DTT 1mM GTP 5mM MgCl$_2$). YopE obtained purified from postdoctoral fellow Milos Vujanac was added to a separate Rac1 reaction as a control. The final reaction conditions were 3.6μM Rac1, 48.5μM YpkA, and 43.5μM YopE. The reaction was initiated by incubation at room temperature and 5μL aliquots were taken every 5minutes for a total of 20 minutes and placed on ice in 500μL assay buffer (20mM Tris pH7.5 50mM NaCl 5mM MgCl$_2$). The entire reaction was then visualized using the Bio-blot apparatus (Biorad) and exposure on a Storage Phosphor Screen / Typhoon Scanner (Molecular Dynamics).

2.10b YpkA Kinase Activity Assay

Purified YpkA constructs were tested for autophosphorylation and phosphorylation of MBP (Mylein basic protein) in the presence of G or F rabbit muscle actin (Cytoskeleton, Inc.). The differing actin forms were prepared as per the company protocol. Reactions were prepared with an excess of MBP relative to YpkA and preincubated in either G reaction buffer (20mM HEPES pH7.5 10mM $MgCl_2$ 1mM DTT and 0.1mM ATP 1μCi γ-^{32}P ATP) or F reaction buffer (20mM HEPES pH7.5 50mM KCl 10mM $MgCl_2$ 1mM DTT and 0.1mM ATP 1μCi γ-^{32}P ATP). Each assay was initiated with an access of either G or F actin relative to YpkA and then incubated at room temperature for 30 minutes. The reactions were then visualized by SDS-PAGE and exposure on a Storage Phosphor Screen / Typhoon Scanner (Molecular Dynamics, Inc.).

2.10c YpkA Inhibitor Screening and IC_{50} Determination

The 45 top hits from virtual screening were tested for inhibitory activity against full length YpkA initially at 450 μM in the presence of G actin (2.10b). The results of each trial were visualized using the Bio-blot apparatus (Biorad) and exposure on a Storage Phosphor Screen / Typhoon Scanner (Molecular Dynamics). The IC_{50} of the best inhibitors was determined by creating serial dilutions of inhibitor from 0 μM to 181 μM (final reaction concentration) spread over 5 to 6 independent points measured in triplicate. A total of 200ng of YpkA 1-732 was used per 16 μL reaction volume. After visualization (2.10b) the intensity of each band was intergraded by the program

ImageQuant (Molecular Dynamics). This data with the program, LSW Data Analysis Tool Box version 1.1.1 in MicroSoft Excel, was used to generate curves for IC_{50} calculation. MAPK kinase and Protein kinase C (PKC) were obtained from Calbiochem and the same protocol followed as YpkA to determine inhibitor IC_{50} values. Each was added in the same molar ratio as YpkA in the final reaction conditions.

2.11 Sequence Alignment and Homology Modeling of the YpkA Kinase Domain

The following section was performed in collaboration with postdoctoral fellow Xin Hu. The minimum kinase domain of YpkA (sequence (115-428)) was used. The sequence alignment was performed using FUGUE, 3D-PSSM, and SUPERFAMILY, with alignment adjustments during optimization. Five homology models of YpkA were constructed using MODELLER version 7.0, the representative model selected had lowest value of MODELLER target function. All the structural models were validated by checking their quality with the programs PROCHECK and PROSA II. The 3D structures of YpkA built by MODELLER were further modeled in the presence of ATP. The ATP molecule was extracted from the CDK2 crystal structure (1HCK) and superimposed into the putative ATP-binding pocket of YpkA structures. The protein/ATP complexes were subjected to a full energy minimization through a series of independent steps. Energy minimizations were carried out employing Sybyl 7.0. In each step, MMFF94s force field was applied with 0.05 kcal/Å convergence and 5000 steps using Powell method.

2.12 Database and Virtual Screening of YpkA Inhibitors

The following section was performed in collaboration by postdoctoral fellow Xin Hu. A diversity of chemical databases was collected from academic commercial providers and compounds were pre-processed to convert to an appropriate format with 3D atomic coordination that is suitable for virtual screening. In addition, the databases were also filtered using "Lipinsky's rule of 5": (MW: 200-500; H-bond donor <= 5; H-bond acceptor <= 10; ClogP <= 5; Rotatable bonds <= 10). Sybyl modules SELECTOR, UNITY, and CONCORD were used in the database processing. A total of more than 200,000 unique and drug-like compounds were generated and used in the virtual screening. Database virtual screening was carried out using FlexX running on a 36-processor LINUX cluster in parallel. We sampled the different conformations of YpkA based on model A and model B, which were generated from comparative modeling and possess a distinctive ATP binding site. The key residues in the ATP binding pocket, including residues Arg211, Lys163, His148, Lys272, Asp215, were adjusted using the rotamer library in FlexE. Finally, a total of 10 YpkA conformations were applied for the multiple conformation screening. The active site of YpkA was defined within a 6.5 Å radius sphere centered on the ATP molecule. FlexX was then used for the subsequent docking using the default parameters. The original FlexX scoring function and the CScore was used and the top 30 solutions were retrieved. The top scoring poses and the conformations of the protein active site associated with the selected pose as generated by FlexX were visually examined.

BIOCHEMICAL CHARACTERIZATION AND CRYSTAL STRUCTURE OF THE YPKA GTPASE BINDING DOMAIN

3.1 Domain Characterization of YpkA

Despite the discovery of YpkA over a decade ago, little has been established regarding its structure and function. In order to better understand this essential virulence factor of the plague bacterium, we have taken a biochemical and structural approach in characterizing the different domains of YpkA. To delineate the functional and structural domain boundaries in YpkA, we created a series of constructs based on secondary structure predictions, threading results, and limited proteolysis experiments (Figure 3.1) (Rost 1996; Fischer, Elofsson et al. 2001; Fiser and Sali 2003). The vast majority of constructs were not soluble as recombinant polypeptides expressed in *E coli*. The results of limited proteolysis on the soluble constructs allowed us to further demarcate domain boundaries, and to identify likely surface accessible loops linking separate structural domains in the YpkA polypeptide. These experiments resulted in our division of YpkA into a C-terminal, Rho GTPase and actin-activation domain spanning residues (434-732), a kinase homology domain spanning residues 115-431, and an N-terminal region. These results combined with the crystal structure discussed below, finalized our working model.
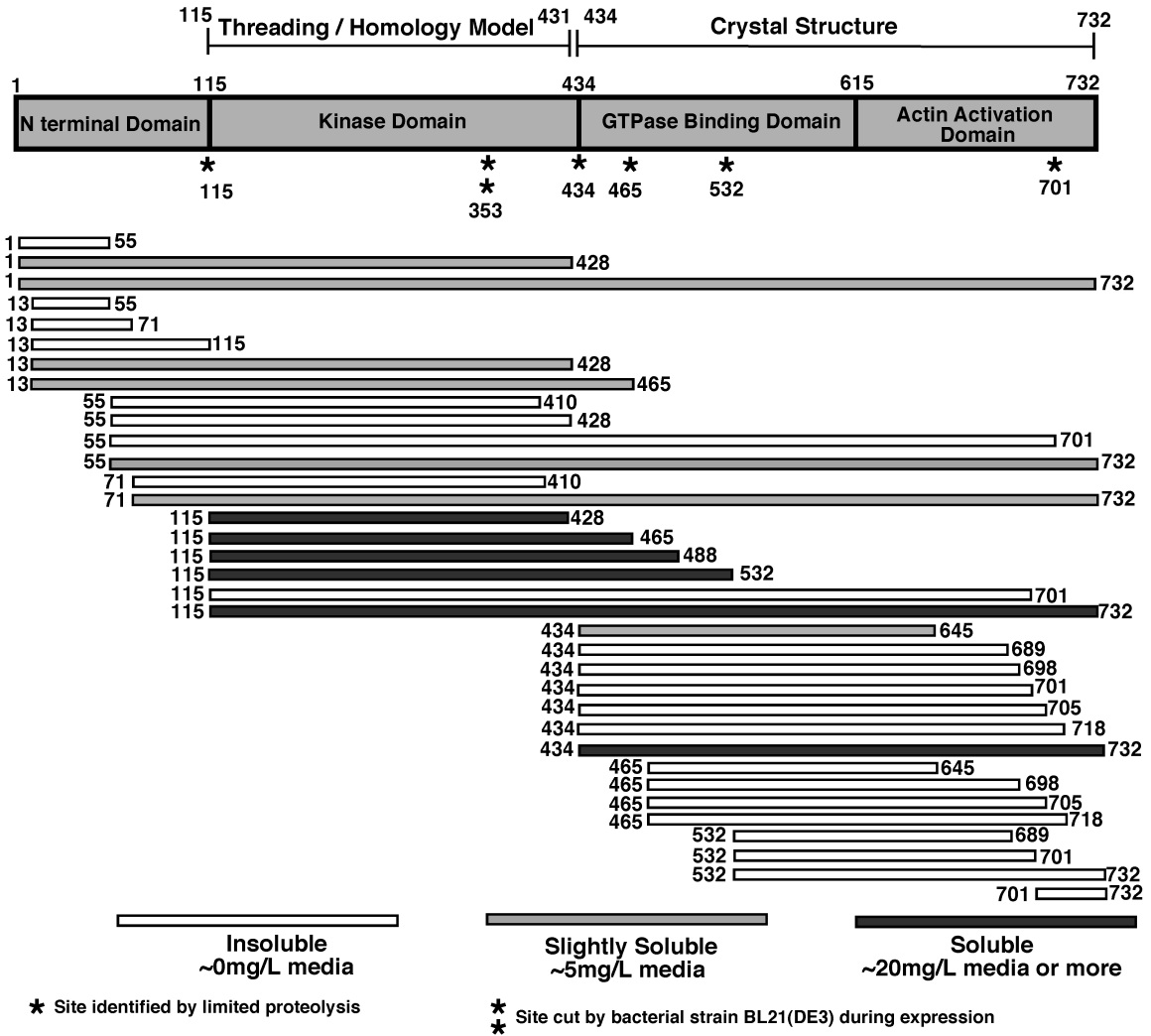
Figure 3.1: Domain Delineation of the *Yerisnia* Virulence Factor YpkA

The domain model of YpkA shows four distinct regions that are demarcated by the dividing residue, and regions of interest are marked by residue number. The region used for the creation of the homology model and sequence of the crystallized construct are outlined, in addition to all locations identified by limited proteolysis experiments are marked with an asterisk. The full domain model was created considering all the known biochemical and structural data.
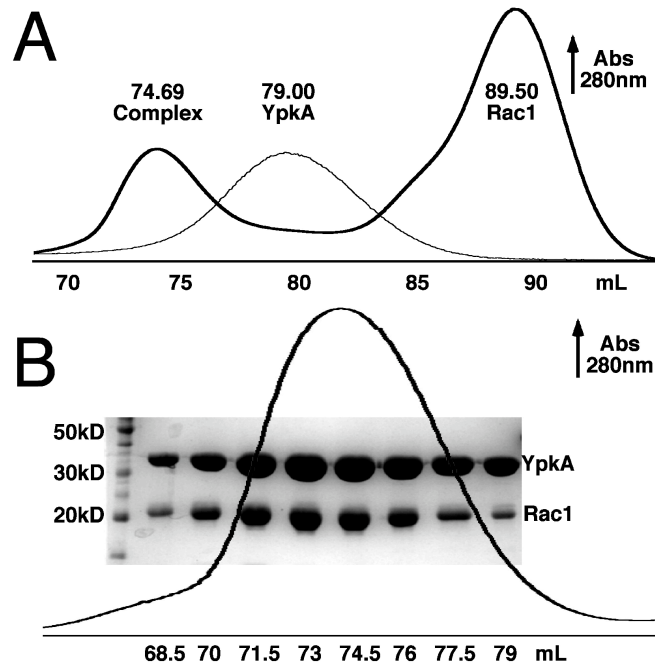
Figure 3.2: YpkA (434-732) Forms a Stable Complex with Human Rac1

(A) Purified YpkA (434-732) and Rac1 bound to GDP form a stable complex as assayed by gel filtration chromatography. The chromatograph of the complex and individual protein peaks are shown in addition to their elution volumes.

(B) The profile of the peak complex fractions re-run on the gel filtration column under the same conditions, overlaid with the fractions visualized by coomassie stained SDS-PAGE.

To analyze the mechanisms behind the activity of YpkA, the soluble domain fragments identified above were purified and assayed for the appropriate biochemical activity. The C-terminal region of YpkA, residues (434-732), was highly soluble and was seen to be active *in vitro* to bind to Rac1 (Figure 3.2). An N-terminal region of this fragment (residues 543-640) has been shown previously in a yeast two-hybrid assay to be required for binding to small GTPases like Rac1, and by sequence similarity was expected to emulate the ACC finger domains of some of their downstream effector kinases(Hakansson, Galyov et al. 1996). Additionally, this C-terminal domain also

contains a potential actin-binding segment, residues 710 to 732, identified again through deletion studies to be required for actin-dependent activation of the YpkA kinase (Juris, Rudolph et al. 2000). The biochemically active construct, YpkA (434-732), was screened for crystallization, yielding diffracting crystals, as discussed below.

3.2 Overall Structure of the C-terminal GTPase Binding Domain

YpkA (434-732) was crystallized in a variety of conditions, but did not yield diffracting crystals until the protein was reductively methylated (Methods) (Figure 3.3) (Rypniewski, Holden et al. 1993). This modification did not affect Rac1 binding (section 3.3), yielded diffraction quality crystals, and the structure was solved by single wavelength anomalous diffraction using methylated, selenomethionine substituted protein (Methods; Table 1).

The overall structure reveals an elongated, all-helical molecule consisting of two distinct sub-domains connected by a 65Å long "backbone" or "linker helix" (Figure 3.4). The N-terminal sub-domain contains most of the sequence-identified, ACC finger-like repeats that resemble elements required in host factors for small GTPase binding, whereas the C-terminal sub-domain contains the sequence implicated in actin activation. The overall surface of the molecule is highly charged, containing a large basic patch in the GTPase binding domain and a large acidic patch in the actin-activation domain. The basic patch stretches around the surface of the GTPase binding domain, and includes the surfaces of helices $\alpha 1$, $\alpha 2$, and $\alpha 3$, in addition to the N-terminus. The acidic patch is comprised of residues from helices $\alpha 9$, $\alpha 10$, and includes the C-terminus (Fig. 3.4). The

Figure 3.3: YpkA (434-732) Crystals

YpkA crystals as imaged under plane-polarized light.

structure of YpkA consists of several disordered regions on the N-terminal face of the molecule. No density or very poor density is seen at the N and C termini, for loops connecting helices α2 and α3 (residues 487 to 488), helices α3 and α4 (residues 515 to 532), and helices α5 and α6 (residues 582 to 584).

YpkA consists of six helices organized into two three-helix bundles packed against each other. Each of the bundles is stabilized by hydrophobic zippering in the core, and extensive hydrophobic packing is observed between the bundles. Although the helices do not form ACC finger domains per-se, this domain possesses a molecular surface that mimics several aspects the RhoA binding surfaces of host ACC finger-like proteins. Previous sequence analysis had identified eight short segments in the GTPase binding domain of YpkA as having similarity to RhoA binding motifs in protein kinase N (PKN)(Juris, Rudolph et al. 2000). The structure suggests that many of these segments possess properties that make them unlikely to serve as GTPase binding elements in a manner similar to ACC finger surfaces. Some of these previously predicted sequences show very little solvent accessibility while others are not located in helical regions but in connecting loops. Other ACC finger-like regions show good accessibility, but are located in regions of the protein which would likely render them unable to bind the GTPase.

A

N Terminal
GTPase Binding Domain

C Terminal
Actin Activation Domain

α2 α3 N α1 α7 α10 α8 C α5 α4 α9 α6

Backbone Helix

B

Basic Patch

Acidic Patch

C

```
                    α1                          α2                        α3
          ooooooooooooooooooooo    oooooooooooooooooooooooooo      oooooooooooooooo
          440        450        460        470        480        490        500
>YpkA  STDVRRITPKKLRELSDLLRTHLSSAATKQLDMGGVLSDLDTMLVALDKAEREGGVDKDQLKSFNSLILK
acc
```

```
                                   α4                        α5
          ooooooooo    ooooooooooooooooo    oooooooooooooo    ooooooooooooo
          510        520        530        540        550        560        570
>YpkA  TYRVIEDYVKGREGDTKNSSTEVSPYHRSNFMLSIVEPSLQRIQKHLDQTHSFSDIGSLVRAHKHLETLL
acc
```

```
                                   α6                        α7
          ooooo    ooooooooooooooooooooooooooooooo    oooooooooooooooooo
          580        590        600        610        620        630        640
>YpkA  EVLVTLSQQGQPVSSETYGFLNRLAEAKITLSQQLNTLQQQQESAKAQLSILINRSGSWADVARQSLQRF
acc
```

```
              α8                        α9
          oooo    oooooooooooooooooo    ooooooooooooooooooo
          650        660        670        680        690        700        710
>YpkA  DSTRPVVKFGTEQYTAIHRQMMAAHAAITLQEVSEFTDDMRNFTVDSIPLLIQLGRSSLMDEHLVEQREK
acc
```

```
          α10
          ooooooooooooooo
          720        730
>YpkA  LRELTTIAERLNRLEREWM
acc
```

Disordered Region
Basic Patch
Acidic Patch

Solvent Exposed
Solvent Accessible
Buried Surface

51

Figure 3.4: Crystal Structure of YpkA (434-732)

(A) A ribbon diagram of YpkA (434-732) is shown. The domain break between the ACC finger like domain and the actin binding domain is indicated by the divide between the two domain labels. Both the N and C termini are indicated, in addition to backbone helix and the two charged patches.

(B) The molecular surface of the YpkA (434-732) crystal structure. Areas of red indicate positive charge, or acidic regions, whereas blue indicates negative charges or basic regions.

(C) Sequence-structural analysis of YpkA (434-732). Secondary structure for given residues is indicated above the sequence. The relative solubility of each residue in the structure is denoted by a colored box beneath it spanning dark blue (solvent exposed) to white (buried from solvent). The residues involved in each charged patch of YpkA are labeled in violet and red for basic and acidic, respectively. Disordered regions are shown in yellow, and putative contacts to Rac1 (see text) are indicated with an orange circle above the residue.

Table I YpkA (434-732) Data collection, phasing and refinement statistics

	Native (single crystal)	SeMet (single crystal)
Data collection		
Space group	$P6_522$	$P6_522$
Cell dimensions		
a, b, c (Å)	60.0, 60.0, 402.02	59.9, 59.9, 402.00
α, β, γ (°)	90, 90, 120	90, 90, 120
Wavelength	0.979	0.979
Resolution (Å)	33.88-2.0	48.45-2.9
R_{sym}	4.5 (36.0)	7.0 (32.0)
$I / \sigma I$	33.1 (3.2)	7.2 (2.3)
Completeness (%)	98.5 (91.7)	99.4 (99.0)
Refinement		
Resolution (Å)	33.88-2.0	48.45-2.9
No. reflections	270529	204549
No. unique reflections	28539	10450
R_{work} / R_{free}	20.9 / 23.8	**Mean Figure of Merit for Phasing SAD**
No. atoms		SOLVE 0.39 (0.35)
Protein	2169	RESOLVE 0.77 (0.56)
Water	276	
B-factors		
Protein	44.9	
Water	57.3	
R.m.s deviations		
Bond lengths (Å)	0.022	
Bond angles (°)	1.744	

These include the N-terminus of the construct which may be partially buried or sterically blocked by the kinase domain in the full length protein, as well as a sequence within the actin-activation sub-domain. Conformational changes or unexpected binding modes may of course invalidate this analysis, but our structural and sequence alignments suggest that residues 489-514 and 533-581 appear to be the best candidates to form an ACC finger-like surface. As shown in Figure 3.5, residues 489-514 align both structurally and by sequence to the known ACC finger structure of PKN. These residues are predicted to
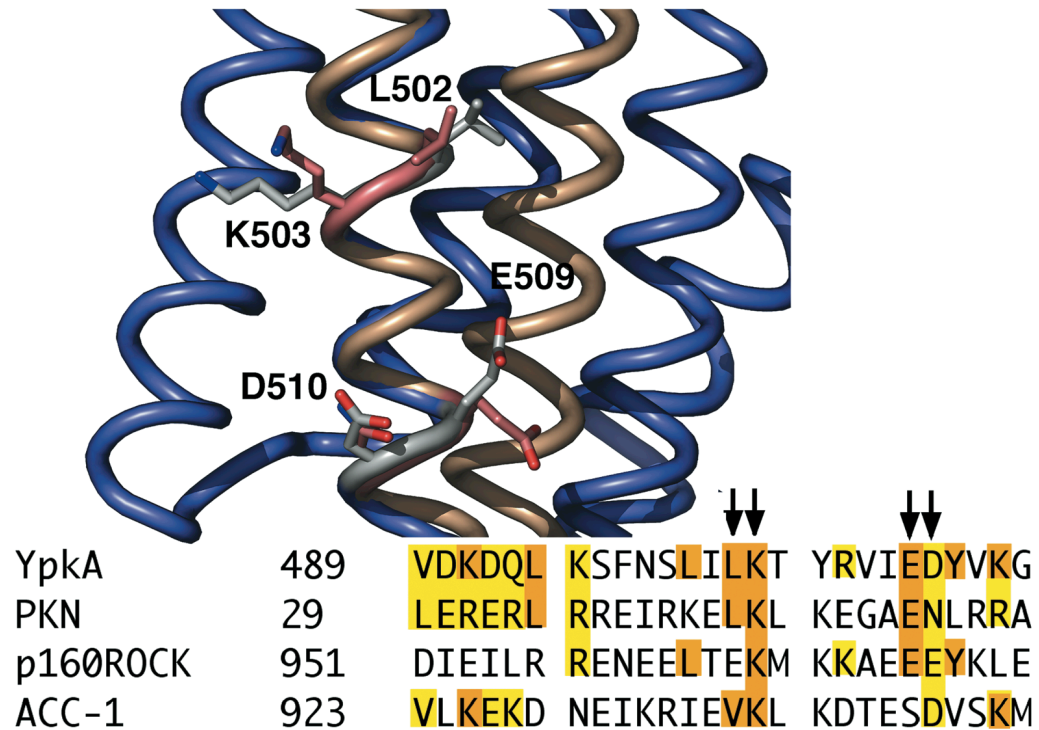
Figure 3.5: Alignment of YpkA with ACC fingers

The residues that align well between YpkA (grey) and PKN (pink) that form part of the PKN/RhoA binding surface are highlighted. The residue numbers of YpkA are indicated. The alignment between the two proteins in this region is shown underneath. Yellow indicates homologous residues and orange shows identical residues. The arrows mark which residues are shown in the structural alignment. Additionally, other ACC finger domains are presented in the sequence alignment. YpkA is *Yersinia* Protein kinase A, PKN is Protein Kinase N, p160Rock is human p160ROCK and ACC-1 is mouse kinectin.
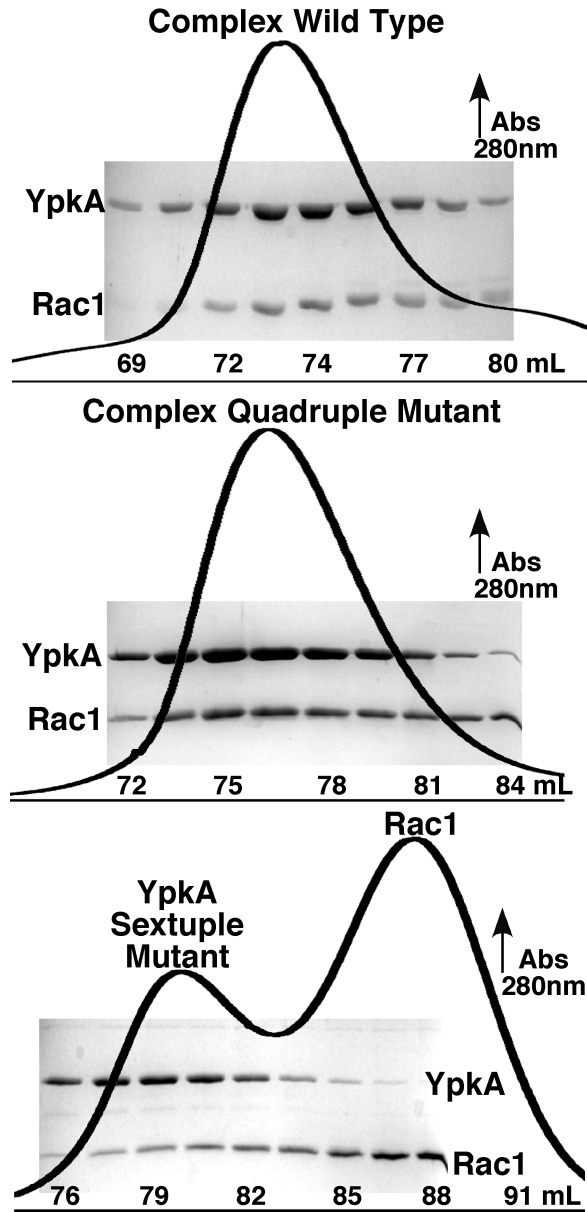
Figure 3.6 ACC finger-like surface mutants do not significantly effect the YpkA-Rac1 Complex

YpkA (434-732) wild type, quadruple mutant (Leu502Gly. Lys503Ala, Glu509Ala, Asp510Ala), and sextuple mutant (Leu502Gly. Lys503Ala, Glu509Ala, Asp510Ala, His566Ala, Glu570Ala) were incubated with Rac1 and analyzed for complex formation. Samples were run on a gel filtration column equilibrated in buffer at 500mM NaCl. The elution profile of the samples from the gel filtration column is shown overlaid with the fractions visualized by coomassie stained SDS-PAGE.

form the α2 helix of the ACC finger, and several of the key residues of PKN that form the interaction with RhoA map to a solvent accessible area of YpkA (Figure 3.4 and Figure 3.5). Associating residues 489-514 with the α2 helix of PKN, residues 533-581 could then comprise a similar surface to the α3 helix of PKN. In this model residues H566 and E570 of YpkA align well with the residues S81 and D85 of PKN, respectively.

To test the importance of the predicted surfaces we performed site directed mutagenesis on a few of the most promising possible GTPase contacting residues of YpkA. We created multiple point mutations, which were used to test binding to Rac1 as compared to wild-type YpkA by a pull-down assay and gel filtration (Methods). All single point mutants were able to form complex with Rac1 at conditions similar to those normally used to assay for binding by pull down assay (data not shown). This is not surprising as in the PKN/RhoA structure the interface consists of 25 hydrogen bonding contacts involving 15 residues (Maesaki, Ihara et al. 1999). As such, limited mutagenesis may not remove enough contacts to prevent complex formation. To address this possibility, we created quadruple and sextuple mutants and assayed their binding by pull down assay and gel filtration in increasing salt concentrations (Figure 3.6). The quadruple mutant consisted of Leu502Gly, Lys503Ala, Glu509Ala, and Asp510Ala (Figure. 3.5) and the sextuple mutant added His566Ala and Glu570Ala. As shown in Fig. 3.6, only at 500mM NaCl, the sextuple mutant complex is destabilized relative to the wild type and quadruple mutant complex, showing that these mutants do not significantly effect the stability of the YpkA-Rac1 complex.

3.3 The Actin Activation Domain

The far C-terminal subdomain of YpkA, containing the polypeptide implicated in actin activation (residues 705-732, corresponding to helix α10), is a novel and elongated fold consisting of four helices clustered into two pairs which only moderately interact with each other. The final two helices of YpkA (α9 and α10), are proximal to the N-terminal subdomain and interact with the backbone helix (α6), creating a small bundle. Toward one end of this bundle, the other helix pair in the C-terminal subdomain (α7 and α8) protrudes out into solution. These two helices (α7 and α8) are separated by an extended loop (residues 640-660) that represents one of the most solvent exposed regions of the protein (Figure 3.4). A loop spanning residues 690-710 (connecting the last two helices) extends toward the N-terminal sub-domain making minor interactions with it. The backbone helix and helix α9 form a surface groove in which the α10 helix rests. The α10 helix (residues 705-730) has been predicted play a role in the interaction of YpkA with actin, as it shows high homology to the actin binding protein coronin (Figure 3.7). Past work has shown that the deletion of the region eliminates both the ability of YpkA to bind to actin, and kinase activity (Juris, Rudolph et al. 2000). As shown in Figure 3.7, the alignment between coronin and YpkA is very high, with several of the conserved residues mapping to solvent accessible areas of the structure (Figure 3.4).

Due to the solvent accessibility of this region of sequence similarity with coronin, we attempted to show an interaction between our C-terminal construct YpkA (434-732) and actin. Both small scale pull-down assays (using GST-YpkA), as well as incubation of purified proteins for the isolation of a complex by gel filtration, similar to the
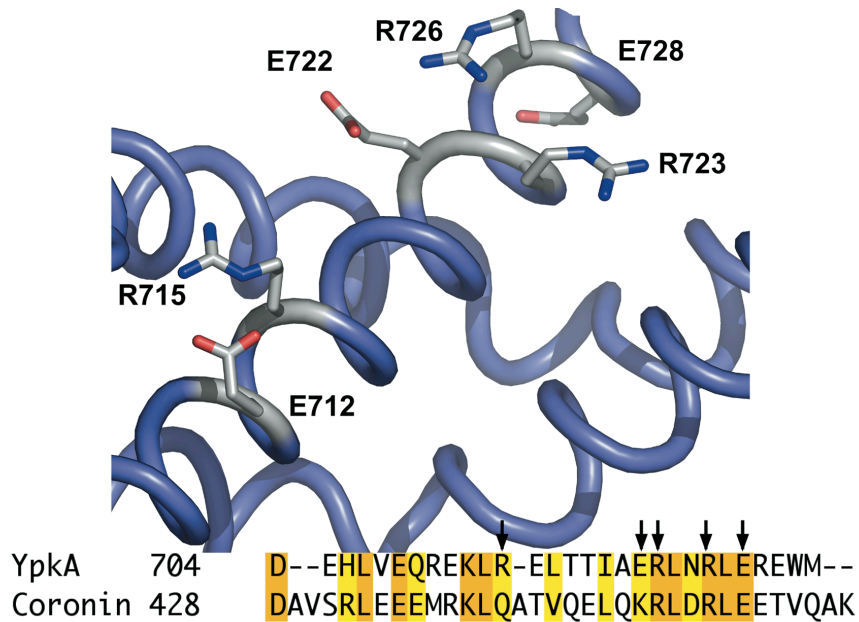
Figure 3.7: Coronin Homology Region

Surface exposed residues in YpkA (434-732) that may form an interaction with actin based on the alignment with coronin are shown. The residue numbers of YpkA are also shown. The arrows mark which residue side chains are shown. Yellow indicates homologous residues and orange shows identical residues.

interaction of YpkA (434-732) and Rac1, were performed (Methods, Fig. 2). These attempts to form a stable complex between YpkA (434-732) and actin were performed in the buffer used in radiological assays, in which it has been shown that actin does interact with and activate full length YpkA (Figure 3.8) (Dukuzumuremyi, Rosqvist et al. 2000; Juris, Rudolph et al. 2000). In contrast to our binding assays with Rac1, these experiments show that no stable complex was formed between YpkA (434-732) and actin (data not shown). This either indicates that the interaction between YpkA and actin is weak or transient, or that residues (434-732) are not sufficient for a stable interaction with actin.
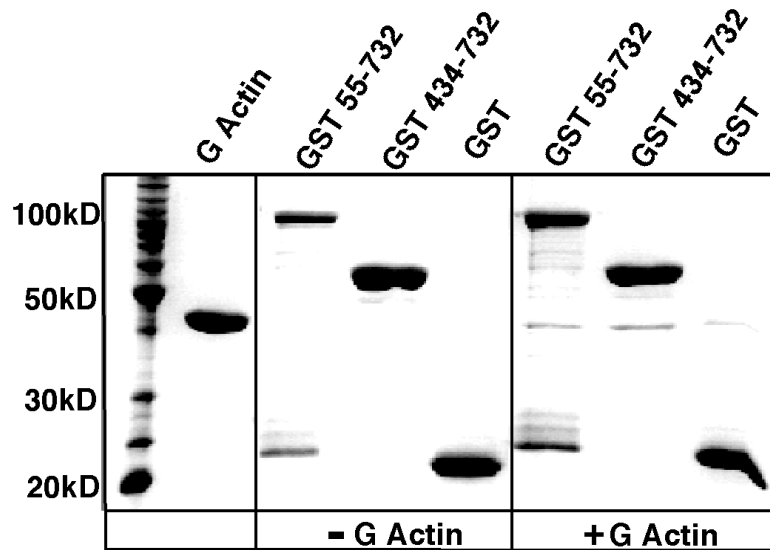
Figure 3.8: YpkA (434-732) forms a weak complex with G Actin
GST-YpkA constructs were bound to affinity resin and incubated in kinase activity buffer (see methods) in the presence of G actin. Beads were precipitated, washed, and the remaining material run on an SDS-page gel followed by coomassie stain.

Previous work to isolate the residues that are responsible for an interaction with actin utilized deletion mutagenesis in the context of the full-length protein, in which the removal of the last 20 amino acids of YpkA rendered it unable to bind or to be activated by actin (Juris, Rudolph et al. 2000). As discussed above, we have found that all of the C-terminal deletion constructs are poorly soluble and are thus most likely poorly folded (Figure 3.1). Specifically, we created several C-terminal deletions removing the last helix, all of which were significantly destabilized. The structure reveals that a removal of this helical region would extensively expose the hydrophobic core and remove a large segment stabilizing the fold of this sub-domain. The α10 helix buries hydrophobic residues in all three of the other helices in the C-terminal subdomain. This data strongly

suggests that the removal this segment destabilizes the fold, and/or induces aggregation, preventing association with actin and thereby explaining the loss of kinase activity. In agreement with these results, recent results have show that the last 20 residues of YpkA in fact only forms part of the actin-interaction surface, and therefore would be necessary but not sufficient for actin binding (Trasak, Zenner et al. 2007).

ANALYSIS OF A YPKA-RAC1 COMPLEX CRYSTAL STRUCTURE

4.1 Overall Structure of a YpkA-Rac1 Complex
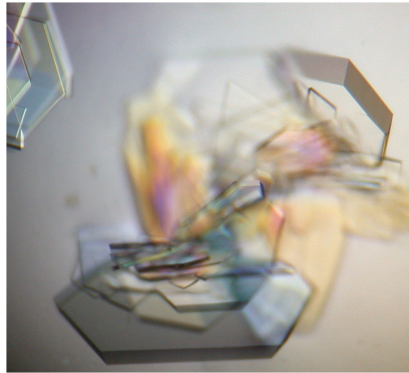
Figure 4.1: YpkA-Rac1 Complex Crystals

YpkA Rac1 crystals as imaged under plane polarized light.

As the crystal structure of the YpkA GTPase binding domain alone was insufficient to determine the Rac1 binding surface, and thus the biochemical function of YpkA, crystals of a YpkA-Rac1 complex were grown (Figure 4.1; Table II). Although YpkA (434-732) and Rac1 form a dimer in solution (Figure 3.2), the complex crystallized as a 2:2 packing (YpkA:Rac1) in the asymmetric unit with a non-crystallographic two-fold axis of symmetry (Figure 4.2). Each molecule of YpkA (434-732) therefore makes contacts to two different Rac1 molecules (Figure 4.2). The sites of these contacts we have labeled "A" and "B." Contact A interacts nearly exclusively with the regulatory Switch I and Switch II regions of the GTPase, whereas Contact B interacts with two C-terminal helices adjacent to the RhoGTPase "insertion" (relative to the Ras-family small
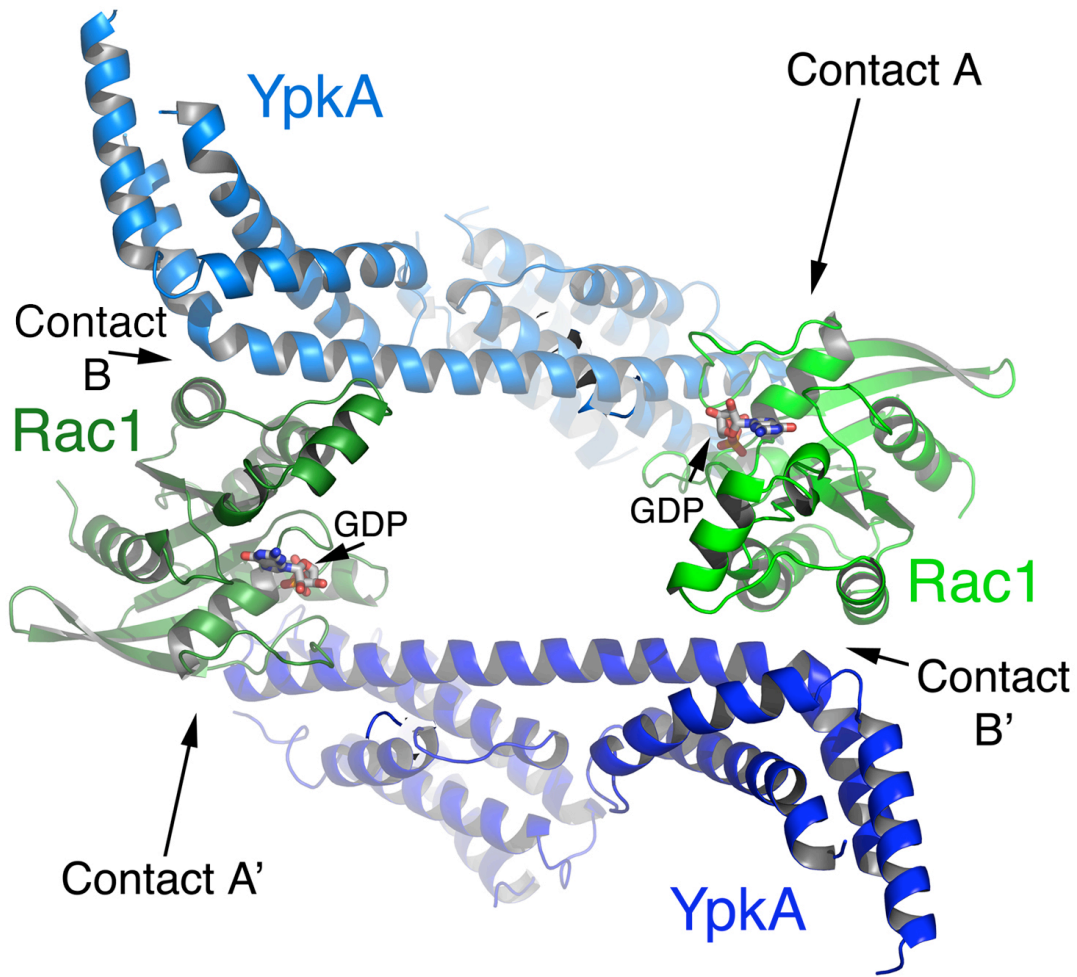
Figure 4.2: Overall Structure of the YpkA-Rac1 Complex

2:2 YpkA:Rac1 contents of the asymmetric unit shown as ribbon diagrams. The two separate contacts of YpkA to Rac1 are shown, as is the nucleotide present, GDP.

GTPases), distal to the nucleotide and Switch regions. Four facts argue for contact A being the biologically relevant interaction: (1) it has nearly twice the buried surface area as contact B, (2) it contacts the critical Switch I and Switch II regions of Rac1, whereas contact B does not (Figure 4.4), (3) mutations of contact A disrupt YpkA-Rac1 interactions, whereas mutations in contact B (N627A, R628A, S631A) do not (see below), and (4) a heterodimer appears to be the biological unit as judged by biochemical

experiments (Figure 4.5). Additionally, a comparison between YpkA (434-732) crystallized alone and YpkA (434-732) in complex with Rac1 shows that there are very few conformational changes in the GTPase binding domain structure (Figure 4.3). Most differences are located in the C-terminal subdomain of YpkA, and involve a slight overall displacement in the positioning of the $\alpha7$ and $\alpha8$ helices, as well as alterations in the conformation and relative disorder of solvent exposed loops. Rac1 is little altered by the binding of YpkA, except for the Switch regions as discussed below.

Table II YpkA (434-732) Rac1 complex data collection and refinement statistics

Data collection	
Space group	P1
Cell dimensions	
a, b, c (Å)	66.4, 75.5, 99.8
α, β, γ (°)	92.1, 103.4, 115.8
Resolution (Å)	99.5-2.60
R_{sym} or R_{merge}	9.0 (41.8)
$I / \sigma I$	11.8 (3.0)
Completeness (%)	99.4 (99.4)
Redundancy	3.7
Refinement	
Resolution (Å)	34.86-2.60
No. reflections	48710
R_{work} / R_{free}	22.2 / 25.7
No. atoms	
All atoms	7191
Water	75
B-factors	
All atoms	62.65
Water	62.59
R.m.s deviations	
Bond lengths (Å)	0.017
Bond angles (°)	1.666
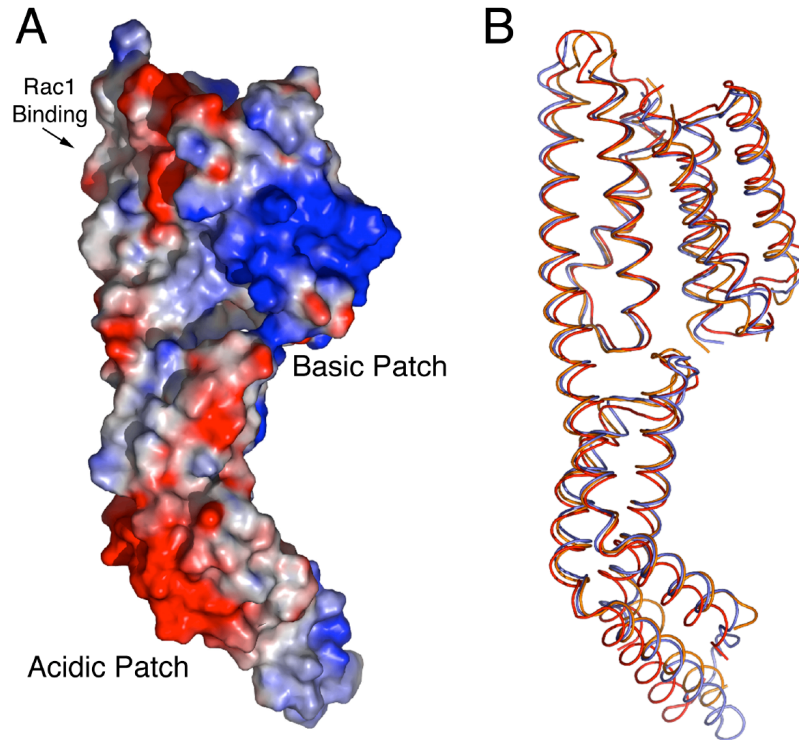
Figure 4.3: Crystal Structures of the C-terminal Domain of YpkA

(**A**) The molecular surface of YpkA (434-732) with charges is shown colored by electrostatic potential, where blue indicates negatively charged regions and red indicates positive charged regions.

(**B**) The structural alignment of the YpkA monomer structure (blue) and both YpkA molecules in the asymmetric unit of the crystal structure (one chain in red and the other in orange).

4.2 The YpkA-Rac1 Interface

YpkA and Rac1 form an interface burying roughly 1,600 Å^2 and limited to
residues 573-601 of YpkA (spanning the helices α5 and α6) contacting the key
regulatory Switch I and Switch II regions of Rac1 (Figure 4.4, A-E). Switch I and
Switch II together create a concave pocket into which the α6 (backbone) helix of YpkA
inserts (Figure 4.4B), which along with the clustering of the Switch II helix with the α5
and α6 helices, cement the interaction tightly.

The Switch I contacts involve a large number of hydrophobic/van der Waals
interactions with YpkA, as well as several hydrogen bonds at this surface (e.g., residues
Asp 38 (Rac1) and Arg 596 (YpkA), Glu 599 (YpkA) and Tyr 32 and The 35 of Rac1;
Figure 4.4, C and E). A striking aspect of this interaction with Switch I is that YpkA
contacts the conserved Thr 35 residue, which is normally involved in magnesium ion
coordination (Dvorsky and Ahmadian 2004), resulting in a stable coordination network
between the side chain of YpkA Glu 599, Rac1 Tyr 32, the hydroxyl group of Thr 35, its
main chain carbonyl oxygen, the magnesium ion, and the hydroxyl group of Thr 17
(Figure 4.4C).

At Switch II, the molecular interface between YpkA and Rac1 is quite extensive,
forming the largest portion of the protein-protein interactions (Figure 4.4, A, B, D and E).
Most of the contacting residues between YpkA and Rac1 are involved in hydrophobic,
van der Waals interactions forming a hydrophobic interface between the α5 and α6
helices of YpkA and the alpha helix and surrounding loops of Switch II. Residues Leu
573, Val 577, and Val 586 of YpkA along with Leu 67 and Leu 70 of Rac1 contribute
significantly to this hydrophobic interface, making what resembles a hydrophobic
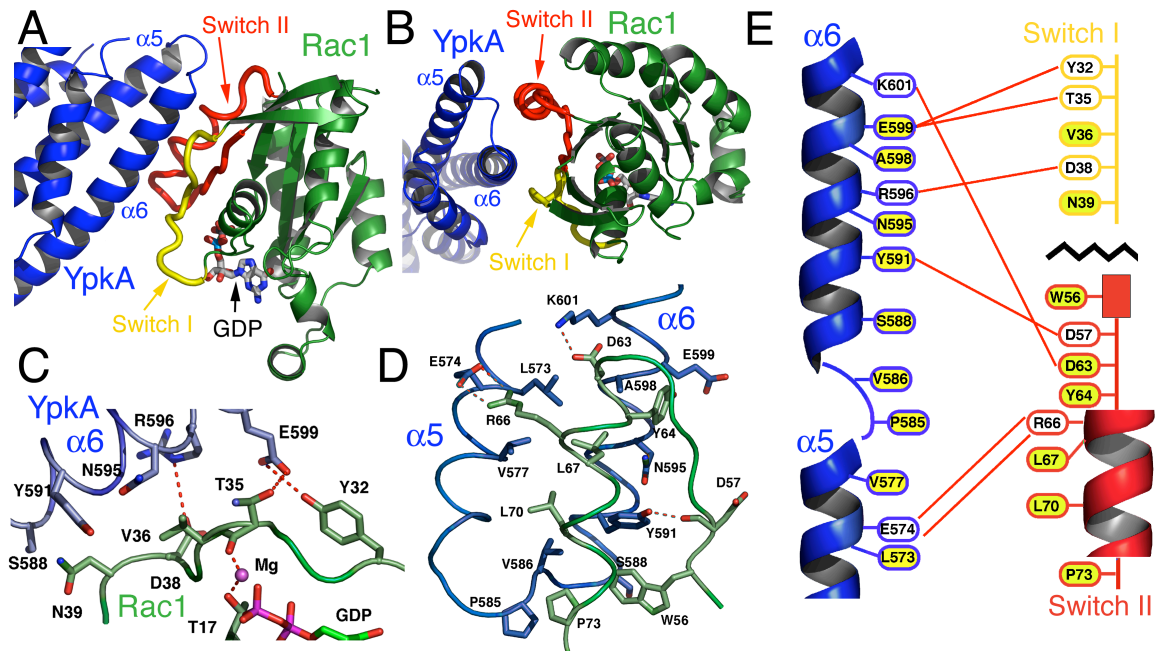
Figure 4.4: The YpkA-Rac1 Interface

(**A**) Ribbon diagram view of the contact A interaction between YpkA and Rac1. Switch I (yellow) and Switch II (red) are highlighted, and GDP is noted.

(**B**) The image in panel (A) rotated by 90 degrees about a horizontal axis.

(**C**) Close up of the Switch I interactions with the α6 helix of YpkA. Hydrogen bonds are denoted by dashed red lines.

(**D**) Close up of the Switch II interactions with the α5 and α6 helices of YpkA. Hydrogen bonds are denoted by dashed red lines.

(**E**) Schematic of the interactions of the α5 and α6 helices of YpkA with Switch I and Switch II of Rac1. Hydrogen bonds are indicated by red lines between the interacting residues, and hydrophobic interactions are shown with a yellow background.

zippering motif between the α5 and α6 helices of YpkA and the helical element of Switch II (Figure 4.4, A, B, and D). This interface between the molecules also includes YpkA Pro 585 contacting Rac1 Pro 73 at one end, "pinching" off the zipper, along with Tyr 591, Asn 595, and Ala 598 (YpkA) and Trp 56 and Asp 63 (Rac1), making several van der Waals contacts between each other and the hydrophobic zippering motif.

Additionally, Rac1 Tyr 64 makes van der Waals contacts with YpkA Glu 599, possibly helping to position it for the hydrogen bonding interactions to Tyr 32 and Thr 35 in the Switch I region of Rac1 described above. Although the binding surface is primarily hydrophobic in nature, it is also stabilized by three polar interactions. The main chain carbonyl oxygen of Rac1 Asp 57 forms a hydrogen bond with the phenol group hydroxyl of YpkA Tyr 591, the acidic group of Rac1 Asp 63 with the amino group of YpkA Lys 601, and a salt bridge is formed between Rac1 Arg 66 and YpkA Asp 574 (Figure 4.4D).

To examine the importance of these contacts to complex stability, we mutated three residues of YpkA at the contact A interaction surface and tested the purified protein for binding to Rac1 and RhoA. The triple mutant consisted of loss-of-contact mutations to alanine (Tyr 591Ala, Asn595Ala, and Glu599Ala). Although these mutations do not destabilize YpkA itself, they completely abolish complex formation with both Rac1 and RhoA (Figure 4.5), strongly supporting the crystallographic analysis. In contrast, mutations in the second binding site in the crystals (contact B, Figure 4.2) had no effect on complex formation (Figure 4.5). The mutated residues at contact B selected (N627A, R628A, S631A) comprise almost the entire interaction surface.

A

Adolase 158kD
Albumin 67kD
Ovalbumin 43kD
Chymotrypsinogen A 25kD
Ribonuclease A 13.7kD

YpkA 434-732 Rac1
YpkA 434-732
Rac1

Abs 280nm

mL

B Contact A Mutant

C Contact B Mutant

D RhoA

E Cdc42
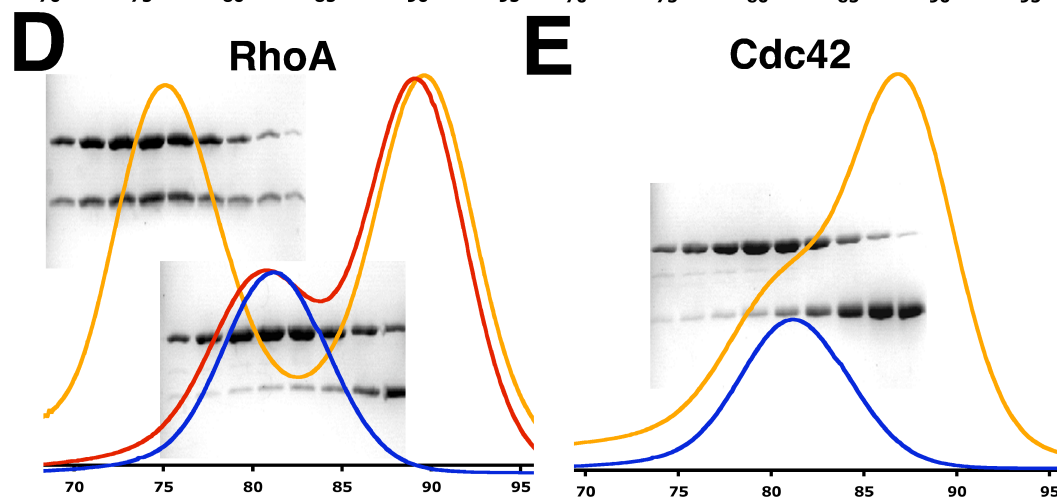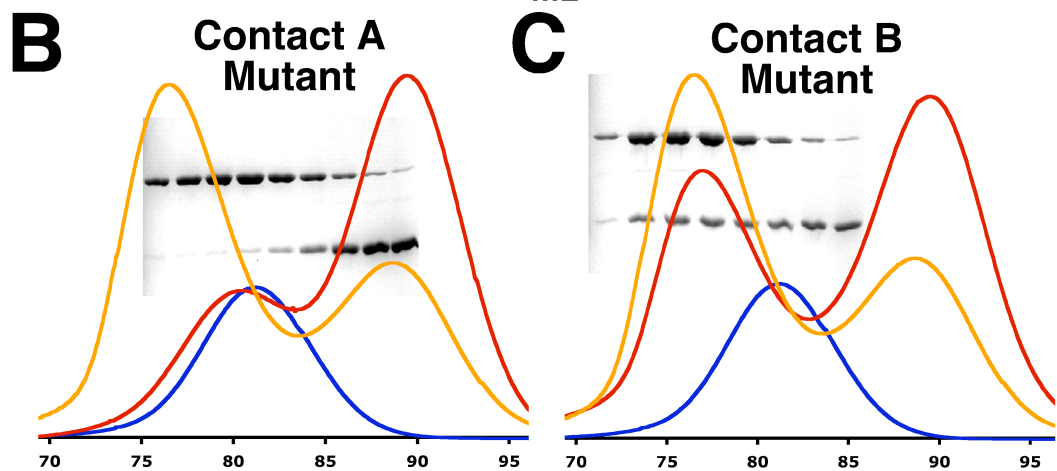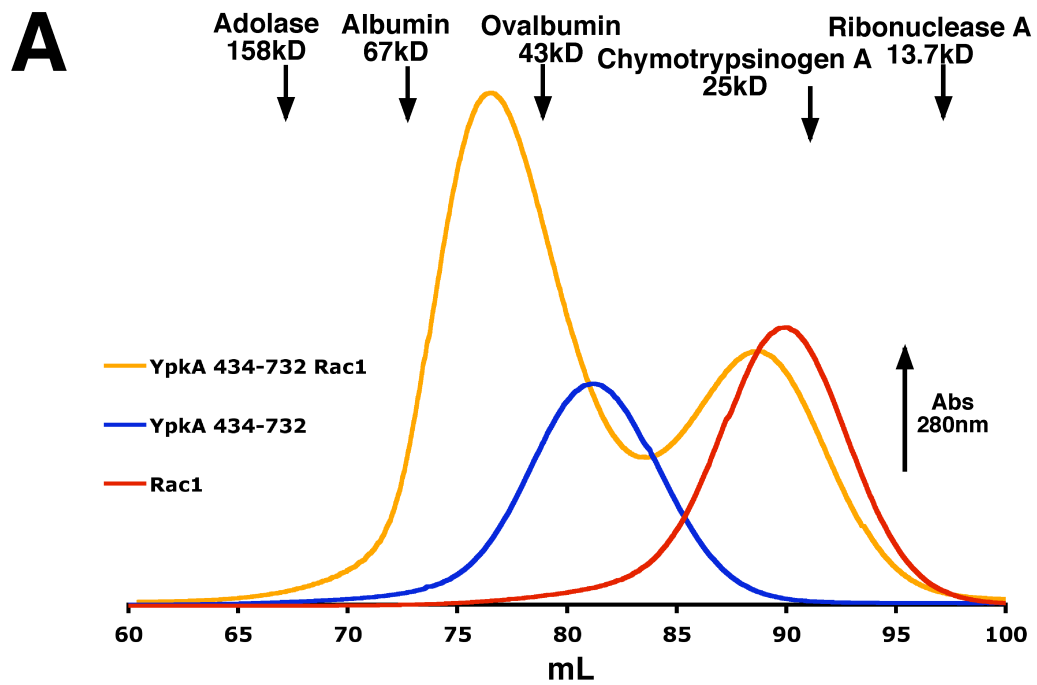
YpkA 434-732 + Small GTPase
YpkA 434-732 mutant + Small GTPase
YpkA 434-732

68

Figure 4.5: Gel Filtration and Mutagenesis

(A) The gel filtration profile of YpkA (434-732). The molecular weight standards used in the column calibration are labeled above the graph with the arrows indicating their approximate elution volume. YpkA (434-732) is shown in blue, Rac1 in red, and the YpkA (434-732) and Rac1 complex in orange.

(B) The gel filtration profile of the YpkA (434-732) Contact A mutant.

(C) The gel filtration profile of the YpkA (434-732) Contact B mutant.

(D) The gel filtration profile of the YpkA (434-732) interaction with RhoA.

(E) The gel filtration profile of the YpkA (434-732) interaction with Cdc42.

In panels (B) and (C), YpkA (434-732) contact A or contact B mutant incubated with Rac1 is shown in red, the YpkA (434-732) and Rac1 complex is shown in orange. In panel (D), YpkA (434-732) incubated in the presence of RhoA is shown in orange and the YpkA (434-732) Contact A mutant and RhoA complex is shown in red. In panel (E) YpkA (434-732) in the presence of Cdc42 is shown in orange. All panels contain the elution peak of YpkA (434-732) shown in blue and the SDS-PAGE analysis (stained with Coomassie blue) of fractions containing the eluted material superimposed on the corresponding chromatograms.

4.3 YpkA Mimics Host GDI Proteins

As described previously, the manner in which YpkA binds to Rac1 is by the use of a region in its long "linker-helix" to interdigitate between the two switch regions (Figure 4.4D). This preliminary observation was interesting, as YpkA seems to be at least structurally mimicking a common mode of binding to the small GTPases. A large number of eukaryotic host factors bind the small GTPases in this manner, by the interdigitation of a helix between Switch I and Switch II (Cherfils 2001). Examples include, Arfaptin (which binds to Rac1 GDP or GTP or Arf GTP to mediate cross-talk between those small GTPase pathways) (Tarricone, Xiao et al. 2001), Tiam1 (which is a host GEF) (Worthylake, Rossman et al. 2000), Rabphillin (which binds to the Rab small GTPases and is involved in vesicle trafficking and neurotransmitter release) (Ostermeier and Brunger 1999), and the kinases PKN and ROCKI, which are downstream effectors of the Rho family of GTPases that are activated upon the binding of the GTP form of the GTPase (Maesaki, Ihara et al. 1999; Dvorsky, Blumenstein et al. 2004). The proteins listed above are only a few of the possible examples, and despite their common binding motifs, they all have different biochemical and biological roles within the cell. Structurally, Arfaptin and Tiam1 both hold Switch I in an open state, the difference being that through additional elements Tiam1 can catalyze nucleotide exchange, whereas arfaptin simply primes Rac1 (binding either the GDP or GTP form) for exchange by a GEF in response to the binding of an activated (GTP bound) small GTPase from the Arf family (Tarricone, Xiao et al. 2001) (Cherfils and Chardin 1999; Cherfils 2001). Rabphillin is recruited to the cell membrane by binding to activated (GTP bound) Rab3A,
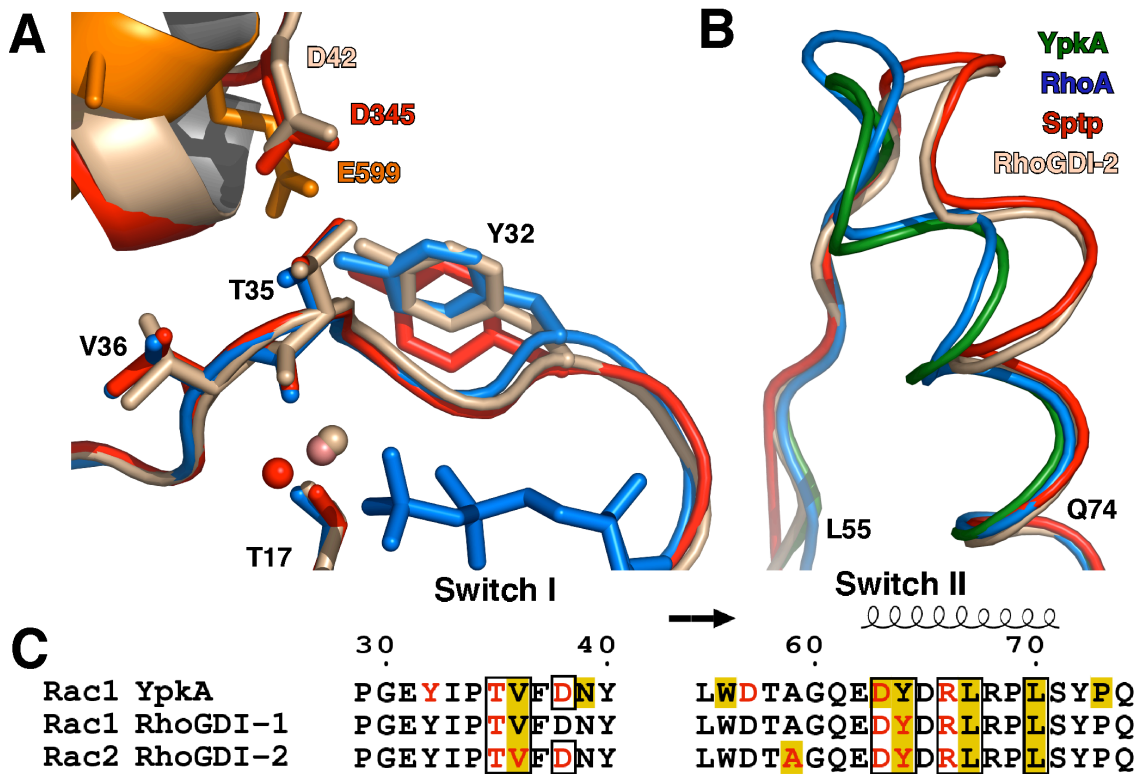
Figure 4.6: YpkA Mimics the Binding of Host Cell GDIs

(**A**) Structural alignment between YpkA-Rac1 with the RhoGDI-1(α)-Rac1 and the RhoGDI-2 (LyGDI)-Rac2 complexes at Switch I. YpkA is indicated in orange, Rac1 bound to YpkA is indicated in blue, the RhoGDI-1(α) Rac1 structure in red and the RhoGDI-2(LyGDI) Rac2 structure in tan. The GDP molecule from the YpkA-Rac1 complex structure is shown in blue.

(**B**) Structural alignment between the YpkA-Rac1, RhoAGDP, SptP-Rac1, and RhoGDI-2 (LyGDI)-Rac2 complexes at Switch II. The YpkA complex is indicated in green, the RhoA-GDP structure in blue, the SptP-Rac1-GDP-AlF$_3$ complex in red, and the RhoGDI-2 (LyGDI)-Rac2 complex in tan. The residues at the N and C terminus of the region in Rac1 shown are labeled.

(**C**) The conserved interactions of YpkA and GDI proteins with Rho family GTPases is shown. Both Switch I and Switch II are indicated with the secondary structural elements of the YpkA-Rac1 structure drawn above. Residues involved in hydrogen bonding are drawn in red and those residues making hydrophobic or Van der Waals contacts are highlighted in yellow. Similar interactions between YpkA and the RhoGDI proteins are boxed.

71

to assist in the merging of vesicles for neurotransmitter release (Ostermeier and Brunger 1999), and both PKN and ROCKI are kinases activated by binding to the GTP bound form of RhoA, although PKN appears to have two binding motifs, one of which binds across a region near Switch I and another which is the interdigitated helix (Maesaki, Ihara et al. 1999). Since the interdigitation of a helix between Switch I and Switch II is such a common and biologically diverse binding motif, to examine the effect of YpkA on Rac1 from a structural perspective, Rac1 was instead used as the search model, as described below.

A comparison of the *Yersinia* YpkA interaction with Rac1 reveals intriguing similarities to the interactions between host cell guanine nucleotide dissociation inhibitors (GDIs) and their interactions with the Rho family GTPases. An alignment of the complexes of RhoGDI-1(α)-Rac1 (Grizot, Faure et al. 2001) and RhoGDI-2(LyGDI)-Rac2 (Scheffzek, Stephan et al. 2000) with YpkA-Rac1 reveals that the Switch I polypeptide adopts a very similar conformation in all of these structures (Figure 4.6A). In addition, the residues of Switch I that make contacts with the host cell GDIs possess nearly identical conformations in the Rac1 complex with YpkA (Figure 4.6, A and C). YpkA even appears to mimic a hallmark of the RhoGDI-small GTPase interaction by using an acidic residue to contact Thr 35 of the small GTPase to form a highly stable coordination network involving YpkA Glu 599, the hydroxyl group of Thr 35, the carbonyl oxygen of Thr 35 with the magnesium ion, and the hydroxyl oxygen of Thr 17 (Scheffzek, Stephan et al. 2000; Grizot, Faure et al. 2001). This structural stabilization of Switch I in the GDP bound conformation results in an inhibition of nucleotide exchange.

YpkA, therefore, possesses the key contacts seen between host RhoGDIs and their target GTPases at Switch I (Figure 4.6, A and C).

YpkA contacts many of the same residues in Switch II as RhoGDI, although the details of the molecular interactions differ (Figure 4.6C). For example, the conformation of Switch II in the YpkA Rac1 structure is different from that seen in RhoGDI complexes with small GTPases, where Switch II is found in a conformation nearly identical to the GTP bound forms of the small G proteins. YpkA instead contacts these residues in Switch II in such as manner as to stabilize the region in a conformation similar to the structures of the GTPases bound to GDP. Indeed, both Switch I and Switch II of Rac1 in the YpkA structure are almost identical to the crystal structure of RhoA bound to GDP (Wei, Zhang et al. 1997), indicating that YpkA acts to lock the small GTPase in the GDP bound, or physiologically "off" conformation. RhoGDIs are so named for their ability to maintain small GTPases in an "inactive" physiological state, specifically through the inhibition of nucleotide exchange in the small GTPases. This inhibition is achieved by stabilizing the coordination of the magnesium ion and preventing both the ion and the bound GDP/GTP from dissociating (Scheffzek, Stephan et al. 2000). This prevents both intrinsic exchange of the GDP/GTP nucleotide and that catalyzed by guanine nucleotide exchange factors, or GEFs.

Given the similarities at the structural level with host RhoGDIs, we therefore examined whether YpkA could inhibit the intrinsic and catalyzed exchange of nucleotides for Rac1, RhoA, and Cdc42. As shown in Figure 4.7A, when Rac1, RhoA or Cdc42 are incubated in the presence of mant-GTP, a fluorescent nucleotide analog, there is an increase of fluorescence signal over time, indicating that increasing amounts of the
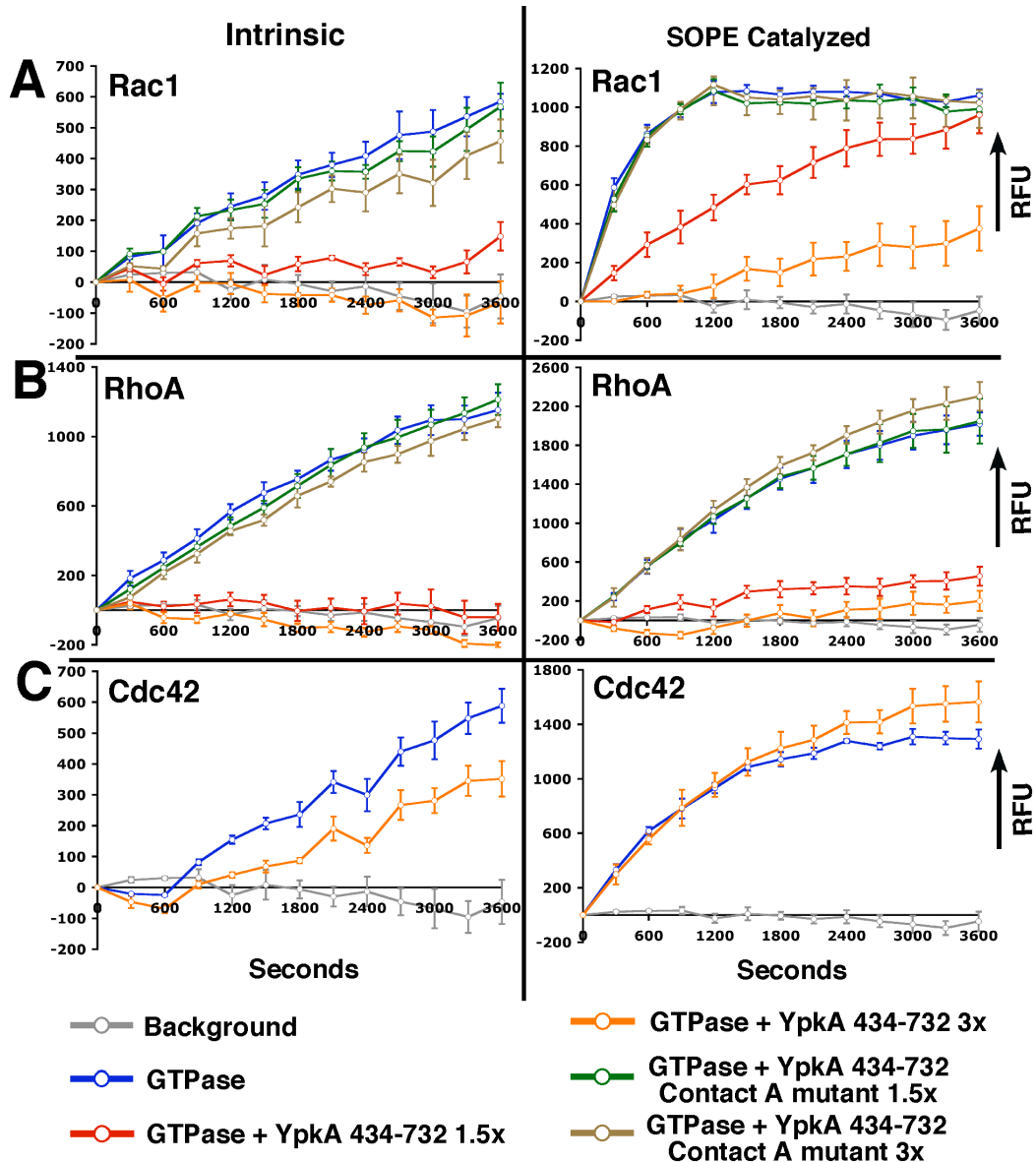
Figure 4.7: YpkA Inhibits Nucleotide Exchange in Rac1 and RhoA

The intrinsic and GEF catalyzed rate of mant-GTP exchange into the small GTPases Rac1, RhoA, and Cdc42 was monitored as described in Experimental Procedures. The left column shows the intrinsic rate inhibition data, and the right column the rate inhibition data when challenged with the *Salmonella* GEF SopE. Panel (A) is Rac1, panel (B) is RhoA, and panel (C) is Cdc42. The experiments are shown with the increase of relative fluorescence units (RFU) over time in seconds. The labels are described in a legend below the Figure, where 1.5x and 3x are relative molar concentrations of YpkA or YpkA Contact A mutant above the small GTPase.

74

labeled nucleotide have been bound (Experimental Procedures). The addition of YpkA (434-732) results in a marked decrease in nucleotide exchange for Rac1 and RhoA, but only a modest decrease for Cdc42 (Figure 4.7A). In contrast, the triple mutant (contact A mutant) of YpkA that is impaired in Rac1 and RhoA binding was unable to inhibit the exchange of GTP (Figure 4.7A). Similar results are observed if we challenge the system with a potent GEF, *Salmonella* SopE (Hardt, Chen et al. 1998; Rudolph, Weise et al. 1999; Friebel and Hardt 2000; Buchwald, Friebel et al. 2002). SopE, one of the most active exchange factors studied to date (Rudolph, Weise et al. 1999; Friebel and Hardt 2000), quickly catalyzed the exchange of the bound GDP of Rac1, RhoA and Cdc42 with the mant-GTP assay (Figure 4.7B). Incubation of Rac1 or RhoA with increasing amounts of YpkA leads to a dose-dependent decrease in exchange when challenged by SopE (Figure 4.7B). In contrast, YpkA was unable to inhibit exchange in Cdc42 catalyzed by SopE (Figure 4.7C). As was observed with the intrinsic nucleotide exchange of Rac1 and RhoA, the triple mutant of YpkA was completely inactive against SopE. We therefore conclude that the YpkA C-terminal domain is a potent inhibitor of nucleotide exchange for Rac1 and RhoA *in vitro*.

4.4 The YpkA GTPase Binding Domain Disrupts the Actin Cytoskeleton

We then sought to establish the *in vivo* significance of binding to RhoA and Rac1, and thereby, presumably, the biological significance of the GDI-like activity of YpkA. To address this, we transfected cultured human intestinal epithelial cells (Henle407) with six YpkA constructs: 1-732 (wild type), 1-732 (K272A, kinase active site mutant), 1-732
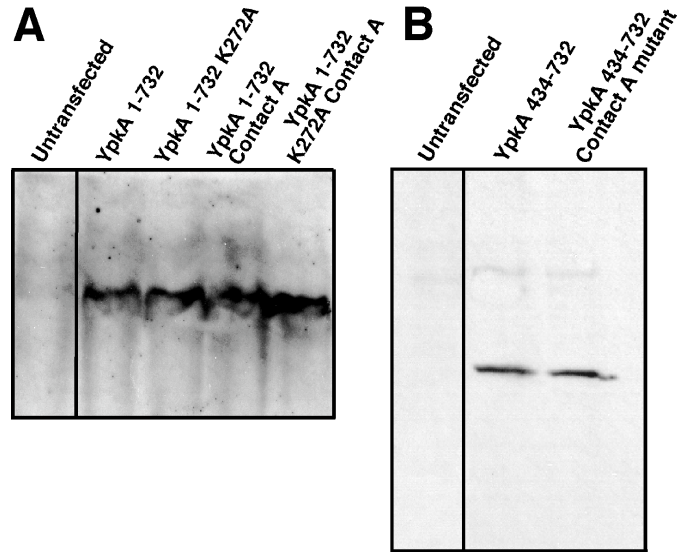
Figure 4.8: Relative *in vivo* stability of transfected YpkA constructs

(A) Transfected and untransfected Henle407 cells expressing N-terminal FLAG-tagged YpkA full length constructs as analyzed by western blot (methods).

(B) Transfected and untransfected Henle407 cells expressing either N-terminally FLAG-tagged YpkA (434-732) or YpkA (434-732) Contact A mutant as analyzed by western blot (methods)

(contact A mutant), 1-732 (K272A + contact A mutant), (434-732), and (434-732) contact A mutant (Experimental Procedures). After 24 hours the cells were immuno-stained for the YpkA constructs with antibodies against an N-terminal FLAG epitope (Experimental Procedures), and the cytoskeleton visualized by staining with Rhodamine Phalloidin. Additionally, each construct was tested for its stability in mammalian cells showing that transfected YpkA is expressed and present within the experiment at similar levels (Figure 4.8). As has been reported (Hakansson, Galyov et al. 1996; Juris, Rudolph et al. 2000; Nejedlik, Pierfelice et al. 2004), transfection with wild type YpkA induces extensive cytoskeletal disruption, leading to a loss of actin stress fibers and a distortion in the cellular morphology. In fact, transfection of YpkA into mammalian cells induces two

observable effects, the first causing the disappearance of actin stress fibers, and the second resulting in severe cellular deformation in addition to the loss of stress fiber formation (Figure 4.9). This cellular deformation has been described as a "rounding up" of the cells, but maintaining focal adhesions (Juris, Rudolph et al. 2000). Full length YpkA shows the ability to both prevent stress fiber formation and causes a large number of the Henle cells to lose their shape, resulting in the described cellular deformation or "wild type" effect (Figure 4.9B). Transfection with the N-terminal deletion construct YpkA (434-732), which lacks the serine-threonine kinase domain, results in an attenuated or "intermediate" phenotype. YpkA (434-732) is able to cause the disappearance of actin stress fibers in most cells, although its ability to cause cellular deformation is reduced as compared to the wild type construct (Figure 4.9A). The YpkA (434-732) contact A mutant, which is deficient for both Rac1 and RhoA binding in addition to GDP dissociation inhibition *in vitro*, leads to a complete loss of cytoskeletal disruption (Figure 4.9B).

In the context of the full length protein, these same mutations are intriguing. To begin, YpkA with the kinase active site mutation K272A is only slightly attenuated in cytoskeletal disruption, losing most of the "wild type effect" but maintaining a significant amount of the intermediate effect (Figure 4.9A). In fact, this mutant is nearly identical in its effect on cells to the deletion of the entire kinase domain. In contrast, the YpkA contact A mutant, defective for binding to Rac1 and RhoA (and defective for GDI-like activity) causes a very low level of cytoskeletal alterations, and adding to this mutant the K272A kinase mutation does not appreciably change the phenotype. Although the full length YpkA Contact A mutant and the K272A Contact A mutant do not cause significant

A
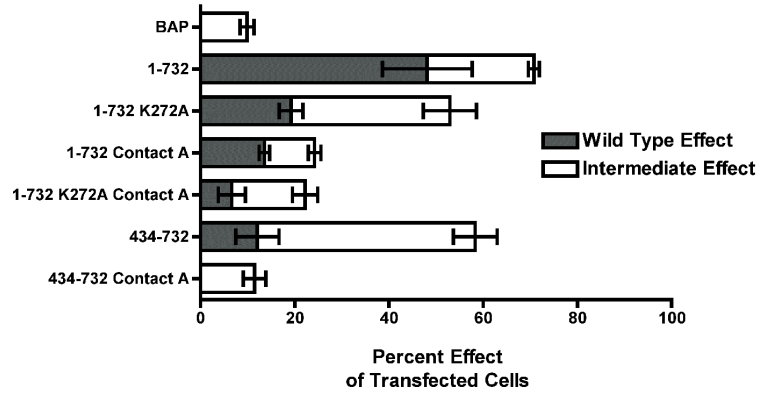
B
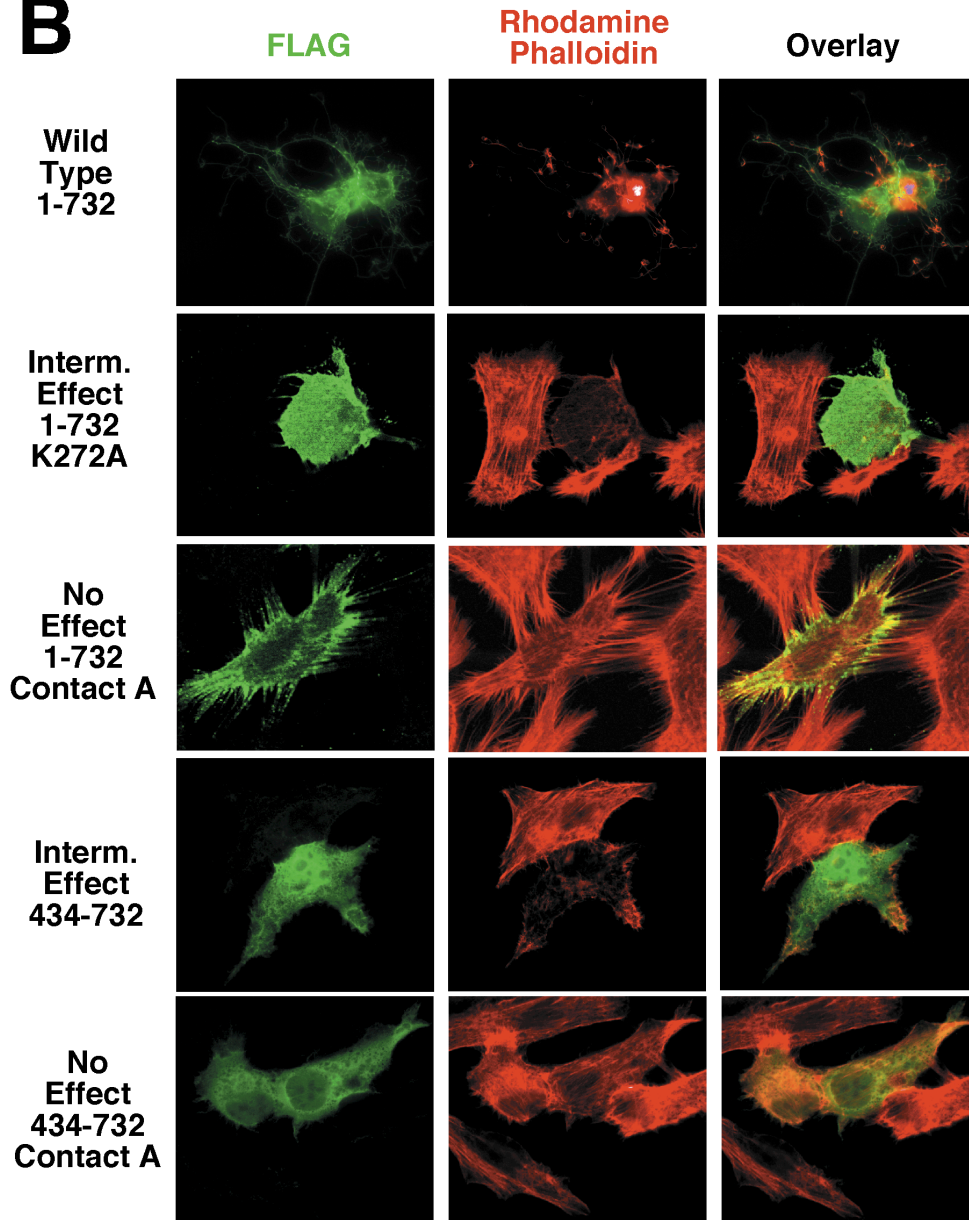
Figure 4.9: GDI Activity is Critical for YpkA to Promote Cytoskeletal Alterations

(A) The bar graph represents the percentage of transfected cells with cytoskeletal alterations and the contribution of the observed wild type phenotype and the observed intermediate phenotype to the total percentage. The wild type phenotype is indicated in grey and the intermediate phenotype in white. Bacterial Alkaline Phosphatase, BAP, is the negative control; 1-732 is wild type YpkA; 1-732 K272A is a kinase inactive mutant of YpkA; 1-732 Contact A is YpkA with the Contact A mutant; 1-732 K272A Contact A is the kinase inactive mutant plus the Contact A mutant; (434-732) is the crystallized construct; (434-732) Contact A is the Contact A mutant (RhoA/Rac1 binding deficient structure-based mutant).

(B) Examples of the visualization of YpkA by the immuno-staining of transfected Henle407 cells. Each of the three observed phenotypes is shown with staining for both the epitope, FLAG-tagged YpkA and for the actin cytoskeleton (Experimental Procedures). Each row is labeled by the construct transfected. The first column shows the presence of expression of the FLAG-tagged YpkA constructs and is colored in green. The second column shows the immunostaining of the actin cytoskeleton by rhodamine phalloidin in red, and the last column is the overlay of the fluorescent signals.

cytoskeletal alterations, they are observed to isolate to the membrane (Figure 4.9B, data not shown). As both the (434-732) and the (434-732) Contact A mutant are seen uniformly distributed throughout the cell, this supports other reported results that membrane localization is not dependent upon small GTPase binding and may be dependent upon a signal in the N-terminus of YpkA (Dukuzumuremyi, et al 2000) (Letzelter, Sorg et al. 2006).

Altogether, these results strongly suggest that the C-terminal, GDI-like activity of YpkA is the more significant contributor to cytoskeletal effects. The kinase domain does appear to work synergistically with the GDI-like domain, however, as evinced in both the reduced wild type effect of the kinase active site mutant, as well as the complete loss of activity in the YpkA (434-732) GDI (contact A) mutant. Interestingly, the double mutant of kinase and GDI does not completely abolish cytoskeletal effects, perhaps due to residual kinase activity present in the K272A mutation (Dukuzumuremyi et al., 2000, Juris et al., 2000).

Our data presented suggests an explanation of previous results obtained in the work by Juris et al (Juris, Rudolph et al. 2000). In their research, HeLa cells were transfected with vectors expressing wild type YpkA, a YpkA kinase inactivated mutant (K269A of *Yersinia entercolitica* YopO, equivalent to K272A of *Yersinia pseudotuberculosis* YpkA), and a C-terminal deletion construct that prevented kinase activity, reporting that, compared to the wild type phenotype, the kinase inactive mutant exhibited an intermediate phenotype where actin stress fiber formation was disrupted, but the actin microfilament system partially remained. This observation is consistent with our observations that YpkA (434-732) alone, as well as YpkA K272A full length, are

sufficient to disrupt stress fiber formation as well as to cause a low level of cellular deformation. Results with the C-terminal deletion constructs show that no cytoskeletal disruption activity is observed, although they contain the elements necessary for Rac1/RhoA binding. Our biochemical and structural results suggest that this may be explained by the aggregated and misfolded nature of these constructs (Figure 3.1), and not simply a lack of binding to actin. Taking into account our observations and the work by Juris *et al*, it is apparent that the full effect of YpkA function is achieved by both the kinase activity and the GDI activity of the C-terminal domain, although our point mutants in cell culture argue that the GDI-like activity is the greater contributor.

4.5 The YpkA GTPase Binding Domain is Essential for Virulence

In order to establish the relevance of these biochemical and cell biological data to infection, we examined the virulence phenotypes of *Y. pseudotuberculosis ypkA* mutants in a mouse infection assay. Two such mutants were tested, a *ypkA* null mutant, in which the entire *ypkA* reading frame was deleted, and a *ypkA* contact A mutant, in which the Tyr 591Ala, Asn595Ala, and Glu599Ala codon substitutions were introduced onto the virulence plasmid. Analysis by SDS-PAGE of the Yops secreted by these strains showed that the null mutant did not secrete YpkA protein, while the contact A mutant secreted a full-length polypeptide (Figure 4.10A). Immunoblot analysis of secreted Yops showed that YopT and YopE were secreted at native levels by the null mutant and contact A
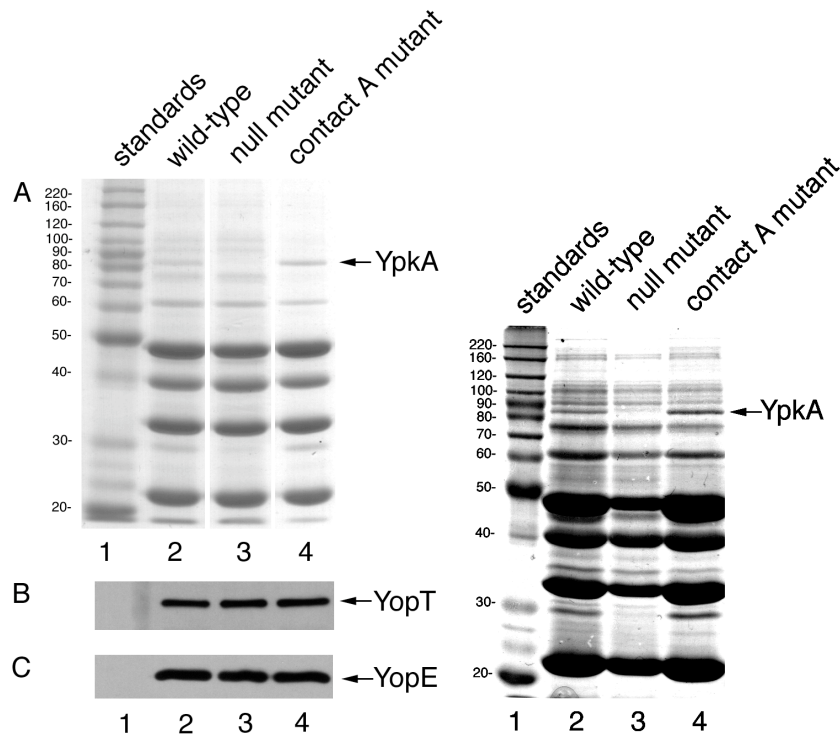
Figure 4.10: Analysis of Yops secreted by *Y. pseudotuberculosis* strains

Y. pseudotuberculosis wild-type, *ypkA* null mutant, and *ypkA* contact A mutant strains were grown in LB under conditions that stimulate Yop secretion (37°C and low Ca^{2+}). Aggregates (filaments) of Yops were collected from the cultures, washed once in LB, and solubilized by boiling in Laemmli sample buffer (100 mM Tris pH 6.8, 200 mM DTT, 4% SDS, 0.2% bromophenol blue, and 20% glycerol). Samples of the solubilized proteins were resolved on a 12% SDS-polyacrylamide gel, and detected by staining with GelCode Blue Stain Reagent (Pierce) (panel A), or by immunoblotting with antibodies specific for YopT (panel B) or YopE (panel C). The position of bands corresponding to the YpkA proteins secreted by the wild-type and contact A mutant strains is indicated by arrow on right of panel A. Panel A is a composite of lanes taken from a single gel that had similar levels of secreted YopT and YopE proteins, as shown by the similar signal intensities of the YopT and YopE bands in panels B and C. Sizes in kDa of molecular weight standards are indicated on the left. These experiments were performed by Maya I Ivanov and James B. Bliska.
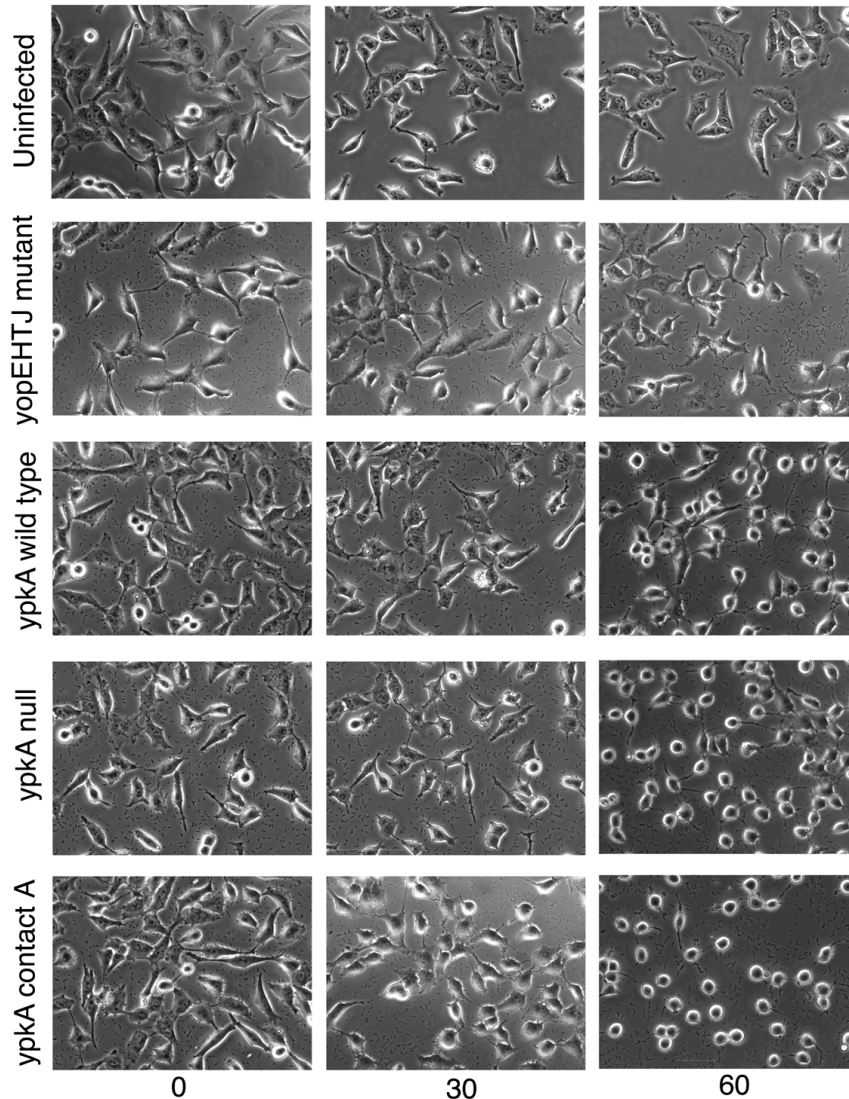
Figure 4.11: Analysis of HeLa cell rounding induced by infection with Y. pseudotuberculosis strains HeLa cells adherent to the bottom of 24-well tissue culture dishes were left uninfected or were infected at a multiplicity of 50 with the *Y. pseudotuberculosis* wild-type, *ypkA* null mutant, or *ypkA* contact A mutant strains. HeLa cells were also infected in parallel with an isogenic *yopEHTJ* mutant for comparative purposes. At the indicated times post infection (in minutes) the live cells were examined by phase contrast microscopy using a Zeiss Axiovert S100 equipped with a 32X objective. Representative images were captured using a digital microscope and a composite image was created using Adobe Photoshop. These experiments were performed by Maya I. Ivanov and James B. Bliska.

mutant (Figure 4.10B and C). In addition, the results of a HeLa cell rounding assay indicated that the null mutant and contact A mutant translocated YopT and YopE into host cells at normal levels (Figure 4.11). Mice were infected intragastrically with one of the mutants, or the isogenic parental strain, and the animals were monitored for survival over a 14-day period. As shown in Figure 4.12, all 8 mice infected with the wild-type strain succumbed to the infection by day 9, while only two mice infected with the contact A mutant died, one on day 7 and one on day 12. The survival curves for the mice infected with the wild-type or contact A mutant strains were significantly different (P=0.0003) as determined by a logrank test. Interestingly, all mice infected with the *ypkA* null mutant died by day eight (Figure 4.12), a survival rate not significantly different from the wild-type control (P=0.7333). Other groups studying pathogenesis of *Yersinia ypkA* null mutants in mouse infection assays have recently reported similar findings, in that strains lacking the *ypkA* gene appear to be as virulent as parental strains (Logsdon and Mecsas, 2003; Trulzsch et al., 2004). In contrast, *Y. pseudotuberculosis* strains that express the altered YpkA protein lacking the GDI activity are clearly attenuated for virulence (Figure 4.12). Taken together, these findings indicate that the GDI activity of YpkA is critical for *Yersinia* virulence.

In the initial work that reported the identification of YpkA, Galyov et al (Galyov et al. 1993) demonstrated that two different *Y. pseudotuberculosis ypkA* mutants were attenuated for virulence in a mouse infection model. One attenuated mutant resulted from an in-frame deletion of *ypkA* codons 207 to 388, which removed a major portion of the kinase domain and produced a protein that was secreted by the T3SS but lacked kinase activity (Galyov et al. 1993). The other mutant resulted from the integration of a
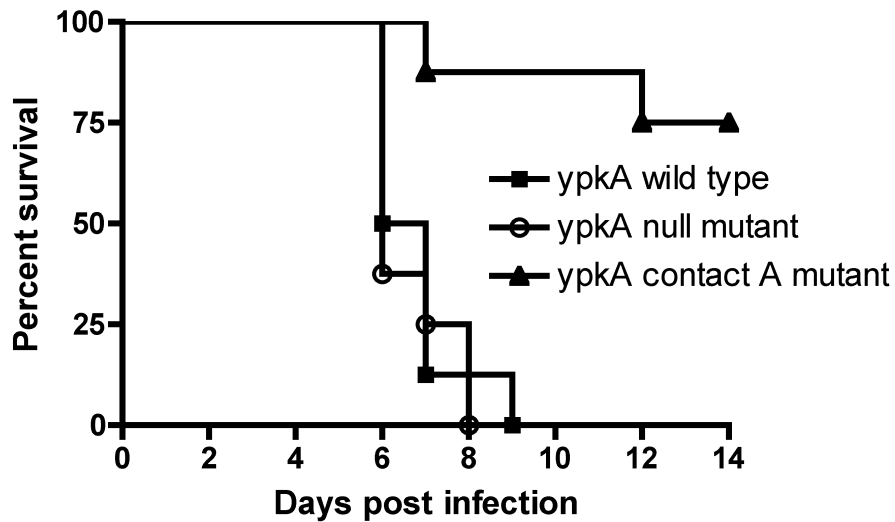
Figure 4.12: The GDI activity of YpkA is critical for *Yersinia* virulence
Groups of mice were infected intragastrically with 5×10^9 CFU of wild-type *Y. pseudotuberculosis,* a *ypkA* null mutant, or a *ypkA* contact A mutant. Survival of the mice was recorded over a 14 day period. Results shown are compiled from two independent experiments that were performed with groups of 4 mice. The survival curves for the mice infected with the wild-type or contact A mutant strains were significantly different (P=0.0003) as determined by a logrank test.

"suicide" plasmid into the *ypkA* gene, and produced a protein truncated after residue 548 that was also competent for secretion by the T3SS (Galyov et al. 1993). Taken at face value, these results suggested that both the kinase activity and a C-terminal region of YpkA were important for *Yersinia* virulence. However, it remained unclear how the C-terminal region of YpkA contributed to *Yersinia* virulence. Moreover, in recent studies, little or no role for YpkA in virulence could be found when *Yersinia* mutants containing a larger in frame deletion of *ypkA* (codons 21 to 712) (Logsdon and Mecsas, 2003), or a complete deletion of the *yopO* reading frame (Trulzsch et al., 2004), were utilized in

mouse infection assays. By constructing a *Y. pseudotuberculosis* point mutant specifically defective for GDI function, we now demonstrate that this activity within the C-terminal domain of YpkA is critical for *Yersinia* virulence. We further confirmed the findings that mutations that remove most, or all, of the *ypkA* reading frame do not lead to attenuation of virulence in a mouse infection model (Logsdon and Mecsas, 2003; Trulzsch et al., 2004). One possible explanation for these results is that null mutations in *ypkA* result in increased translocation of other Yop virulence factors *in vivo*, which compensates for the loss of YpkA function. If true, this type of phenomenon further underscores the power and importance of employing mutations that ablate the activity, and not the expression, of a suspected bacterial virulence determinant.

4.6 Collaborative Inactivation of Small GTPases by *Yersinia* Outer Proteins

Interestingly, despite the fact that YpkA mimics key aspects of GDI function, it does not possess all the activities of its host cell counterparts. Host cell RhoGDIs are also able to slow the intrinsic rate of GTP hydrolysis by small GTPases, and, more importantly, have a specialized domain which the protein uses to bind the isoprenylated tail of small GTPases to remove them from the cell membrane. Although YpkA can bind to both the GDP and GTP conformations of RhoA and Rac1 (Dukuzumuremyi, Rosqvist et al. 2000) and inhibit nucleotide release, it only has a moderate effect on the intrinsic rate of GTP hydrolysis (Figure 4.13) Moreover, YpkA does not possess any clear analog of the β-sheet motif necessary to bind an isoprenyl group. While it is possible that another region of the protein may indeed harbor such a motif, it seems more likely that
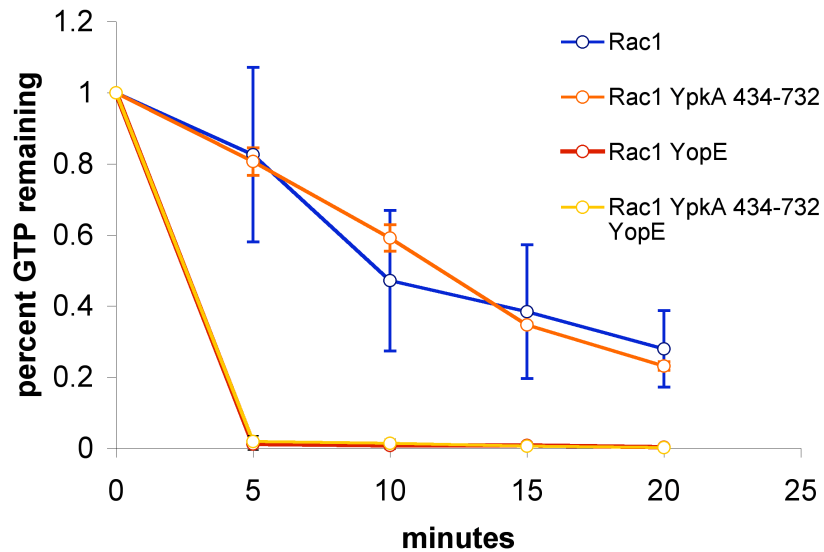
Figure 4.13: GTPase Activity Assay

The ability of YpkA to inhibit the intrinsic and YopE catalyzed GTPase activity of γ-^{32}P GTP Rac1 was measured as described in materials and methods.

YpkA has no such function. This may in fact be reasonable, as another *Yersinia* virulence factor that is translocated into host cells along with YpkA, YopT, removes any selective pressure for such a membrane removing GDI-like activity. This is because YopT, a cysteine protease that specifically cleaves off the C-terminal tail residues of the RhoGTPases Cdc42, Rac1, and RhoA, removes the small GTPases from the membrane in a highly efficient manner (Shao, Merritt et al. 2002). Moreover, *Yersinia* also translocates into host cells YopE, a GAP or GTPase activating protein essential for virulence, that quickly catalyzes the hydrolysis of GTP bound to small GTPases of the Rho-family (Black and Bliska 2000). Therefore, between YpkA and YopT, two of the physiological effects of a RhoGDI can be recapitulated. The final activity seen in many host RhoGDIs,
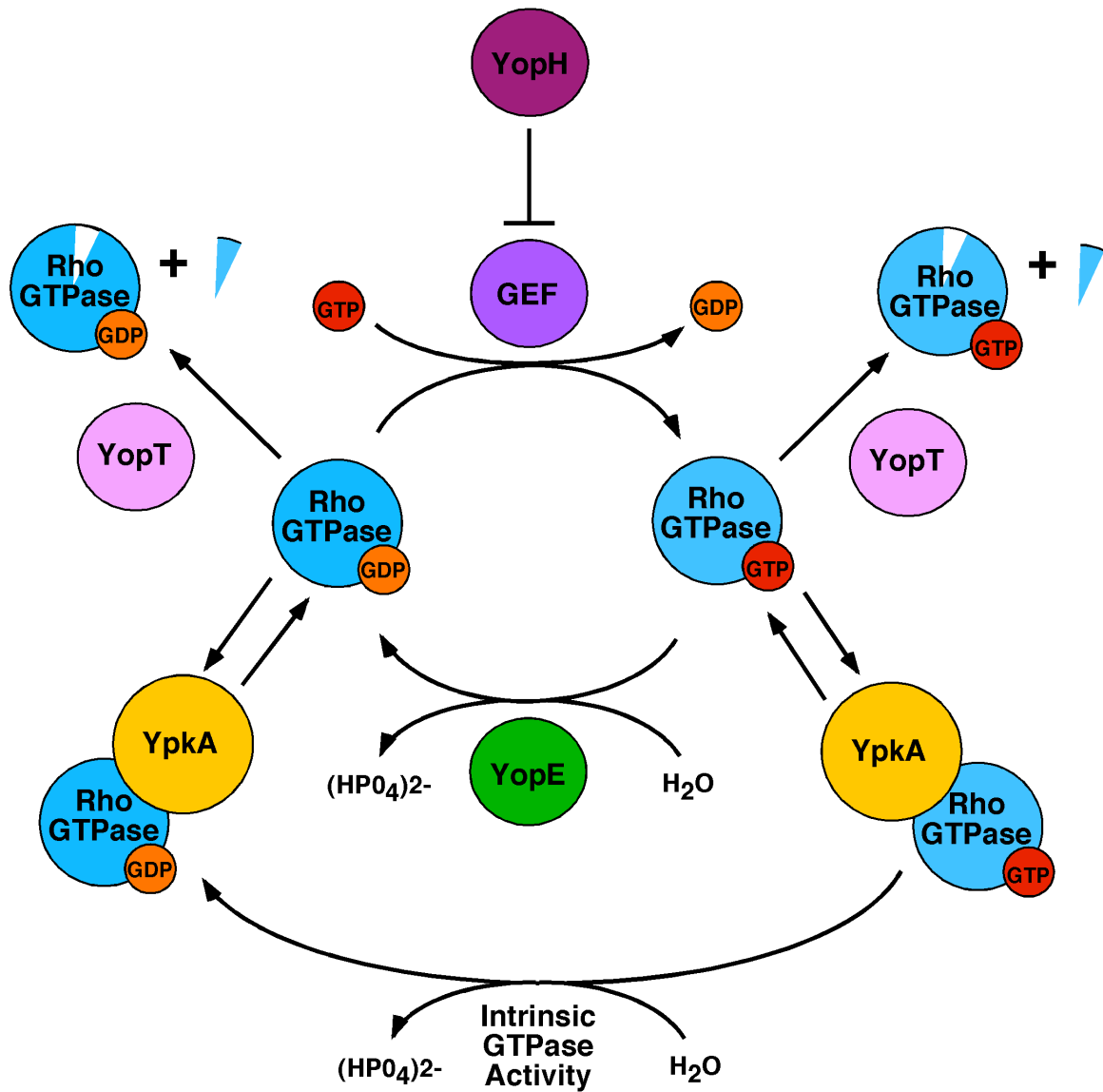
Figure 4.14: *Yersinia* **Outer Proteins work to inhibit Rho-family GTPase Signaling**
The proposed model of Yop action is described taking into account our results and the results summarized in the literature. YopH, YopE, and YpkA work in concert to shunt the entire small GTPase signaling pathway to the GDP bound, or "off" signaling state. YopT cleaves the C-terminal isoprenyl modification of the small GTPases so they can no longer associate with the cell membrane and bind to other regulators or effectors.

the inhibition of GAP activity on the RhoGTPases, would be counterproductive and interfere with the function of YopE. Thus, YpkA appears to be perfectly engineered to work in concert with YopE and YopT. A mechanism of the proposed Yop modulation of the small GTPases is shown in Figure 4.14.

It is becoming increasingly clear that a common strategy used by bacterial pathogens to modulate host cell biology is the mimicry of eukaryotic biochemical processes(Stebbins and Galan 2001). This has been especially true of virulence factors that target the Rho-GTPases in order to manipulate host cytoskeletal structure. We have presented data here that the *Yersinia* virulence factor YpkA, in addition to its host-like serine/threonine kinase activity, possesses an additional host mimicry by harboring key functions of the RhoGDI proteins, preventing nucleotide exchange in RhoA and Rac1 and thereby disrupting the host cytoskeleton. This fascinating interaction we show to exert a virulence effect, revealing another example of host mimicry in the virulence strategies of bacterial pathogens.

CHAPTER FIVE:

YERSINIA PROTEIN KINASE AND INHIBITOR DESIGN

5.1 Biochemical Characterization of the *Yersinia* Protein Kinase

YpkA is a serine/threonine kinase, capable of autophosphorylation as well as phosphorylating common serine/threonine kinase substrates *in vitro*, and kinase activation is dependent upon the presence of actin (Galyov, Hakansson et al. 1993; Juris, Rudolph et al. 2000). As illustrated in Figure 3.1, soluble constructs containing the putative ser/thr kinase domain could be isolated and purified. Minimal constructs were highly soluble and well purified, as well as a construct containing the kinase domain plus the GTPase binding and actin activation domains YpkA (115-732). Interestingly, the minimal kinase constructs were highly susceptible to specific degradation by an unknown protease in *E. coli* (Figure 3.1). Purified protein preps with GST-(115-465) yielded two smaller fragments of approximately 13kD and 30kD. N-terminal sequencing revealed that YpkA was cleaved at residue F353 which maps to a predicted loop region connecting the N and C lobes of the kinase model (Figure 5.2). To remove this degradation product, the minimal kinase constructs (115-465) and (115-428) were refolded removing the cleaved products as described in materials and methods. After purification the soluble constructs were tested for kinase activity as described in materials and methods (Figure 5.1).

As expected and in agreement with previous work, the minimal constructs were inactive and a mutation in the predicted kinase active site (K272A) was also inactive
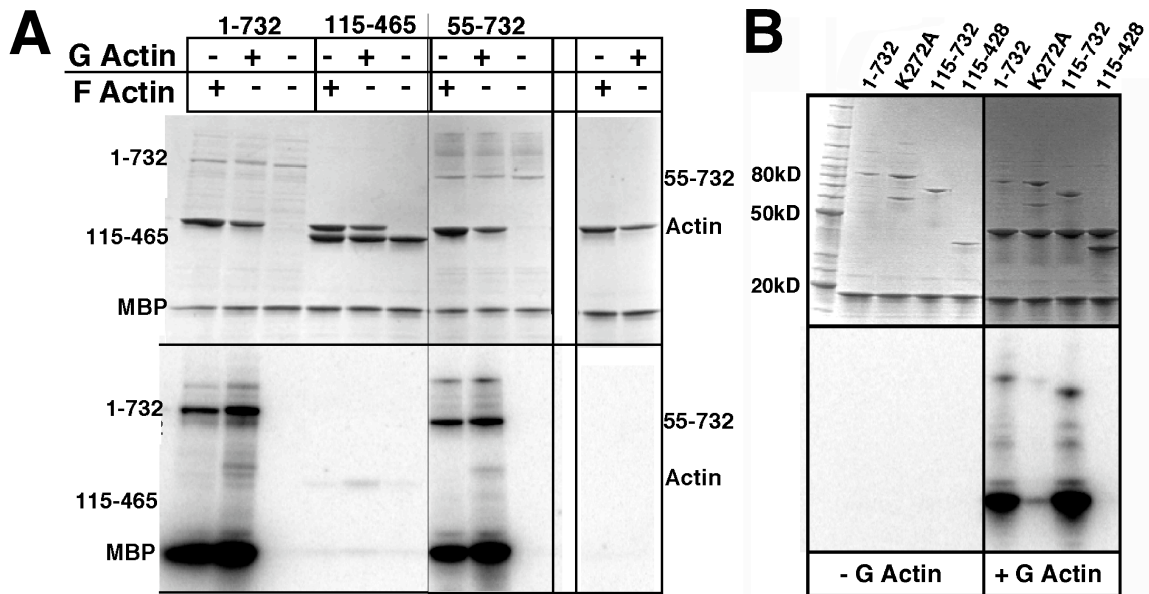
Figure 5.1*: Activity assay of purified YpkA constructs*

YpkA was incubated in G reaction or F reaction buffer with MBP and G or F actin and YpkA was assayed for autophosphorylation and phosphorylation of MBP with radio-labeled γ-^{32}P ATP (materials and methods). The upper gel shows the coomassie stain of the experiment whereas the bottom gel shows the autoradiograph of the assay. Positions of the various proteins are labeled on the left of the gel images. The presence or absence of G or F actin form is indicated with a + or – respectively.

relative to the wild type protein (Figure 5.1). What is interesting is both (55-732) and (115-732) are just as active as the wild-type YpkA, although the work by Trasak and colleagues shows that this should not be the case (Trasak, Zenner et al. 2007). In their work, they show that YpkA is only activated upon interaction with G-actin, but not F-actin, and that full kinase activity requires elements in the N-terminus of the protein, specifically the auto-phosphorylation of residues S90 and S95. In contrast, our results show that YpkA (115-732) is just as active as YpkA 1-732 and seems to also autophosphorylate. This implies that S90 and S95 are not needed for full activity, and that these are not the only sites of phosphorylation on YpkA. Additionally, our results

show that YpkA can also be activated by F-actin. The observed activity with F-actin can be reconciled with the results of Trasak et al, due to the fact that F and G actin are in a state of equilibrium. Most likely in our case, F actin was in equilibrium with a pool of G actin which was enough to activate YpkA. Although further studies are needed to explore the biochemical properties of YpkA, it is clear that the activity of YpkA seems to be regulated in some fashion by association with G actin and that the kinase domain by itself is not sufficient for kinase activity.

5.2 *Yersinia* Protein Kinase A Homology Model

Due to the importance of this virulence factor in disease caused by *Yersinia* spp., and the lack of any structural knowledge regarding the kinase domain of YpkA, a collaboration was set up with a post-doctoral associate, Xin Hu, where we constructed a homology model for this region. As the kinase activity of YpkA, has been shown to directly correlate to overall virulence, inhibition of YpkA could yield new anti-plague therapeutics (Wiley, Nordfeldth et al. 2006). A large number of eukaryotic homolog structures are known, which provided a means for evaluating the proposed model. The construction of this model by Xin Hu and its subsequent structural analysis (Section 5.2 and 5.3) was used for the development of potential YpkA inhibitors.

As the structure of YpkA is unavailable, we constructed the 3D model based on the template MAPK. YpkA shares about 20% homology to mammalian Ser/Thr kinases (Figure 5.2). If only considering the residues near the ATP binding site, the sequence
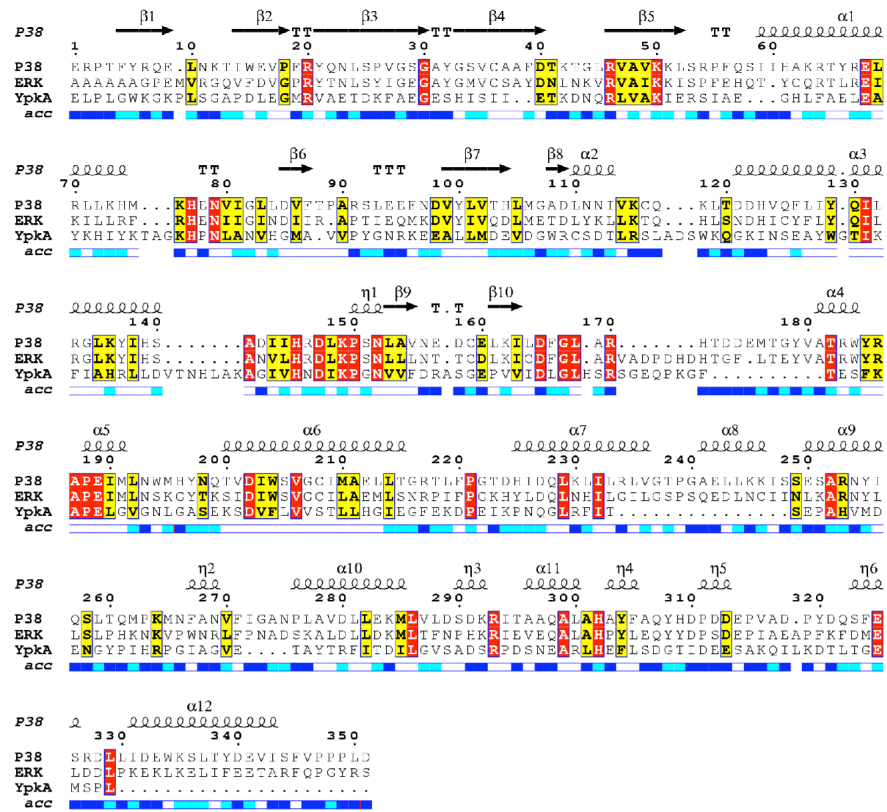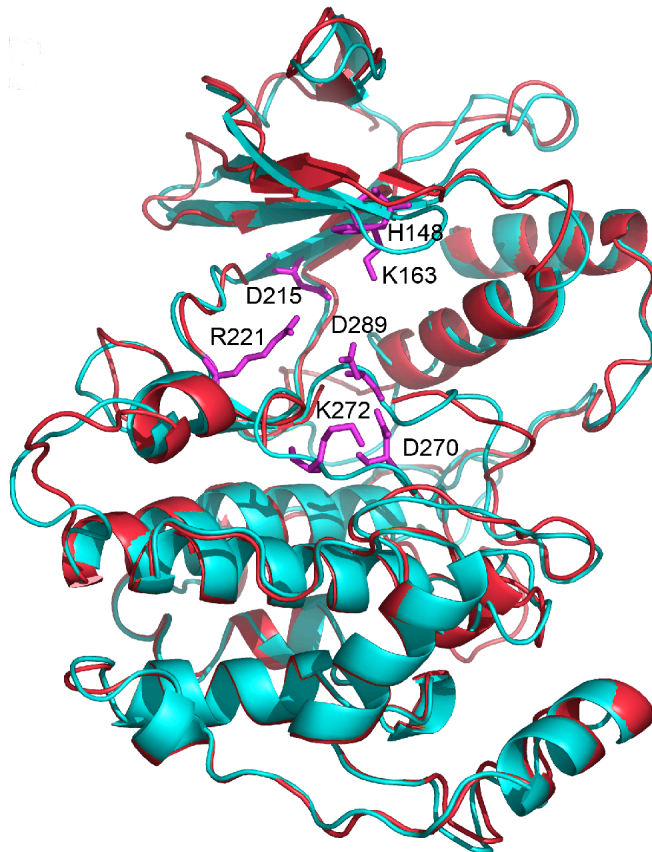
Figure 5.2: Homology Model of the YpkA Kinase Domain

(A) Sequence alignment of YpkA (115-431) with protein kinases P38 and ERK. Strict sequence conservation is shown in red background, and strong sequence conservation in yellow. The solvent-accessibility of each residues in the P38 structure is indicated in the bar at the base of the sequences, with white representing buried residues, dark blue representing solvent-accessible residues, and light blue representing an intermediate value. The secondary structural elements are also indicated according to the structure of P38.

(B) Structural alignment of the two homology models of YpkA kinase domain. Model A (red) represents a conformation of YpkA with an open ATP-binding pocket, while model B (cyan) has a closed ATP-binding pocket. The key residues to the ligand binding are shown in magenta.

identity to MAPK is 60%. Therefore, there is enough sequence similarity to build a reliable model of YpkA focusing on the catalytic site. We decided to construct two structural models based on different templates of MAPK. Model A used the apo structures of p38 (PDB id 1p38 and 1erk), while model B adopted ligand-bound complexes with induced fit at the ATP binding site (PDB id 1a9u and 3erk). An overall comparison of the ATP binding site between YpkA and MAPK reveals that most of the interactions that stabilize the ATP in YpkA are very similar to those observed in mammalian serine/threonine kinases. In the YpkA model, for example, the phosphate group of ATP is stabilized by at least two potential hydrogen bonds interacting with the conserved residues Lys163, and the backbone residues of glycine-rich loop. As shown in Figure 5.2, structural differences can be seen within these two models. Model A possesses a more open ATP binding pocket at the Glycine loop, while the catalytic site in model B is closed with the G-loop flipping down. YpkA possesses a DLG motif, rather than the conserved DFG in mammalian kinases at the beginning of activation loop, which is critical for function and inhibitor binding. Specifically, recent studies shown that the conformational change of "Phe-in" and "Phe-out" in the DFG motif plays an important role in substrate interaction (Mol, Fabbro et al. 2004). The substitution of leucine for phenylalanine at this site in YpkA would likely create important differences in inhibitor specificity. As the conformational changes of the G-loop are very sensitive to ligand perturbation, both are valid conformations for our structural analysis and inhibitor design.

To further examine the structural features of YpkA, we performed molecular dynamic simulations of the two YpkA structures in the apo and ATP bound forms. The
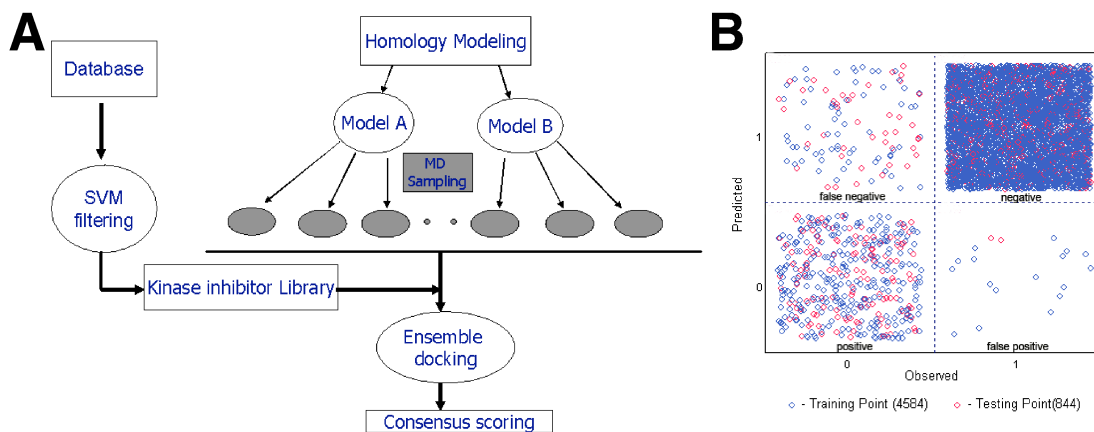
Figure 5.3: Virtual Screening Strategy and SVM Model

(A) Machine-leaning SVM model derived from kinase inhibitors using ADMET/Predictor (SimulationsPlus Lancaster, CA).

(B) Machine-leaning SVM model derived from kinase inhibitors using ADMET/Predictor (SimulationsPlus Lancaster, CA). This data was generated by Xin Hu.

simulations were carried out *in vacuo* permitting more extended conformational changes of modeled systems. Analysis of the dynamics of the protein at different states revealed that a number of active site residues that exhibited high flexibility (Figure 5.2). In order to sample a good representation of protein conformations for the subsequent ensemble docking, 500 conformers were extracted from 2.0 ns MD simulations and clustered according to a defined residue center at the active site. Five major clusters were obtained with model A and three clusters were found with model B. From the MD simulations and the docking studies we believe that the conformational changes of these key active site residues represent to some extent the plasticity of the ATP binding site upon ligand binding, and are thus crucial for consideration in inhibitor design.

96

5.3 Database Screening for *Yersinia* Protein Kinase A Inhibitors

Protein kinase inhibitor design remains a challenging problem because of the high similarity and plasticity of the catalytic site of the ATP (Cavasotto and Abagyan 2004; Muegge and Enyedy 2004; Noble, Endicott et al. 2004; Scapin 2006). In this study, we applied an approach combining machine learning method and multiple conformational high throughput docking for the discovery of YpkA inhibitors. The screening strategy employed was illustrated in Figure 5.3A. First, we developed a machine learning SVM model using a data set of known kinase inhibitors from a diverse kinase collection. The ligand-based SVM model was used as a kinase filter to prioritize the large size of chemical databases and a target-focused library was obtained. Second, we constructed homology models of YpkA based on the MAPK templates, and further performed MD simulations to sample different protein conformations characterized in the catalytic site to account for protein flexibility. Finally, with an ensemble of protein structures and the kinase inhibitor-enriched library, multiple conformational high throughput docking was performed and a number of potent and selective inhibitors of YpkA have been successfully identified.

In order to develop a general kinase model for large-scale database filtering, 364 kinase inhibitors were selected from a diverse kinase collection. These active compounds were seeded into a data set of a non-kinase chemical library comprising 4220 inactive compounds serving as the training set. Molecular descriptors were calculated with ADMET/Predictor consisting of 276 descriptors from the 3D structure. The use of ADMET molecular descriptors was anticipated to improve the drug-likeness property of

identified compounds, which is a crucial aspect in the late stage of drug development. The SVM model was derived from the molecular descriptors of the training set in distinguishing the active and non-active compounds. As shown in Figure 5.3B, 319 out of 364 inhibitors were classified in the "positive" region, while only 15% active compounds were mis-classified as false negative. To validate the model, a testing data set comprising 175 known kinase inhibitors and 669 inactive compounds was applied using the SVM model. 127 out of 175 active compounds were predicted correctly, yielding an enrichment of 70%. This result (Ford, Pitt et al. 2004; Briem and Gunther 2005)is quite promising, comparable to many other machine-learning models published recently . Given the high efficiency of the SVM model, we then screened our in-house database collections consisting of more than 2 million compounds, and a kinase-focused library of ~200,000 compounds was obtained.

With the structural model of YpkA and the SVM-enriched kinase inhibitor library, we then performed a multiple conformational high throughput docking for the search of potent and selective inhibitors for YpkA. The program FlexE was used, which is designed to accommodate multiple conformations of protein in docking by forming new structural representatives (Claussen, Buning et al. 2001). A total of eight conformers of YpkA sampled from MD simulations were used in FlexE docking. The focused library consisting of ~200,000 compounds were subsequently docked to the ensemble of protein structures and ranked according to the FlexX score. The top-ranked 1000 compounds bound in the active site were visually inspected. To improve the hit selection, we also applied consensus scoring on the FlexX-docked complexes. The top 5% compounds were extracted and re-ranked using X-Score, which proved to be a reliable consensus scoring
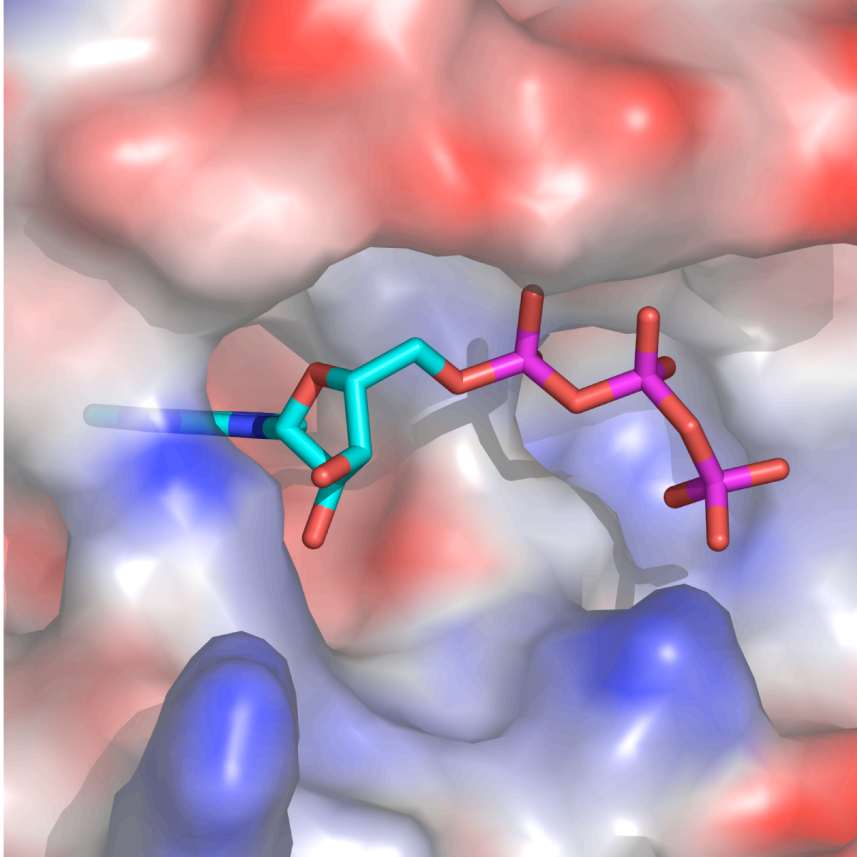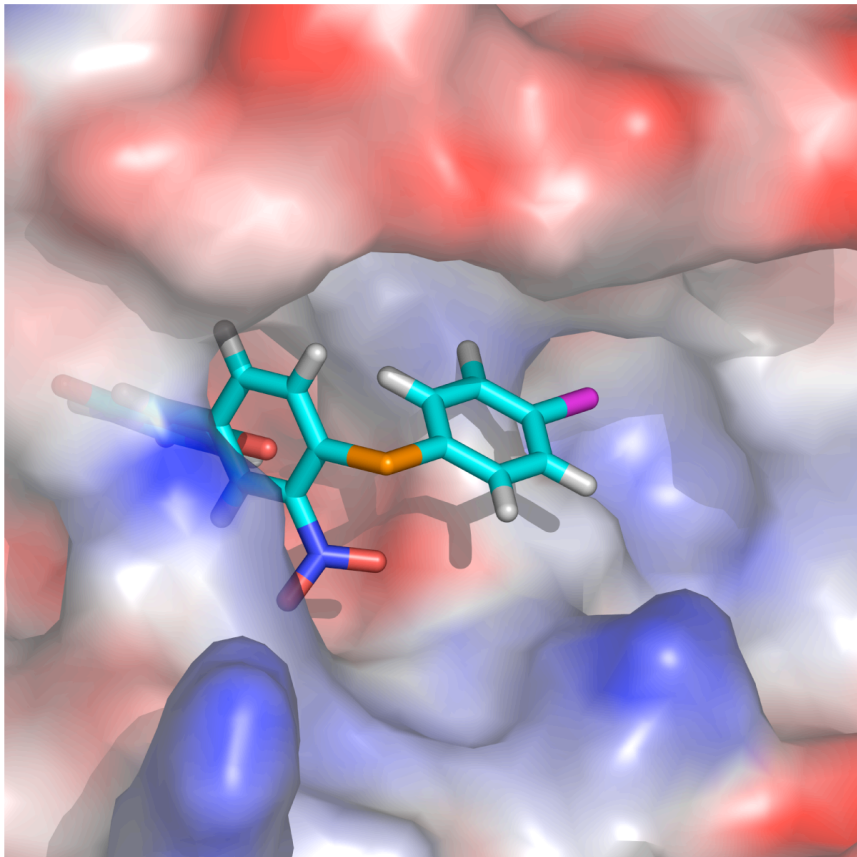
Figure 5.4: Comparison of ATP and Inhibitor Binding

(A) The fit of ATP in the active site of the YpkA homology model.

(B) The fit of compound four in the active site of the YpkA homology model.

ATP was placed by MODELLER and each compound was placed in the active site by virtual screening experiments (see materials and methods). The active site conformation in each case is the homology model B, or closed ATP pocket. The surface of the YpkA kinase domain was calculated and the charges on the surface and in the active site pocket are colored by charge. Red indicates areas of positive charge, or acidic areas, whereas blue represents areas of negative charge or basic areas.

function and has been successfully used in many applications for drug discovery (Wang, Lai et al. 2002). The top 1000 compounds were also visually inspected in terms of overall fit, key interactions in the binding site, as well as the structural complexity and diversity of compounds. An example of the final fit of one of the inhibitors is shown in Figure 5.4.

5.4 *Yersinia* **Protein Kinase A Inhibitors**

A total of 45 compounds were finally selected to experimentally test the inhibitory activity against YpkA. Initial screening was done in duplicate and assayed by blotting onto a nitrocellulose filter at high inhibitor concentrations. Seven of the 45 initial compounds showed complete inhibition at the high screening concentrations of 225μM to 450μM, yielding a hit rate of 15%. The IC_{50} values of these compounds were determined by radiological assay with three compounds exhibiting inhibitory activities below 10μM at 1.81μM, 5.87μM, and 9.72μM, and the remaining four having IC_{50} values below 50 μM (Figure 5.5). Those compounds shown represent distinct families of compounds, three of which were redundant and thus omitted. Examination of these active compounds revealed a diversity of chemical structure, as represented in Figure 5.6. Compound 1 possesses a scaffold of indolin-2, which is commonly found in the derivatives of CDK2 inhibitors. Compound 2 belongs to the class of anthiraquinone, the potent inhibitor of casein kinase-2. The structures of compounds 3 and 4 are quite interesting, as they possess a novel functional group of pyrimidine-2,4,6-trione. Analysis of these compounds bound in the catalytic site of the ATP suggested that the pyrimidine
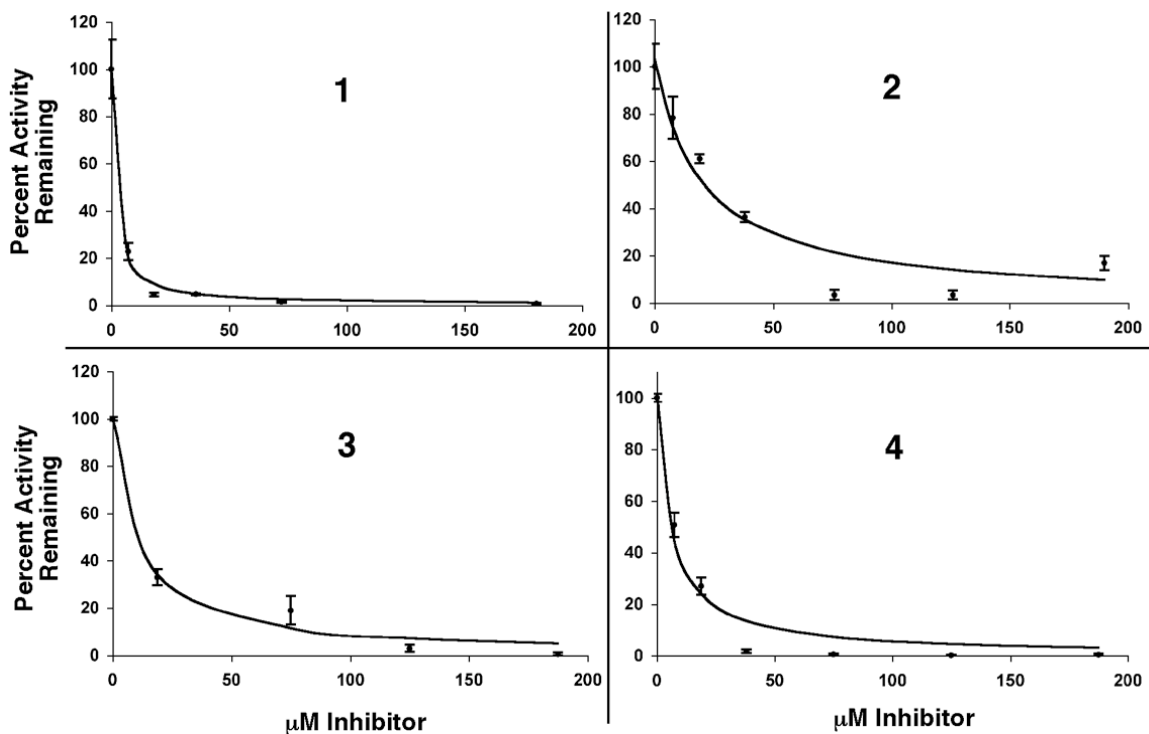
Figure 5.5: Inhibitor data of YpkA and the four top inhibitors

The activity of each compound was assayed as percent total phosphorylation remaining versus inhibitor concentration. The data was generated and the curve fit was performed as described in experimental methods. Each curve is labeled with its respective compound from Figure 5.6.

derivative resemble the adenosine moiety of the cofactor, involving in H-bonding interactions with hinge residue Asp218.

We further evaluated the selectivity of these identified YpkA inhibitors by testing against other two kinases, MAPK and protein kinase C (PKC). It is not surprising that some compounds showed comparable inhibitory activities to MAPK, from which the homology models of YpkA were derived. For example, compound 1 showed the best inhibition to YpkA with IC_{50} of 1.81 μM, and also exhibited similar activity to MAPK with IC_{50} of 2.45 μM. However, compounds 2, 3, 4 are highly selective to YpkA over
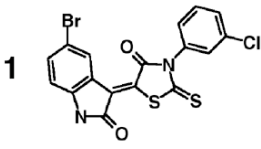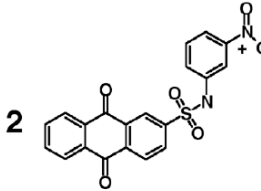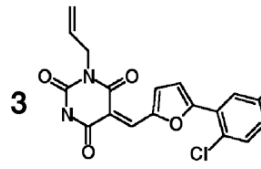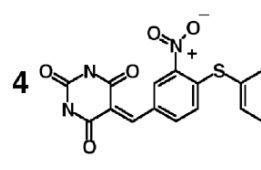
Compound	IC$_{50}$ (μM)		
	YpkA	MAPK	PKC
1	1.81 ± 0.33	2.45 ± 0.53	> 200
2	19.84 ± 4.32	> 200	165 ± 61
3	9.72 ± 1.66	54.46 ± 19.4	205.16 ± 47.65
4	5.87 ± 0.81	63 ± 21	> 200

Figure 5.6: YpkA Inhibitor IC$_{50}$ Values

The chemical structure of each compound is listed along with its IC$_{50}$ value (at YpkA concentrations of 0.15μM).

MAPK and PKC with 5 to 10 fold better inhibition (Figure 5.6). The discovery of *Yersinia* protein kinase A inhibitors by the use of a combination of both ligand-based and structure-based knowledge of protein kinases has demonstrated high screening efficiency and reasonable speed. This integrated approach therefore provides a practical method to account for protein flexibility in a large-scale database for virtual screening of effective inhibitors of therapeutic targets. Although further studies are required to validate and characterize the inhibition of YpkA by these compounds, the discovery of these potential YpkA inhibitors provides a staring point for the design of more potent and selective inhibitors as anti-plague drugs.

CHAPTER SIX:

A RAC1-GDP-ZINC COMPLEX

6.1 Overall Structure of a Rac1 GDP Zn Complex

The initial crystal screening was performed with the YpkA-Rac1 complex and produced a condition containing a high concentration of zinc (15mM Zn SO_4), which caused intense precipitation followed by subsequent crystal formation (Figure 6.1). Due to the inability to completely separate the aggregated material from the crystalline material the exact protein species within the crystal could not be determined. A data set was taken on a spinning copper anode source with a Rigaku IV++ detector, and processed to a final resolution of 2.2Å (data not shown). To generate initial phases and to reveal with protein species were present, molecular replacement using the program Phaser was employed. Molecular replacement searches used both a Rac1 search model (1MH1 pdb accession code) and the YpkA structure. The final solution structure contained three Rac1 molecules arranged as a trimer with an R/Rfree of 19.2/23.6. Based on the crystallization conditions and visual inspection, four sites of zinc coordination were initially found and magnesium was modeled into its predicted site at Switch I. To verify the presence of zinc within the crystal, another and primary data set was taken at the zinc K-absorption edge (see materials and methods) and the zinc sites found by single anomalous dispersion using the program SOLVE to analyze the anomalous differences. The initial molecular replacement solution was refined against the new data set with the experimental zinc sites placed (Table III).
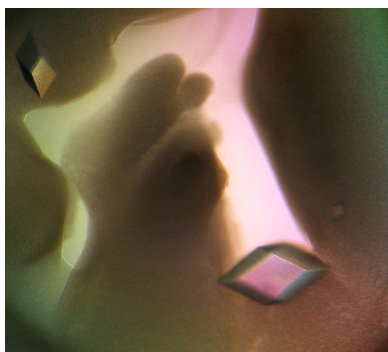
Figure 6.1: Rac1 GDP Zinc Complex Crystals

Crystals grow by hanging drop were imaged under plane polarized light at 100x magnification.

Table III Rac1 data collection and refinement statistics

Data collection	
Space group	$P3_221$
Cell dimensions	
a, b, c (Å)	89.7, 89.7, 191.6
α, β, γ (°)	90.0, 90.0, 120.0
Resolution (Å)	77.62-1.90
No. reflections	548625
No. Unique reflections	70803
R_{sym} or R_{merge}	10.0 (51.2)
$I / \sigma I$	18.2 (3.7)
Completeness (%)	99.9 (100.0)
Redundancy	7.7
Refinement	
Resolution (Å)	77.62-1.90
No. reflections	67009
R_{work} / R_{free}	17.4 / 20.8
No. atoms	
All atoms	5129
Protein	4141
Water	896
GDP	84
Zinc	8
B-factors	
All atoms	27.6
Protein	33.1
Water	42.6
GDP	20.7
Zinc	31.3
R.m.s deviations	
Bond lengths (Å)	0.014
Bond angles (°)	1.638
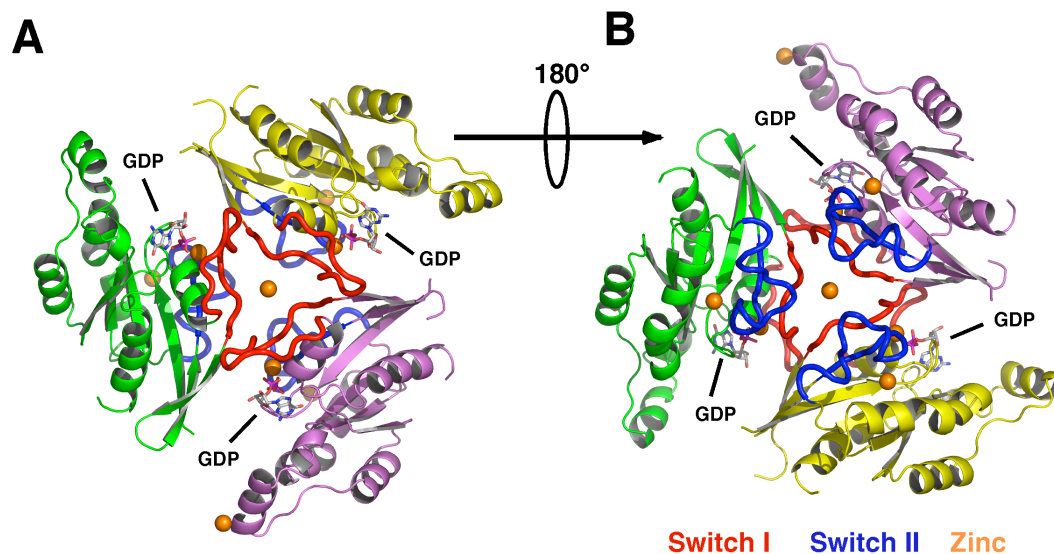
Figure 6.2: The Overall Crystal Structure of the Rac1 GDP Zinc Complex

(A) The trimeric Rac1 complex is shown with chain A in violet, Chain B in yellow, and Chain C in green. Switch I is indicated in red, Switch II in blue and zinc in dark orange. The three GDP molecules are labeled.

(B) The trimeric Rac1 complex after a 180° rotation from panel A.

The Rac1 GDP complex consists of three Rac1 molecules in the asymmetric unit related to each other by a three-fold non-crystallographic axis of symmetry. Each monomer not only contains the expected GDP, but also contains two zinc ions (Figure 6.2). At the center of the non-crystallographic three-fold axis there is an additional zinc that is coordinated by residues from Switch I (Figure 6.3A), and a final zinc that links Chain A and Chain C from symmetry related molecules (Figure 6.2 and Figure 6.3D). The Rac1 molecules in the asymmetric unit are almost identical, having low root mean square deviations as calculated from the Cα trace. The highest RMSD calculated using the Cα trace was between chain A and chain B, having a value of 0.42Å. Although

highly identical, there are small differences between the Rac1 monomers in areas of high disorder. None of the Rac1 monomers contain density for the N-terminal cloning artifact present in the protein sequence from the GST affinity tag linker (see materials and methods) and all three chains have poor density from residues 45 to 50. In this loop region connecting β-sheets two and three, Chain A is missing density for residues 45, 46 and 47, and Chain B is missing density for residues 49. Additionally, both Chain B and Chain C have poor side chain density in an α-helical region from residues 121-130 (Chain B) and residues 121-124 (Chain C). Finally, no density is seen in any of the monomers C-terminal to residue 179.

6.2 Zinc Coordination by Rac1

In the original molecular replacement structure to 2.2Å, four zinc sites were placed due to the crystallization conditions and by homology to known zinc coordinating protein structures. To verify these assignments experimentally another data set at the zinc absorption edge was taken to find the heavy atom sites using the anomalous differences. As shown in Figure 6.2 and Table III, the anomalous signal in the data set revealed that the crystal in fact contained eight zinc sites. Four of the zinc sites were those previously assigned (Figure 6.3A, and 6.3C), but the data also revealed four additional anomalous scatterers (Figure 6.3B and 6.3D). Although Rac1 has not previously been shown to bind zinc, it is clear that there are zinc atoms present in the crystal structure based on the crystallization conditions, the anomalous data, calculated electron density maps, and comparison to known crystal structures containing zinc. As
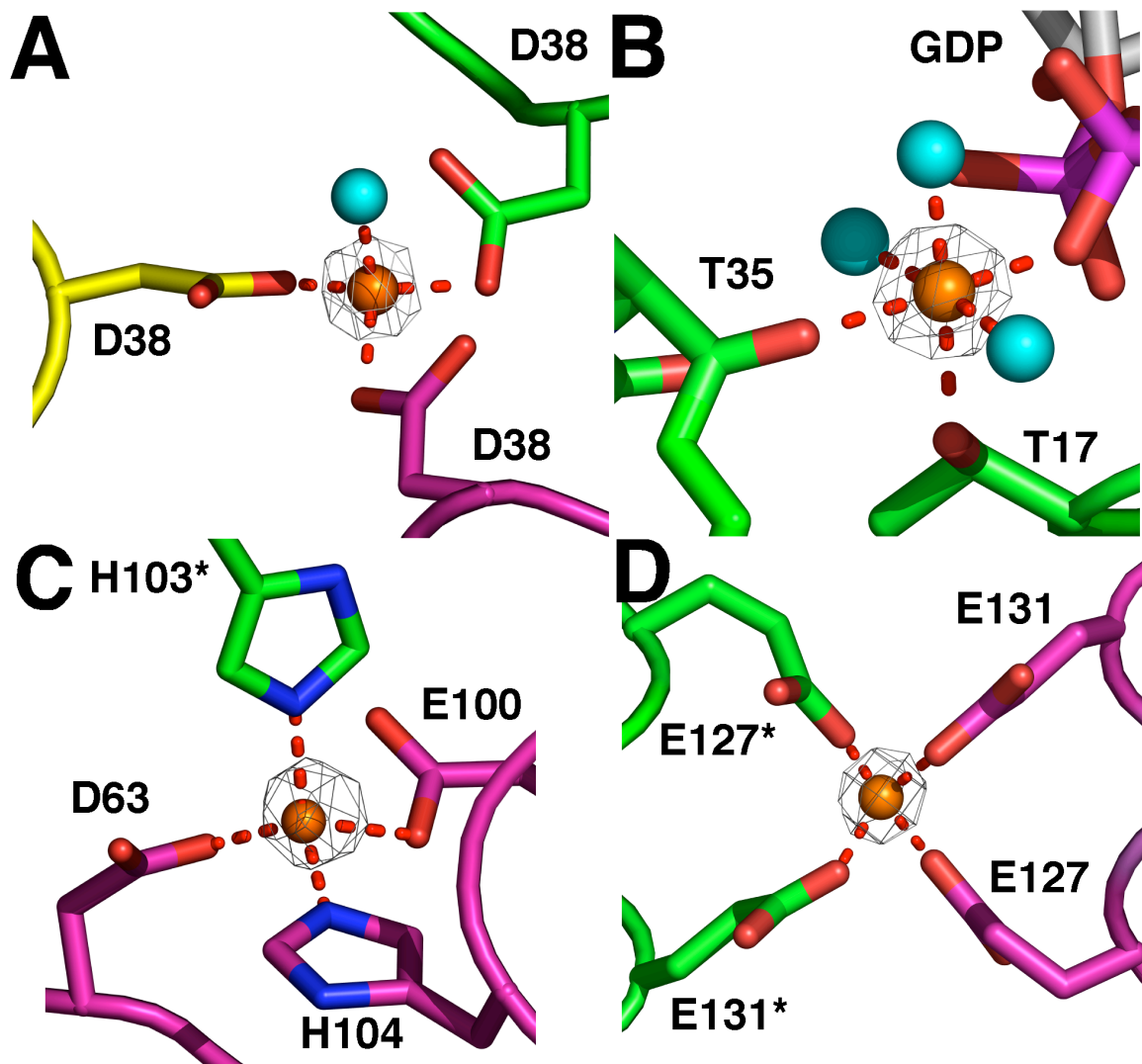
Figure 6.3: Rac1 zinc Coordination Sites

(A) The coordination of zinc by Switch I. Chain A is shown in magenta, chain B in yellow, and chain D in green. Zinc is shown in dark orange and the coordinated water molecule in light blue. The dashed red lines indicate the zinc coordination interactions. The Fo-Fc map is show in grey at 20σ and calculated without zinc.

(B) Zinc replaces magnesium. The octahedral coordination of zinc at the normal magnesium binding site in Chain C shown. Zinc is shown in dark orange and the coordinated water molecules in light blue. The dashed red lines indicate the zinc coordination interactions. The Fo-Fc map is show in grey at 20σ and calculated without zinc.

(C) Coordination of Zinc by Switch II. Each residue participating in the coordination of zinc (dark orange) is drawn from both Chain A (magenta) and a symmetry related Chain A (green). The coordination of Zinc is indicated by dashed red lines. The Fo-Fc map is drawn in grey at a contour level of 20σ and calculated without the zinc atom. The asterisk indicates a residue donated from a Rac1 in a crystallographically related asymmetric unit.

(D) Crystal Packing interaction involving Zinc. Residues from Chain A are shown in magenta and residues from a symmetry related Chain C are shown in green. The coordination of Zinc is indicated by dashed red lines. The Fo-Fc map is drawn in grey at a contour level of 9σ and calculated without the zinc atom. The asterisk indicates a residue donated from a Rac1 in a crystallographically related asymmetric unit.

stated in materials and methods, high concentrations of $ZnSO_4$ was used in the crystallization conditions, and could not be removed or replaced with other salts such as $MgSO_4$. Additionally, the Fo-Fc density is drawn as grey wire in Figure 6.3 for each assigned site, showing the clear presence of large difference density. The Fo-Fc map was generated without the zinc atom and is drawn in blue at a contour level of twenty sigma for panels 6.3A-3C, and at nine sigma for panel 6.3D. The modeling of zinc in the Rac1 structure caused an overall drop in both the Rwork and Rfree by 3.1% and 3.5% respectively. Similar effects on the refined R-values have been observed previously for macromolecular structures containing zinc (Papageorgiou, Acharya et al. 1995)

As expected and demonstrated by our data, most of the zinc sites are tetrahedrally coordinated, the most common mode of zinc coordination in proteins (Alberts, Nadassy et al. 1998; Dudev and Lim 2003). The central zinc (Figure 6.2 and Figure 6.3A) is coordinated by the carboxyl group of an aspartic acid residue from each monomer (asp38), which is completed by a water molecule to generate a tetrahedral geometry. In Figure 6.3C and 6.3D, a similar configuration is observed, with all donating atoms belonging to side chain residues. In contrast, Figure 3B shows that zinc is coordinated in an octahedral geometry by T35, T17, GDP and three water molecules. This was in contrast to the original model as magnesium was assigned to this site. Normally Rac1 uses magnesium to coordinate and bind GDP, and that this coordination contributes significantly to the binding energy of the nucleotide and overall stability(Vetter and Wittinghofer 2001; Hakoshima, Shimizu et al. 2003). Based on these considerations it was at first surprising to see magnesium displaced, but the GDP molecule still binding tightly. Although initially unexpected, zinc is known to adopt octahedral coordination

and, more importantly, zinc has been shown to displace magnesium in both protein and RNA structures, as these atoms have similar coordination properties and ionic radii (Dudev 2001; Ennifar, Walter et al. 2001; Dudev and Lim 2003). Additionally, the work of Dudev and colleagues clearly demonstrates that zinc has a higher affinity than magnesium for the octahedral coordination sites that magnesium adopts in protein structures (Dudev 2001). Due to the high concentration of zinc in the crystallization buffer (methods) and the contribution of the Rac1 P-loop to nucleotide binding (Hutchinson and Eccleston 2000), zinc which is hexahydrated normally in solution most likely replaced magnesium before the GDP could be released. Moreover, these zinc sites represented the largest peaks in the anomalous signal. This in conjunction with our data demonstrates that zinc can fill the role of magnesium in GDP coordination of Rac1 molecules.

6.3 Intermolecular Interactions at Switch I stabilize the Rac1 Trimer

Rac1 bound to GDP behaves as a monomer in solution as assayed by gel filtration (Stebbins and Galan 2000; Prehna, Ivanov et al. 2006), making the Rac1-GDP-Zn complex crystal structure unexpected. Although, there are several packing interactions that stabilize the crystal, the asymmetric unit has extensive intermolecular interactions that promote the trimer structure. Most of the intersubunit contacts within the asymmetric unit are isolated to Switch I and occur at residues 31 to 35, 39 to 42, and residue 37 (Figure 6.4). As described in Figure 6.4, at each apex of the triangle like structure formed by Switch I due to the tetrahedral coordinaton of zinc (Figure 6.2), the

111
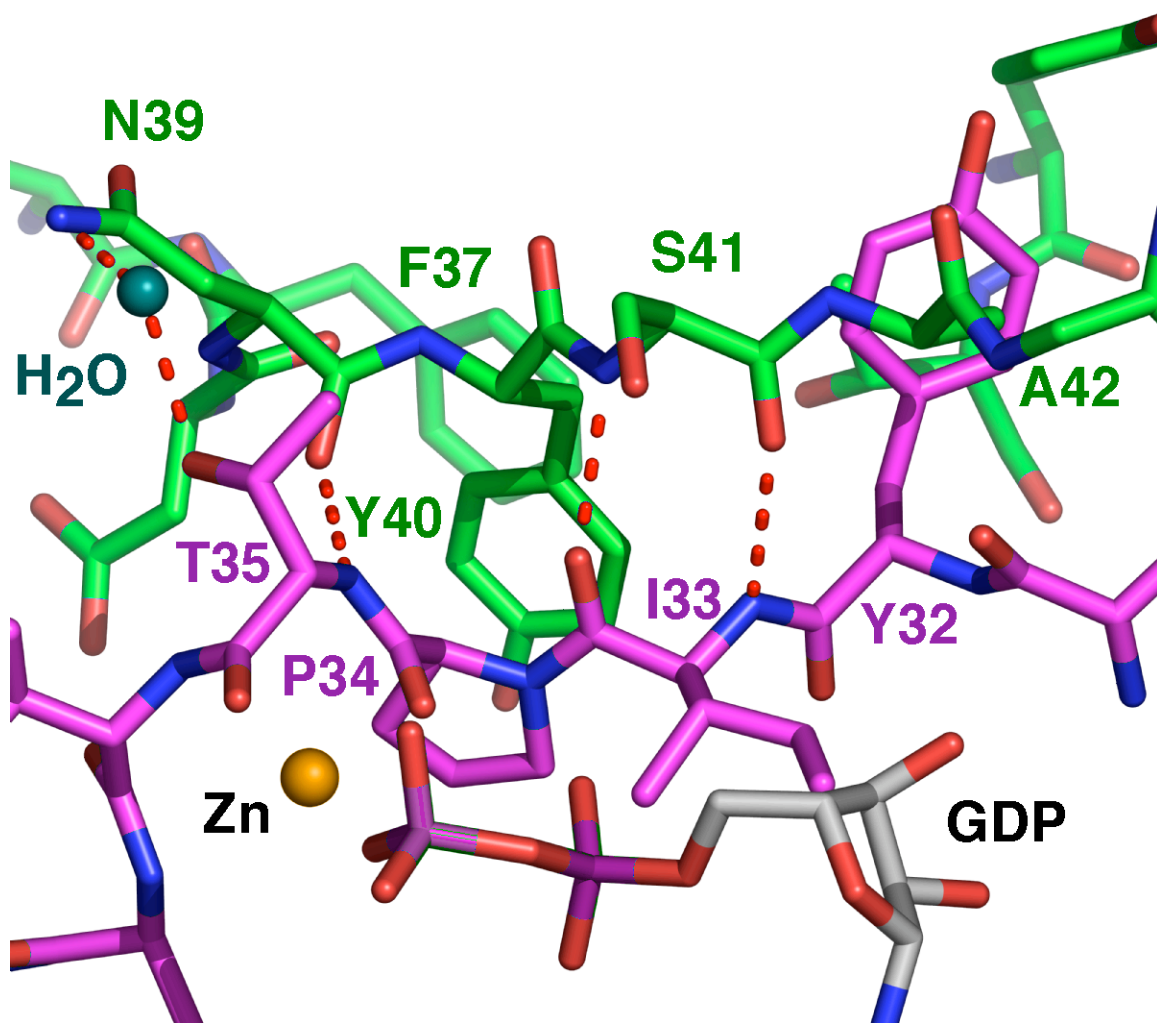
Figure 6.4: Intermolecular Interactions at Switch I

The intermolecular interactions at Switch I as viewed between chain A (violet) and chain C (green). All residues from chain A are labeled in violet and all residues from chain C are labeled in green. Water molecules are labeled in dark green and zinc in orange. All non-carbon atoms in the protein chains are colored as red indicating oxygen and blue indicating nitrogen. Hydrogen bonding interactions are indicated by dotted red lines.

switch region from each Rac1 monomer forms an intermolecular β-sheet with the subsequent monomer in the trimer. This is a continuation of the β-sheet formed between β-sheets 1, 2, and 3 in each Rac1 monomer. Overall, the intersubunit interactions provide a contact area of approximately $900Å^2$ between each monomer for a total of $2699Å^2$. An example of this interaction is shown in Figure 6.4 using chain A in violet and chain C in green. Main chain hydrogen bonds are formed between the nitrogen of I33 of chain A and the carbonyl oxygen of S41 in chain C, the carbonyl oxygen of I33 in chain A and the nitrogen of S41 in chain C, the nitrogen of T35 in chain A and the carbonyl oxygen of N39 in chain C. One side chain polar interaction stabilizes this interface and consists of a water mediated hydrogen bond that is formed between the hydroxyl of T35 in chain A, a water molecule, and the nitrogen atom of the carbonyl group of N39 in chain C. Additionally, other residues near Switch I complete the binding surface. Y32 from chain A makes several main-chain and side chain van der waal interactions with the chain A residues Y23, A42, and L55. Residue F37 of chain C makes intramolecular van der waal contacts with Y40 to position it to make a hydrophobic contact with the main chain atoms of I33 and the side chain of P34, both from chain A.

In addition to the protein-protein interactions that stabilize the Rac1 trimer, the trimer itself seems to be formed and stabilized by the coordination of a zinc atom by the Switch I regions of each of the Rac1 monomers (Figure 6.3A). Specifically, as illustrated in Figure 6.3A, and described previously. D38 is involved in the tetrahedral coordination of a zinc atom. It seems apparent from this configuration that the resulting Rac1 trimer is a result of the tetrahedral coordination of the zinc atom by the Switch I regions of each monomer.

6.4 The Coordination of Zinc Stabilizes the Rac1-GDP-Zn Crystal Structure

The Rac1-GDP-Zn complex crystal is stabilized by symmetry related packing interactions at three surfaces, two of which involve zinc coordination. One crystallographic contact does not contain zinc but involves interactions at the C-terminal helix and consists of residues 117, 121, 127, 131, 138 to 140, 156, 163, 167, and 170 of each Rac1 molecule. These surfaces mainly constitute a hydrophobic van der waal packing surface with the same residues in a symmetry related molecule of Rac1. In contrast, the other two intermolecular interactions are primarily mediated through zinc coordination. The first is shown in Figure 6.3D and involves the tetrahedral coordination of zinc. Two glutamic acid residues (E127 and E131) from each symmetry related chain A and C, form a tight interaction stabilizing crystal formation by zinc binding. The second site is more extensive, involving zinc coordination at Switch II and several van der waal and polar interactions (Figure 6.5). At the surface, each Rac1 monomer forms the same bonding interactions with a related Rac1 molecule across a 2-fold crystallographic symmetry axis. For simplicity, the interaction between chain A and another chain A from a symmetry related asymmetric unit is used as an example (Figure 6.5).

The contact surface including Switch II is at a crystallographic 2-fold axis of symmetry and is primarily polar in nature (Figure 6.5). Each tyrosine residue 64 forms a water mediated hydrogen bond with the main chain of residue H104 from a symmetry related Rac1 molecule. Specifically, the hydroxyl group of Y64 in one chain contacts a water molecule which is hydrogen bonded to the main chain carbonyl oxygen of H104
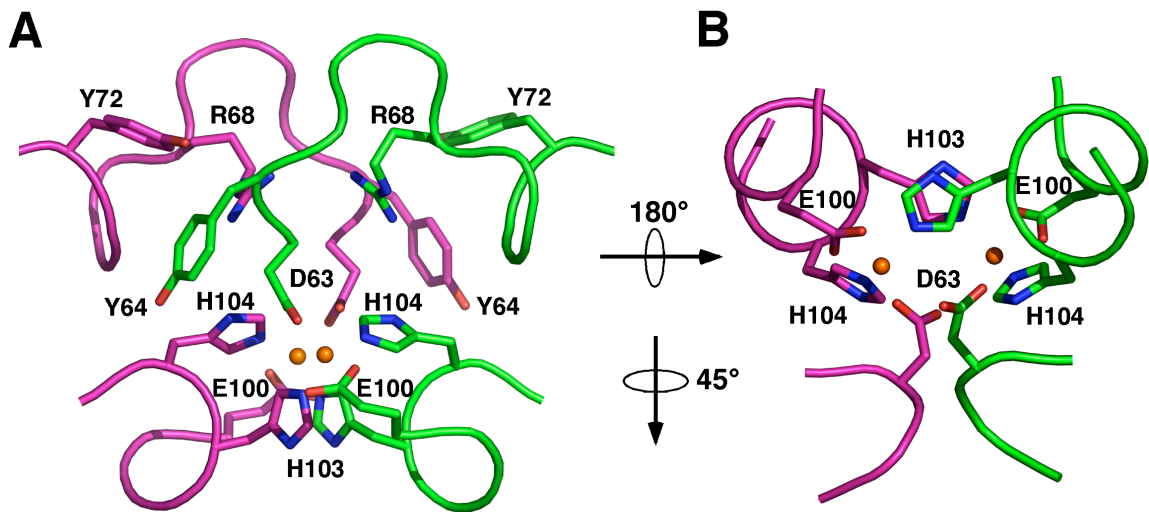
Figure 6.5: Zinc Coordination site at a 2-fold Crystallographic Axis of Symmetry
(A) Chain A (magenta) and a symmetry related Chain A (green) are shown and the intermolecular contacts drawn. Zinc is colored in dark orange and all amino acids are labeled in black.

(B) A view of the residues at the interface that are involved in zinc coordination after a 180° and 45° rotation as described in the figure. Chain A is colored in magenta and a symmetry related Chain A is colored in green.

from the related chain. R68 of each monomer makes a hydrogen bond from its side chain to the main chain carbonyl oxygen of the symmetry related D63, and E62 of chain A is part of a hydrogen bond network with the imidazole ring of H104, a water molecule, and the hydroxyl group of S71 (not shown). This network is completely within the same protein chain, aside for the solvent interactions, and seems to help position H104 for its role in zinc coordination (Figure 6.3C and Figure 6.5).

The core interactions of this surface are residues D63, E100, H104, and H103 from the symmetry related Rac1 molecule, which coordinate a zinc atom (Figure 6.3C and Figure 6.5A). This intermolecular interaction has two coordinated zinc atoms where

each monomer donates residue H103 as a zinc coordinated residue in a "swapping" interaction to its symmetry related element. Each H103 residue is further stabilized by van der waal stacking with the symmetry related H103 residue. This coordination is also tetrahedral and is reminiscent of the zinc coordination seen in the SEC2 crystal structure (Papageorgiou, Acharya et al. 1995). In this structure, zinc was coordinated by two acidic residues and two histidines, with one residue being donated from a crystallographically related molecule. In their studies, zinc was also reported to be needed for efficient crystallization of SEC2.

6.5 Zinc Induced Conformations of Switch I and Switch II

The conformation of Switch I when compared to known structures, shows that it adopts the conical GDP bound conformation as expected (Figure 6.6). Switch I is almost identical in conformation to the YpkA-Rac1 GDP complex, and thus identical to the conformation seen in RhoGDI-Rac1GDP complexes and RhoA GDP (Wei, Zhang et al. 1997; Prehna, Ivanov et al. 2006). What is even more striking is that the zinc coordinated Rac1 trimer, allows T35 to be stabilized by a hydrogen bonding interaction with a water molecule and N39 of Switch I in a subsequent Rac1 monomer (Figure 6.4). Thr35 of Rac1 is important in magnesium binding, and thus GDP binding, and molecules such as GDIs specifically stabilize this residue with a hydrogen bonding interaction to prevent nucleotide dissociation (Scheffzek, Stephan et al. 2000; Grizot, Faure et al. 2001). It seems possible from this analysis that the oligomerization induced by zinc stabilizes Switch I region in the GDP bound state and would prevent GDP dissociation.
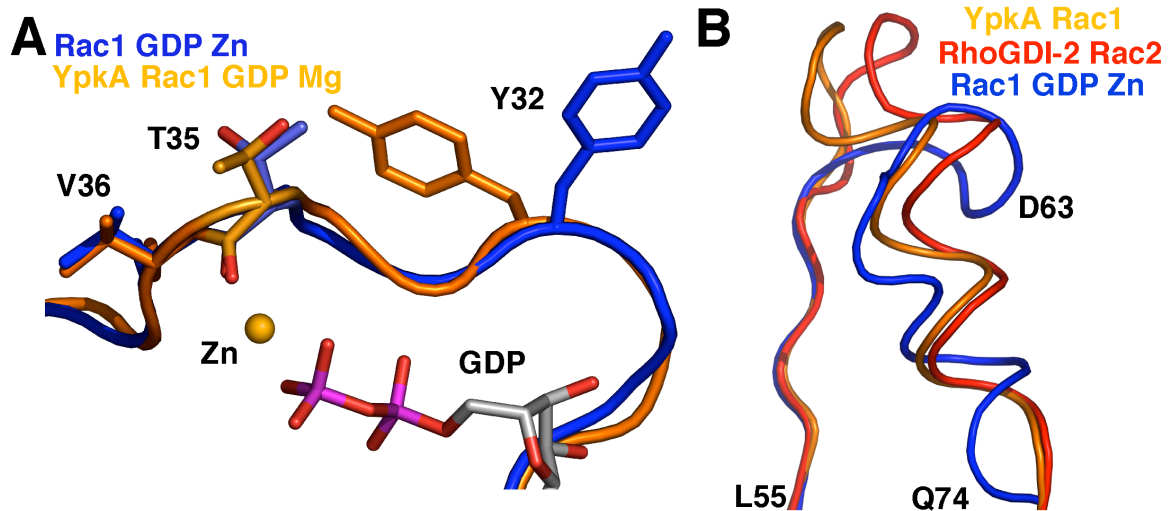
Figure 6.6: The Conformations of Switch I and Switch II

(A) The conformation of Switch I is the conical GDP bound form. The Switch I from the Rac1 GDP Zinc complex is shown in blue and aligned to a YpkA (434-732) Rac1 GDP complex shown in orange. The zinc ion and the GDP are shown from the Rac1 GDP Zinc complex.

(B) Switch II adopts a novel conformation. Switch II (blue) was aligned with a YpkA (434-732) Rac1 GDP complex (orange) and the RhoGDI-2 Rac2 complex (red). The approximate locations of residues in the Switch II structure of the Rac1 GDP Zinc complex are labeled in black.

As previously described in Figure 6.5, Switch II is also involved in zinc coordination, and undergoes a drastic conformational change upon binding zinc. Normally, Switch II adopts one of two discrete canonical conformations based on its bound nucleotide (Figure 6.6). In contrast, Switch II in the Rac1-GDP-Zn complex structure seems to adopt a completely novel fold based on zinc coordination. The coordination of zinc alters the fold of Switch II, rearranging hydrophobic packing interactions and significantly displacing key elements of Switch I. Normally, Y72 participates in the hydrophobic core at Switch II, but as shown in Figure 6.5, it rotates out

to face a symmetry related Rac1 molecule to make packing interactions. Its role in stabilizing the hydrophobic core is filled by L70, which moves into almost the same position that Y72 fills in other Rac1 structures. Residue Q61, which is involved in the intrinsic GTPase activity of Rac1 is moved ~7.4Å, and residues E62 and D63, which is involved directly in zinc coordination, are both moved ~10Å from their position as compared to the GDP bound conformation of Rac1.

6.6 Biological Considerations of Zinc Coordination

The structure of a Rac1-GDP-Zn complex at 1.9Å resolution reveals the surprising ability of Rac1 molecules to form a trimer due to the coordination of zinc. Moreover, zinc coordination sites at packing surfaces stabilize the crystal structure, and zinc coordination influences the conformation of the biologically important Switch regions. Additionally, although magnesium is known to form an octahedral coordination site at Switch I to facilitate the binding of GDP, our data suggests that zinc can replace the role of magnesium in GDP binding. As the switch regions are involved in the coordination of zinc, it is tempting to speculate that in some conditions zinc may be a regulator of Rac1 signaling.

Zinc is found in a large number of crystal structures, and can serve a myriad of biological purposes (Alberts, Nadassy et al. 1998). The coordination of zinc by biological structures has been shown to be important in catalysis, such as the protease thermolysin and the alcohol dehydrogenase htADH (Holden and Matthews 1988; Ceccarelli, Liang et al. 2004), structural stabilization, also in alcohol dehydrogenases and

in the tumor suppressor p53 (Meplan, Richard et al. 2000; Ceccarelli, Liang et al. 2004), and for macromolecular binding, such as zinc fingers (Brown 2005). However, other crystal structures are known where zinc is simply present in the crystal and has no biological function. An example is the zinc containing structure of the protease tonin, which has a zinc coordination site similar to that of SEC2 (Papageorgiou, Acharya et al. 1995). In this structure, the protein was crystallized in zinc containing buffer and it was concluded that zinc had no biological role within the context of the protease (Fujinaga and James 1987).

Despite the observation that zinc is present in our crystal structure and that zinc can replace the role of magnesium in GDP binding, one must consider that the cytosolic concentration of zinc relative to magnesium is extremely low - femtomolar ranges for zinc compared to millimolar ranges for magnesium (Outten and O'Halloran 2001; Dudev and Lim 2003). In the cell, the concentration of zinc is kept far below that of magnesium due to the sequestering of zinc into specialized compartments (Dudev 2001; Eide 2006). Eukaryotic cells have devoted several genes to the regulation of zinc concentration within various cellular compartments, store free zinc in "zincosomes." and zinc concentration has been keyed to apoptosis (Truong-Tran, Ho et al. 2000; Eide 2006). Such considerations argue that it is unlikely that zinc binding would play a role in the regulation of RhoGTPase function *in vivo*, and that the novel conformations induced by zinc binding will not likely be present in the normal physiological state of the cell. Therefore, despite the intriguing possibilities raised for the regulation of RhoGTPase function by zinc, without further biological mandate, our results remain at present tied tightly to the artificial environment of the crystallization conditions.

CHAPTER SEVEN:

CONCLUSIONS

7.1 The *Yersinia* Protein Kinase A GTPase Binding Domain

Virulence in *Yersinia pseudotuburculosis* depends upon the translocated virulence factor YpkA(Galyov, Hakansson et al. 1993), a protein that is very highly conserved in the plague pathogen, *Yersinia pestis*. Despite the known importance to virulence, little has been forthcoming in understanding the mechanism of activity of YpkA. This has been particularly true of the C-terminal domain of the protein, which while known to bind Rho GTPases, has remained enigmatic in terms of function.

The co-crystal structure of a C-terminal domain of YpkA and Rac1 reveals that this bacterial virulence factor mimics host Rho-family GDI proteins in its binding to the GTPase, and also in its ability to inhibit nucleotide exchange. Loss-of-contact mutations in YpkA that impair Rac1 and RhoA binding abolish this GDI-like activity, and severely diminish the cytoskeletal disruption induced by this domain. Furthermore, these mutations severely decrease virulence in a mouse model of infection. Altogether, these data strongly suggest that YpkA mimics host GDI proteins by acting as an "off switch" to modulate the Rac1-associated signaling pathways that regulate host cytoskeletal structure. When considering the role of YpkA with the function of the other Yops, we find that the GDI-like activity of YpkA complements the activities of YopT and YopE. In other words, YpkA, YopT, and YopE drive the Rho-family GTPase signaling pathway to the "off" state which results in the deregulation of the host actin cytoskeleton and therefore an inhibition of phagocytosis and a host immune response.

As described in Chapter 1 and Chapter 3, previous work with YpkA predicted that the small GTPase binding surface might consist of multiple ACC finger-like elements, similar to those found in downstream effector kinases such as PKN (Maesaki, Ihara et al. 1999; Maesaki, Shimizu et al. 1999; Dukuzumuremyi, Rosqvist et al. 2000). The crystal structure of YpkA (434-732) shows that many of the previously predicted ACC finger segments either map to loop regions or are in inaccessible regions (see Chapter 3.2). Additionally, structural alignments with ACC finger elements and subsequent mutagenesis experiments showed that these structural predictions were not sufficient to locate the elements necessary for GTPase binding (see Chapter 3.2). This comparative analysis taken with the crystal structure of a YpkA Rac1 complex shows that YpkA does not contain an ACC coiled-coil finger, and YpkA does not bind across the β-sheet region of the small GTPase (β2 and β3 of RhoA/Rac1). Instead, YpkA contributes to the formation of an inter-molecular helical cluster with Switch II and directly contacts Switch I, influencing their conformations. The YpkA binding to Rac1 is more similar in location to the secondary contact in the asymmetric unit of the crystal structure of PKN/RhoA, although the details of the structures are quite different (Maesaki, Ihara et al. 1999; Maesaki, Shimizu et al. 1999; Dukuzumuremyi, Rosqvist et al. 2000).

In Chapter 4.5 we show that the GDI-like activity is essential for virulence in *Yersinia*, although we find a curious experimental result where YpkA null mutants seem to have no effect on virulence in *Yersinia*. We are not alone in these observations as similar phenotypes have been reported by other groups (Logsdon and Mecsas 2003; Trulzsch, Sporleder et al. 2004). As we have already postulated, one possible explanation for these results is that null mutations in *ypkA* result in increased

translocation of other Yop virulence factors *in vivo*, which compensates for the loss of YpkA function. More precisely, considering that the T3SS is a highly regulated system requiring multiple signals for proper secretion, some of which are encoded within the protein sequence of the Yop (one for secretion by the T3SS and another for interaction with a chaperone) that deletion of an entire gene may unbalance this system. The loss of an entire gene may simply allow greater quantities of the other Yops to be secreted at a higher rate, thus compensating for the loss of YpkA. Another possibility is that there is an unknown Yop that is redundant with the GTPase binding of YpkA. In the case where YpkA is completely deleted the T3SS may compensate by the secretion of higher levels of this redundant Yop, which then shows a phenotype that YpkA is not essential for virulence. On the other hand, point mutants which still allow a full length but inactive protein to be secreted would not be compensated by another Yop, thus showing a loss of virulence. A final possibility is that the deletion of an entire gene on the pYV plasmid itself causes deregulation or other unforeseeable changes in the expression of the Yops. Although other possibilities exist and it is presently difficult to experimentally probe these possibilities, what is certain is that ideally virulence should be judged based on those alterations, such as point mutations in contrast to entire gene deletion, that disturb the system as minimally as possible.

7.2 The *Yersinia* Protein Kinase and Drug Design

In agreement with past results, our experimental results with the kinase domain of YpkA show that activity is dependent upon an interaction with actin (see Section 5.1).

Although, in contrast to past results, we show that the simple deletion of a C-terminal element does not necessary correlate to actin binding. Previous results showed that the removal the C-terminal 20 amino acids abolished kinase activity and actin binding (Juris, Rudolph et al. 2000). Our results show that all of the C-terminal deletion constructs are poorly soluble and are thus most likely poorly folded (Figure 3.1). Specifically, we created several constructs with C-terminal deletions that removed the last helix, all of which were significantly destabilized (Compare Figure 3.1, 115-701 to 115-732 and 434-701/705/718 to 434-732). As described in Chapter 3, the structure reveals that a removal of this helical region would extensively expose the hydrophobic core and remove a large segment stabilizing the fold of this sub-domain. Thus, although we also show in agreement with past analysis that some similarity with the actin binding protein coronin does exists at the sequence level, both our biochemical and structural observations indicate that it is unclear if this C-terminal segment is responsible for actin binding or kinase activation. These arguments show that further studies are required to examine the actin-associated aspects of kinase activation.

Recently it has been described that YpkA is auto-phosphorylated in a region in the N-terminus of YpkA, and that interaction with this N-terminal segment in addition to the C-terminal coronin homology region is necessary for full kinase activity (Trasak, Zenner et al. 2007). Considering our biochemical and structural observations as described above, again, it is unclear if the C-terminal region is in fact interacting with actin. Additionally, although Trasak and colleagues found phosphorylation sites by mass-spectrometry, our results show that these sites of modification are not necessary for activity, nor are at the very least the only sites of phosphorylation (see Chapter 5.1).

Specifically, Trasak et al. found that residues S90 and S95 in YopO (S93 and S98 in YpkA) are sites of auto-phosphorylation and that this modification was necessary for the full activity. In contrast, our construct YpkA (115-732) seems to be just as active as YpkA (1-732) and shows similar levels of auto-phosphorylation as the wild type (Figure 5.1). Not only does our construct display full activity, but does not contain the N-terminal region that was shown to be the sites of auto-phosphorylation. Our data demonstrate that it is still unclear how YpkA interacts with and is activated by actin. What is promising is that the work by Trasak and colleagues has revealed the necessary steps to form a complex between YpkA and G-actin that can be isolated by gel filtration. The solution of an X-ray crystal structure of such a complex seems to be necessary to reveal the mechanism of actin binding and kinase activation.

Although no X-ray crystal structure currently exists for the YpkA kinase domain, we have constructed a model for use in virtual screening for inhibitors. This has allowed us to characterize some potential drug scaffolds for future development into anti-plague therapeutics. Despite the use of a YpkA model based largely on similarities to MAPK, we have managed find two inhibitors (compounds 3 and 4, Figure 5.6) that are specific for YpkA as compared to MAPK. Additionally, as mentioned in Chapter 5, these compounds seem to be novel kinase inhibitors and are thus especially promising. Although these compounds are the first characterized small molecule inhibitors of YpkA, representing a large first step in antibiotic design, there is still much work to be done to create a viable drug. Subsequent rounds of modification and rescreening of the compounds could yield better drugs, in addition to the need of the X-ray crystal structure of the YpkA kinase domain. The solution of this structure would allow a more accurate

picture of the active site for screening, not only producing better inhibitors and more specific inhibitors, but would also open the possibility for the solution of co-crystal structures of YpkA with inhibitors. Such co-crystals would reveal the exact mode of binding and thus allow for structure driven drug design. Finally, an exploration of compounds that can cross the host cell membrane and are viable within a host cell without high toxicity is also necessary for the creation of a final antibiotic product.

7.3 A Rac1-GDP-Zinc Complex

The solution of a Rac1-GDP-Zinc complex suggests a hypothetical biological role for zinc in small GTPase biology. The conformation and binding interactions involving Switch I, Switch II, and zinc, imply that under certain conditions zinc may be able to serve as a signaling regulator. The structure of Switch I is stabilized not only by zinc coordination, but also by significant interactions with other Rac1 monomers within the asymmetric unit. This binding mode would not only shield Switch I from its binding regulators and effectors, but in effect locks down the Switch I conformation in the GDP bound, or signaling inert state. Even more striking is that zinc coordination at Switch II induces a completely novel fold of this region, which would most likely prevent the ability of known GAPs, GEFs, GDIs, or other downstream effectors from binding the small GTPase. The novel conformation of Switch II induced by zinc binding underscores the switch regions propensity for modulation and possibility to conform to other roles as not yet previously characterized. This overall effect could serve for a novel regulation of Rac1, particularly with sensitivity to zinc concentration within the cell, resulting in zinc

based signaling events. This is not completely hypothetical, as mammalian cells have devoted several genes to the regulation of zinc concentration within various cellular compartments(Eide 2006). Any deregulation or stimuli to alter the zinc concentration within the cytosol could result in a novel pathway to alter the conformation of Rac1, and thus serve as a signaling event. This may reflect a new zinc-dependent signaling mode of the GTPase, perhaps binding to unknown Rac1 effectors.

Although zinc coordination by Rac1 offers a potential new facet to small GTPase biology, especially considering that zinc can replace the role of magnesium in GDP binding, one must consider that the cytosolic concentration of zinc relative to magnesium is low (Outten and O'Halloran 2001). Based on this and other data, it is postulated that due to the higher affinity of zinc over magnesium for typical metal coordination sites in proteins, biological molecules that coordinate magnesium only do so due to the evolution of the cellular machinery that governs metal ion concentrations in cellular compartments (Dudev 2001; Eide 2006). Such a postulate would explain why Rac1 normally coordinates magnesium instead of zinc, and fits with our observations that high zinc concentrations cause the replacement of zinc for magnesium, and induce the switch regions of Rac1 to adopt specific, and in the case of Switch II, novel conformations that could potentially interfere with its normal signaling role. Although the possibilities for GTPase zinc regulation are intriguing, as described above and in greater detail in Chapter 6, without further biochemical characterization and considering the low zinc content of the cytosol relative to magnesium, it is unclear if this trimer complex is an artifact of crystallization or if zinc is somehow involved in a biological context.

REFERENCES

(1994). The CCP4 suite: programs for protein crystallography. <u>Acta Crystallogr D Biol Crystallogr</u>. **50:** 760-3.

Aepfelbacher, M. and J. Heesemann (2001). "Modulation of Rho GTPases and the actin cytoskeleton by Yersinia outer proteins (Yops)." <u>Int J Med Microbiol</u> **291**(4): 269-76.

Aepfelbacher, M., R. Zumbihl, et al. (1999). "The tranquilizing injection of Yersinia proteins: a pathogen's strategy to resist host defense." <u>Biol Chem</u> **380**(7-8): 795-802.

Agrain, C., I. Sorg, et al. (2005). "Secretion of YscP from Yersinia enterocolitica is essential to control the length of the injectisome needle but not to change the type III secretion substrate specificity." <u>Mol Microbiol</u> **57**(5): 1415-27.

Akeda, Y. and J. E. Galan (2005). "Chaperone release and unfolding of substrates in type III secretion." <u>Nature</u> **437**(7060): 911-5.

Aktories, K. and J. T. Barbieri (2005). "Bacterial cytotoxins: targeting eukaryotic switches." <u>Nat Rev Microbiol</u> **3**(5): 397-410.

Alberts, I. L., K. Nadassy, et al. (1998). "Analysis of zinc binding sites in protein crystal structures." <u>Protein Sci</u> **7**(8): 1700-16.

Barz, C., T. N. Abahji, et al. (2000). The Yersinia Ser/Thr protein kinase YpkA/YopO directly interacts with the small GTPases RhoA and Rac-1. <u>FEBS Lett</u>. **482:** 139-43.

Birtalan, S. C., R. M. Phillips, et al. (2002). "Three-dimensional secretion signals in chaperone-effector complexes of bacterial pathogens." Mol Cell **9**(5): 971-80.

Bishop, A. L. and A. Hall (2000). "Rho GTPases and their effector proteins." Biochem J **348 Pt 2**: 241-55.

Black, D. S. and J. B. Bliska (2000). The RhoGAP activity of the Yersinia pseudotuberculosis cytotoxin YopE is required for antiphagocytic function and virulence. Mol Microbiol. **37:** 515-27.

Bliska, J. B. (2006). "Yersinia inhibits host signaling by acetylating MAPK kinases." ACS Chem Biol **1**(6): 349-51.

Briem, H. and J. Gunther (2005). "Classifying "kinase inhibitor-likeness" by using machine-learning methods." Chembiochem **6**(3): 558-66.

Brooks, G., Butel, Janet S., and Morse, Stephen A. (2001). Medical Microbiology, McGraw-Hill.

Brown, R. S. (2005). "Zinc finger proteins: getting a grip on RNA." Curr Opin Struct Biol **15**(1): 94-8.

Brunger, A. T., P. D. Adams, et al. (1998). "Crystallography & NMR system: A new software suite for macromolecular structure determination." Acta Crystallogr D Biol Crystallogr **54**(Pt 5): 905-21.

Buchwald, G., A. Friebel, et al. (2002). Structural basis for the reversible activation of a Rho protein by the bacterial toxin SopE. Embo J. **21:** 3286-95.

Burghout, P., R. van Boxtel, et al. (2004). "Structure and electrophysiological properties of the YscC secretin from the type III secretion system of Yersinia enterocolitica." J Bacteriol **186**(14): 4645-54.

Cavasotto, C. N. and R. A. Abagyan (2004). "Protein flexibility in ligand docking and virtual screening to protein kinases." <u>J Mol Biol</u> **337**(1): 209-25.

Ceccarelli, C., Z. X. Liang, et al. (2004). "Crystal structure and amide H/D exchange of binary complexes of alcohol dehydrogenase from Bacillus stearothermophilus: insight into thermostability and cofactor binding." <u>Biochemistry</u> **43**(18): 5266-77.

Cherfils, J. (2001). "Structural mimicry of DH domains by Arfaptin suggests a model for the recognition of Rac-GDP by its guanine nucleotide exchange factors." <u>FEBS Lett</u> **507**(3): 280-4.

Cherfils, J. and P. Chardin (1999). "GEFs: structural basis for their activation of small GTP-binding proteins." <u>Trends Biochem Sci</u> **24**(8): 306-11.

Chromy, B. A., M. W. Choi, et al. (2005). "Proteomic characterization of Yersinia pestis virulence." <u>J Bacteriol</u> **187**(23): 8172-80.

Claussen, H., C. Buning, et al. (2001). "FlexE: efficient molecular docking considering protein structure variations." <u>J Mol Biol</u> **308**(2): 377-95.

Cornelis, G. R. (1998). "The Yersinia deadly kiss." <u>J Bacteriol</u> **180**(21): 5495-504.

Cornelis, G. R. (2000). "Molecular and cell biology aspects of plague." <u>Proc Natl Acad Sci U S A</u> **97**(16): 8778-83.

Cornelis, G. R. (2002). "Yersinia type III secretion: send in the effectors." <u>J Cell Biol</u> **158**(3): 401-8.

Cornelis, G. R. (2002). "The Yersinia Ysc-Yop virulence apparatus." <u>Int J Med Microbiol</u> **291**(6-7): 455-62.

Cornelis, G. R., A. Boland, et al. (1998). "The virulence plasmid of Yersinia, an antihost genome." <u>Microbiol Mol Biol Rev</u> **62**(4): 1315-52.

Dale, C. and N. A. Moran (2006). "Molecular interactions between bacterial symbionts and their hosts." Cell **126**(3): 453-65.

DeLano, W. L. (2002). The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos, CA, USA.

Derewenda, U., A. Mateja, et al. (2004). "The structure of Yersinia pestis V-antigen, an essential virulence factor and mediator of immunity against plague." Structure **12**(2): 301-6.

Dudev, T. and C. Lim (2003). "Principles governing Mg, Ca, and Zn binding and selectivity in proteins." Chem Rev **103**(3): 773-88.

Dudev, T. a. L., C. (2001). "Metal Selectivity in Metalloproteins: Zn^{2+} vs Mg^{2+}." J. Phys. Chem. B **105**: 4446-4452.

Dukuzumuremyi, J. M., R. Rosqvist, et al. (2000). "The Yersinia protein kinase A is a host factor inducible RhoA/Rac-binding virulence factor." J Biol Chem **275**(45): 35281-90.

Dvorsky, R. and M. R. Ahmadian (2004). "Always look on the bright site of Rho: structural implications for a conserved intermolecular interface." EMBO Rep **5**(12): 1130-6.

Dvorsky, R., L. Blumenstein, et al. (2004). "Structural insights into the interaction of ROCKI with the switch regions of RhoA." J Biol Chem **279**(8): 7098-104.

Eide, D. J. (2006). "Zinc transporters and the cellular trafficking of zinc." Biochim Biophys Acta **1763**(7): 711-22.

Ennifar, E., P. Walter, et al. (2001). "An efficient method for solving RNA structures: MAD phasing by replacing magnesium with zinc." <u>Acta Crystallogr D Biol Crystallogr</u> **57**(Pt 2): 330-2.

Evdokimov, A. G., D. E. Anderson, et al. (2001). "Unusual molecular architecture of the Yersinia pestis cytotoxin YopM: a leucine-rich repeat protein with the shortest repeating unit." <u>J Mol Biol</u> **312**(4): 807-21.

Evdokimov, A. G., J. E. Tropea, et al. (2002). "Crystal structure of the Yersinia pestis GTPase activator YopE." <u>Protein Sci</u> **11**(2): 401-8.

Fadouloglou, V. E., A. P. Tampakaki, et al. (2004). "Structure of HrcQB-C, a conserved component of the bacterial type III secretion systems." <u>Proc Natl Acad Sci U S A</u> **101**(1): 70-5.

Fischer, D., A. Elofsson, et al. (2001). "CAFASP2: the second critical assessment of fully automated structure prediction methods." <u>Proteins</u> **Suppl 5**: 171-83.

Fiser, A. and A. Sali (2003). "Modeller: generation and refinement of homology-based protein structure models." <u>Methods Enzymol</u> **374**: 461-91.

Ford, M. G., W. R. Pitt, et al. (2004). "Selecting compounds for focused screening using linear discriminant analysis and artificial neural networks." <u>J Mol Graph Model</u> **22**(6): 467-72.

Frean, J., K. P. Klugman, et al. (2003). "Susceptibility of Yersinia pestis to novel and conventional antimicrobial agents." <u>J Antimicrob Chemother</u> **52**(2): 294-6.

Frean, J. A., L. Arntzen, et al. (1996). "In vitro activities of 14 antibiotics against 100 human isolates of Yersinia pestis from a southern African plague focus." <u>Antimicrob Agents Chemother</u> **40**(11): 2646-7.

Friebel, A. and W. D. Hardt (2000). Purification and biochemical activity of Salmonella exchange factor SopE. Methods Enzymol. **325:** 82-91.

Fujinaga, M. and M. N. James (1987). "Rat submaxillary gland serine protease, tonin. Structure solution and refinement at 1.8 A resolution." J Mol Biol **195**(2): 373-96.

Galan, J. E. and A. Collmer (1999). "Type III secretion machines: bacterial devices for protein delivery into host cells." Science **284**(5418): 1322-8.

Galimand, M., E. Carniel, et al. (2006). "Resistance of Yersinia pestis to antimicrobial agents." Antimicrob Agents Chemother **50**(10): 3233-6.

Galimand, M., A. Guiyoule, et al. (1997). "Multidrug resistance in Yersinia pestis mediated by a transferable plasmid." N Engl J Med **337**(10): 677-80.

Galyov, E. E., S. Hakansson, et al. (1993). "A secreted protein kinase of Yersinia pseudotuberculosis is an indispensable virulence determinant." Nature **361**(6414): 730-2.

Galyov, E. E., S. Hakansson, et al. (1994). Characterization of the operon encoding the YpkA Ser/Thr protein kinase and the YopJ protein of Yersinia pseudotuberculosis. J Bacteriol. **176:** 4543-8.

Ghosh, P. (2004). "Process of protein transport by the type III secretion system." Microbiol Mol Biol Rev **68**(4): 771-95.

Grizot, S., J. Faure, et al. (2001). "Crystal structure of the Rac1-RhoGDI complex involved in nadph oxidase activation." Biochemistry **40**(34): 10007-13.

Gruenheid, S. and B. B. Finlay (2003). "Microbial pathogenesis and cytoskeletal function." Nature **422**(6933): 775-81.

Guarner, F. and J. R. Malagelada (2003). "Gut flora in health and disease." Lancet **361**(9356): 512-9.

Hakansson, S., E. E. Galyov, et al. (1996). "The Yersinia YpkA Ser/Thr kinase is translocated and subsequently targeted to the inner surface of the HeLa cell plasma membrane." Mol Microbiol **20**(3): 593-603.

Hakoshima, T., T. Shimizu, et al. (2003). "Structural basis of the Rho GTPase signaling." J Biochem (Tokyo) **134**(3): 327-31.

Hardt, W. D., L. M. Chen, et al. (1998). S. typhimurium encodes an activator of Rho GTPases that induces membrane ruffling and nuclear responses in host cells. Cell. **93:** 815-26.

Henderson, D. A. (1999). "The looming threat of bioterrorism." Science **283**(5406): 1279-82.

Hirshberg, M., R. W. Stockley, et al. (1997). "The crystal structure of human rac1, a member of the rho-family complexed with a GTP analogue." Nat Struct Biol **4**(2): 147-52.

Holden, H. M. and B. W. Matthews (1988). "The binding of L-valyl-L-tryptophan to crystalline thermolysin illustrates the mode of interaction of a product of peptide hydrolysis." J Biol Chem **263**(7): 3256-60.

Huse, M. and J. Kuriyan (2002). The conformational plasticity of protein kinases. Cell. **109:** 275-82.

Hutchinson, J. P. and J. F. Eccleston (2000). "Mechanism of nucleotide release from Rho by the GDP dissociation stimulator protein." Biochemistry **39**(37): 11348-59.

Inglesby, T. V., D. T. Dennis, et al. (2000). Plague as a biological weapon: medical and public health management. Working Group on Civilian Biodefense. <u>Jama</u>. **283:** 2281-90.

Jaffe, A. B. and A. Hall (2005). "Rho GTPases: biochemistry and biology." <u>Annu Rev Cell Dev Biol</u> **21**: 247-69.

Janjusevic, R., R. B. Abramovitch, et al. (2006). "A bacterial inhibitor of host programmed cell death defenses is an E3 ubiquitin ligase." <u>Science</u> **311**(5758): 222-6.

Juris, S. J., A. E. Rudolph, et al. (2000). "A distinctive role for the Yersinia protein kinase: actin binding, kinase activation, and cytoskeleton disruption." <u>Proc Natl Acad Sci U S A</u> **97**(17): 9431-6.

Juris, S. J., K. Shah, et al. (2006). "Identification of otubain 1 as a novel substrate for the Yersinia protein kinase using chemical genetics and mass spectrometry." <u>FEBS Lett</u> **580**(1): 179-83.

Kerschen, E. J., D. A. Cohen, et al. (2004). "The plague virulence protein YopM targets the innate immune response by causing a global depletion of NK cells." <u>Infect Immun</u> **72**(8): 4589-602.

Letzelter, M., I. Sorg, et al. (2006). "The discovery of SycO highlights a new function for type III secretion effector chaperones." <u>Embo J</u> **25**(13): 3223-33.

Lilic, M., M. Vujanac, et al. (2006). "A common structural motif in the binding of virulence factors to bacterial secretion chaperones." <u>Mol Cell</u> **21**(5): 653-64.

Logsdon, L. K. and J. Mecsas (2003). "Requirement of the Yersinia pseudotuberculosis effectors YopH and YopE in colonization and persistence in intestinal and lymph tissues." Infect Immun **71**(8): 4595-607.

Maesaki, R., K. Ihara, et al. (1999). "The structural basis of Rho effector recognition revealed by the crystal structure of human RhoA complexed with the effector domain of PKN/PRK1." Mol Cell **4**(5): 793-803.

Maesaki, R., T. Shimizu, et al. (1999). Biochemical and crystallographic characterization of a Rho effector domain of the protein serine/threonine kinase N in a complex with RhoA. J Struct Biol. **126:** 166-70.

Marketon, M. M., R. W. DePaolo, et al. (2005). "Plague bacteria target immune cells during infection." Science **309**(5741): 1739-41.

Marra, A. (2006). "Targeting virulence for antibacterial chemotherapy: identifying and characterising virulence factors for lead discovery." Drugs R D **7**(1): 1-16.

McCormick, J. B. (1998). Epidemiology of emerging/re-emerging antimicrobial-resistant bacterial pathogens. Curr Opin Microbiol. **1:** 125-9.

Meplan, C., M. J. Richard, et al. (2000). "Metalloregulation of the tumor suppressor protein p53: zinc mediates the renaturation of p53 after exposure to metal chelators in vitro and in intact cells." Oncogene **19**(46): 5227-36.

Mol, C. D., D. Fabbro, et al. (2004). "Structural insights into the conformational selectivity of STI-571 and related kinase inhibitors." Curr Opin Drug Discov Devel **7**(5): 639-48.

Muegge, I. and I. J. Enyedy (2004). "Virtual screening for kinase targets." Curr Med Chem **11**(6): 693-707.

Murshudov, G. N., A. A. Vagin, et al. (1997). "Refinement of macromolecular structures by the maximum-likelihood method." <u>Acta Crystallogr D Biol Crystallogr</u> **53**(Pt 3): 240-55.

Nejedlik, L., T. Pierfelice, et al. (2004). "Actin distribution is disrupted upon expression of Yersinia YopO/YpkA in yeast." <u>Yeast</u> **21**(9): 759-68.

Noble, M. E., J. A. Endicott, et al. (2004). "Protein kinase inhibitors: insights into drug design from structure." <u>Science</u> **303**(5665): 1800-5.

Ostermeier, C. and A. T. Brunger (1999). "Structural basis of Rab effector specificity: crystal structure of the small G protein Rab3A complexed with the effector domain of rabphilin-3A." <u>Cell</u> **96**(3): 363-74.

Otwinowski, Z., and Minor, W. (1997). "Processing of X-ray diffraction data collected in oscillation mode." <u>Methods Enzymol</u> **276**: 307.

Outten, C. E. and T. V. O'Halloran (2001). "Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis." <u>Science</u> **292**(5526): 2488-92.

Papageorgiou, A. C., K. R. Acharya, et al. (1995). "Crystal structure of the superantigen enterotoxin C2 from Staphylococcus aureus reveals a zinc-binding site." <u>Structure</u> **3**(8): 769-79.

Perrakis, A., R. Morris, et al. (1999). "Automated protein model building combined with iterative structure refinement." <u>Nat Struct Biol</u> **6**(5): 458-63.

Perry, R. D. and J. D. Fetherston (1997). "Yersinia pestis--etiologic agent of plague." <u>Clin Microbiol Rev</u> **10**(1): 35-66.

Persson, C., T. Sjoblom, et al. (2004). "Preferential oxidation of the second phosphatase domain of receptor-like PTP-alpha revealed by an antibody against oxidized protein tyrosine phosphatases." Proc Natl Acad Sci U S A **101**(7): 1886-91.

Potterton, E., P. Briggs, et al. (2003). "A graphical user interface to the CCP4 program suite." Acta Crystallogr D Biol Crystallogr **59**(Pt 7): 1131-7.

Prehna, G., M. I. Ivanov, et al. (2006). "Yersinia virulence depends on mimicry of host rho-family nucleotide dissociation inhibitors." Cell **126**(5): 869-80.

Roggenkamp, A., S. Schubert, et al. (1995). Dissection of the Yersinia enterocolitica virulence plasmid pYVO8 into an operating unit and virulence gene modules. FEMS Microbiol Lett. **134:** 69-73.

Rost, B. (1996). "PHD: predicting one-dimensional protein structure by profile-based neural networks." Methods Enzymol **266**: 525-39.

Rudolph, M. G., C. Weise, et al. (1999). Biochemical analysis of SopE from Salmonella typhimurium, a highly efficient guanosine nucleotide exchange factor for RhoGTPases. J Biol Chem. **274:** 30501-9.

Rypniewski, W. R., H. M. Holden, et al. (1993). "Structural consequences of reductive methylation of lysine residues in hen egg white lysozyme: an X-ray analysis at 1.8-A resolution." Biochemistry **32**(37): 9851-8.

Salyers, A. A. a. W., Dixie D. (2002). Bacterial Pathogenesis: A Molecular Approach, ASM Press.

Scapin, G. (2006). "Protein kinase inhibition: different approaches to selective inhibitor design." Curr Drug Targets **7**(11): 1443-54.

Scheffzek, K., I. Stephan, et al. (2000). "The Rac-RhoGDI complex and the structural basis for the regulation of Rho proteins by RhoGDI." Nat Struct Biol **7**(2): 122-6.

Shao, F., P. M. Merritt, et al. (2002). "A Yersinia effector and a Pseudomonas avirulence protein define a family of cysteine proteases functioning in bacterial pathogenesis." Cell **109**(5): 575-88.

Shao, F., P. O. Vacratsis, et al. (2003). "Biochemical characterization of the Yersinia YopT protease: cleavage site and recognition elements in Rho GTPases." Proc Natl Acad Sci U S A **100**(3): 904-9.

Stebbins, C. E. and J. E. Galan (2000). "Modulation of host signaling by a bacterial mimic: structure of the Salmonella effector SptP bound to Rac1." Mol Cell **6**(6): 1449-60.

Stebbins, C. E. and J. E. Galan (2001). "Maintenance of an unfolded polypeptide by a cognate chaperone in bacterial type III secretion." Nature **414**(6859): 77-81.

Stebbins, C. E. and J. E. Galan (2001). "Structural mimicry in bacterial virulence." Nature **412**(6848): 701-5.

Stebbins, C. E. and J. E. Galan (2003). "Priming virulence factors for delivery into the host." Nat Rev Mol Cell Biol **4**(9): 738-43.

Symons, M. and J. Settleman (2000). "Rho family GTPases: more than simple switches." Trends Cell Biol **10**(10): 415-9.

Tardy, F., F. Homble, et al. (1999). "Yersinia enterocolitica type III secretion-translocation system: channel formation by secreted Yops." Embo J **18**(23): 6793-9.

Tarricone, C., B. Xiao, et al. (2001). "The structural basis of Arfaptin-mediated cross-talk between Rac and Arf signalling pathways." Nature **411**(6834): 215-9.

Terwilliger, T. (2004). "SOLVE and RESOLVE: automated structure solution, density modification and model building." J Synchrotron Radiat **11**(Pt 1): 49-52.

Trasak, C., G. Zenner, et al. (2007). "Yersinia Protein Kinase YopO Is Activated by A Novel G-actin Binding Process." J Biol Chem **282**(4): 2268-2277.

Trulzsch, K., T. Sporleder, et al. (2004). "Contribution of the major secreted yops of Yersinia enterocolitica O:8 to pathogenicity in the mouse infection model." Infect Immun **72**(9): 5227-34.

Truong-Tran, A. Q., L. H. Ho, et al. (2000). "Cellular zinc fluxes and the regulation of apoptosis/gene-directed cell death." J Nutr **130**(5S Suppl): 1459S-66S.

Van Aelst, L. and C. D'Souza-Schorey (1997). "Rho GTPases and signaling networks." Genes Dev **11**(18): 2295-322.

Vetter, I. R. and A. Wittinghofer (2001). "The guanine nucleotide-binding switch in three dimensions." Science **294**(5545): 1299-304.

Viboud, G. I. and J. B. Bliska (2005). "Yersinia outer proteins: role in modulation of host cell signaling responses and pathogenesis." Annu Rev Microbiol **59**: 69-89.

Wang, R., L. Lai, et al. (2002). "Further development and validation of empirical scoring functions for structure-based binding affinity prediction." J Comput Aided Mol Des **16**(1): 11-26.

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

Weidow, C. L., D. S. Black, et al. (2000). "CAS/Crk signalling mediates uptake of Yersinia into human epithelial cells." Cell Microbiol **2**(6): 549-60.

Wiley, D. J., R. Nordfeldth, et al. (2006). "The Ser/Thr kinase activity of the Yersinia protein kinase A (YpkA) is necessary for full virulence in the mouse, mollifying phagocytes, and disrupting the eukaryotic cytoskeleton." Microb Pathog **40**(5): 234-43.

Winn, M. D. (2003). "An overview of the CCP4 project in protein crystallography: an example of a collaborative project." J Synchrotron Radiat **10**(Pt 1): 23-5.

Winn, M. D., M. N. Isupov, et al. (2001). "Use of TLS parameters to model anisotropic displacements in macromolecular refinement." Acta Crystallogr D Biol Crystallogr **57**(Pt 1): 122-33.

Winn, M. D., G. N. Murshudov, et al. (2003). "Macromolecular TLS refinement in REFMAC at moderate resolutions." Methods Enzymol **374**: 300-21.

Worthylake, D. K., K. L. Rossman, et al. (2000). "Crystal structure of Rac1 in complex with the guanine nucleotide exchange region of Tiam1." Nature **408**(6813): 682-8.

Wren, B. W. (2003). "The yersiniae--a model genus to study the rapid evolution of bacterial pathogens." Nat Rev Microbiol **1**(1): 55-64.

Yip, C. K., T. G. Kimbrough, et al. (2005). "Structural characterization of the molecular platform for type III secretion system assembly." Nature **435**(7042): 702-7.

Yip, C. K. and N. C. Strynadka (2006). "New structural insights into the bacterial type III secretion system." Trends Biochem Sci **31**(4): 223-30.

Zhang, Y., A. T. Ting, et al. (2005). "Inhibition of MAPK and NF-kappa B pathways is necessary for rapid apoptosis in macrophages infected with Yersinia." J Immunol **174**(12): 7939-49.