2008

Structural Analysis of Cycle Inhibiting Factor from Pathogenic Escherichia Coli

Yun-Yuan Hsu

STRUCTURAL ANALYSIS OF CYCLE INHIBITING FACTOR FROM PATHOGENIC *ESCHERICHIA COLI*

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Yun-Yuan Hsu

June 2008

STRUCTURAL ANALYSIS OF CYCLE INHIBITING FACTOR FROM PATHOGENIC *ESCHERICHIA COLI*

Yun-Yuan Hsu, Ph.D.

The Rockefeller University 2008

Bacterial pathogens have evolved a sophisticated arsenal of virulence factors to modulate host cell biology. Enteropathogenic and enterohemorrhagic *Escherichia coli* (EPEC and EHEC) use a type III protein secretion system (T3SS) to inject microbial proteins into host cells. These translocated proteins possess a diverse array of biochemical activities, reprogramming eukaryotic cell biochemistry to serve the requirements of the pathogen. Examples of translocated effector proteins hijacking essential host functions such as cytoskeleton assembly, vesicular transport, and apoptosis are numerous (Galan and Wolf-Watz, 2006). However, until recently, less progress had been made with bacteria that also possess virulence mechanisms that target the host cell cycle. Studies have demonstrated the capabilities of pathogenic bacteria in modulating host-cell vital processes with a growing family of bacterial toxins and effectors that interfere with the eukaryotic cell cycle (Nougayrede et al., 2005; Oswald et al., 2005).

The EPEC and EHEC T3SS effector Cycle inhibiting factor (Cif) is able to block host eukaryotic cell cycle progression. We present here the crystal structure of Cif. It is a divergent member of a superfamily of enzymes sharing a common Cys-His-Asp/Asn catalytic triad, which includes cysteine proteases and acetyltransferases. Mutation of these conserved active site residues abolishes the ability of Cif to block cell cycle

progression in different models. We demonstrate that Cif possesses auto-acetylation activity that is dependent on a wild-type catalytic triad, whereas irreversible inhibitors of cysteine protease activity do not affect the cytopathic phenotype. Finally, we identify the cell cycle regulatory protein Nedd8 as an interaction partner for Cif. Cif binds Nedd8 *in vitro* through both the catalytic domain and the N-terminal domain, the latter of which we demonstrate to be absolutely required for Cif activity in the host cell.

To my parents who taught me to go beyond boundaries to fulfill my dreams

獻給偉大的爸媽

ACKNOWLEDGMENTS

I would like to thank anyone who contributed in any way to the work presented here.

I would like to express my sincere gratitude to my advisor, C. Erec Stebbins, for his continuous support and guidance throughout my PhD studies. His insightful advice and encouragement during the challenging stages of my research have been invaluable and greatly appreciated.

Many thanks to my thesis committee members Seth Darst, Vincent Fischetti, and James Bliska for their time and expertise, as well as their guidance.

I want to thank my collaborators at the Ecole Nationale Veterinaire de Toulouse, France. Thank you to Prof. Eric Oswald, Gregory Jubelin, Frederic Taieb, and Jean-Philippe Nougayrede for fruitful discussions and a successful collaboration.

Thanks to everyone from the Stebbins lab. Thank you Marianne for your dedication in keeping the lab running well. Of course, I will miss her and her yummy baked goods. I would like to thank Mira, Milos, and Dragana, who all did a wonderful job teaching me when I first rotated into the lab. Mira, thanks for always being super nice and helpful. Radmila, thanks for your helpful advice as well as the perfect cup-of-joe you freshly brewed every morning. Lucas, thanks for being my go-to person whenever I had any questions and for allowing me to pick your brain for inspiration. Thanks to Gerd, Shyam, and Marshall for their companionship.

Oh Cindy, where can I start with Cindy. She was always there when I needed someone to talk to. Hey LOVE, your friendship and help are heartfelt. After all, we were both victimized by Cif/Sif. Don't worry; you will never be 'out of sight, out of mind'. Hopefully one day we can buy each other lobsters.

I would like to thank Dragana who is a great mentor and friend. I still have a lot to learn from her. She is an immensely sweet and caring person. She is also the one who got me interested in cyclomodulins as well as getting fit! She is one of many that have helped guide me to this point; the list goes on and on.

I would like to thank my undergraduate advisor Andrew D. Ellington and his graduate student Kate Bell for providing an interesting yet challenging environment that fostered my interest in science and research.

I am thankful to have met several great friends at The Rockefeller University. Boo, Andrea, Carl, Taulant, and Mary: I enjoyed every game-night and get-together that we shared. Monica and Martin, who are my M&M; I can never get enough of them.

My deepest thanks go to my wonderful family for their loving support: My father Chin-Wen, my mother Meiling, and my brothers James and Sam.

The End…

Just kidding. Lastly, I would like to thank my Eric. If it weren't for you, I wouldn't have jumped off an airplane or a 160 meters tall suspension bridge. Thanks for teaching me how to live and enjoy life. Thank you for all your love and support you have given me through the years. You have given me the strength through all of the ups and downs to keep striving and reaching for my goals and dreams. I would not be at this point in my life without you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Å	Angstrom
A/E	attaching and effacing
Acetyl-CoA	acetyl coenzyme A
AMC	amidomethylcoumarin
APC/C	anaphase-promoting complex or cyclosome
ATP	adenosine triphosphate
$CaCl_2$	calcium chloride
Cif	Cycle inhibiting factor
CPE	cytopathic effect
DTT	dithiothreitol
DOC	deoxycholate
DMSO	dimethyl sulfoxide
EDTA	ethylenediaminotetraacetic acid
EHEC	enterohemorraghic *Escherichia coli*
EPEC	enteropathogenic *Escherichia coli*
FBS	fetal bovine serum
FPLC	fast performance liquid chromatography
GST	glutathion-S-transferase
HEPES	4-(2-hydroxyethyl)-piperazine-1-ethanesulfonic acid
IPTG	isopropylthiogalacopyranoside
kD	kilodalton
LB	Luria broth
LEE	locus of enterocyte
M	molar
MCS	multiple cloning site
$MgCl_2$	magnesium chloride
MES	2-(N-Morpholino)ethanesulfonic acid
mM	millimolar
MW	molecular weight
μM	micromolar
NaCl	sodium chloride
NAT	arylamine N-acetyltransferase
NP-40	Nonidet P40
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PMSF	phenylmethylsulfonyl fluoride
SAD	Single-wavelength anomalous dispersion
SCF	Skp1/Cullin/F-box protein
SeMet	seleno-methionine
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
T3SS	type III secretion system
TBS-T	Tris-Buffered Saline Tween-20

Tris	tris-(hydroxymethyl)-aminomethane
WT	wild type
R_{work}	$\Sigma\|\|F_{obs}\| - \|F_{calc}\|\|/\Sigma\|F_{obs}\|$; $\|F_{obs}\|$ is the observed and $\|F_{calc}\|$ is the calculated structure factor amplitudes
R_{free}	The same as R_{work} which is calculated for the set of reflections that is not used during the refinement

CHAPTER ONE:

INTRODUCTION

1.1 Foodborne illnesses

Foodborne illnesses are worldwide public health concern. They are generally caused by consumption of food and water contaminated with bacteria, parasites, or viruses. Their symptoms range from mild gastrointestinal distress to severe effects such as diarrhea, vomiting, and fever. Globally, diarrheal diseases kill an estimated 2.1 million people, 1.8 million of those being children under age 5 (www.who.int). Diarrheal diseases often occur in immune compromised individuals, the elderly, and infants. The majority of foodborne illnesses are caused by pathogenic bacteria. The most common bacterial foodborne pathogens are *Campylobacter jejuni, Salmonella ssp, Escherichia coli, Staphylococcus aureus, Listeria monocytogenes, Clostridium perfringens, Vibrio parahaemolyticus, Vibrio vulnificus,* and *Shigella flexneri* (Olsen et al., 2000). The severity of symptoms caused by each pathogen varies from strain to strain.

Emerging health threats in industrialized countries are increasingly associated with pathogenic *Escherichia coli*. Each year the United States alone reports about 76 million foodborne illnesses, resulting in 325,000 hospitalization and 5,000 deaths (DuPont, 2007). The trend of diarrheal illnesses in industrialized nations is due to the increase in transmission vessels of disease causing pathogens. Centralization of food supply allows commercially available products to be mass distributed throughout the country (DuPont, 2007). For example, in 2006 bagged spinach distributed for use in

salad was contaminated with *E. coli*. There were 199 people from 26 states infected and three confirmed deaths with *E. coli* O157:H7 after the spinach consumption (CDC). This incident demonstrated how contaminated food could be quickly spread from one location to multiple regions around the county. The *E. coli* outbreak was not just a public health concern, but it additionally had economic ramifications costing the food industry millions of dollars (Lee, 2006).

1.2 *Escherichia coli*, commensal to pathogenic

Escherichia coli is a gram-negative, facultative anaerobic bacterium that populates the gastrointestinal tract of warm-blooded animals. *E. coli* was first identified in 1885 by the German pediatrician Theodor Escherich when he conducted a series of experiments to study the impact of intestinal bacteria on the gastrointestinal tract (Escherich, 1886). *E. coli* is the most widely studied and best understood bacterial species. Its genome was fully sequenced in 1997 (Blattner et al., 1997). Commensal *E. coli* is a member of the human intestinal microflora, providing the essential vitamin K2 (menaquinone) to its host organism (Bentley and Meganathan, 1982).

In 1945 John Bray and T.E.D Beavan recognized *E. coli* as a versatile gastrointestinal pathogen after multiple occurrences of infant diarrheal outbreaks in the United Kingdom. Bray isolated a strain of pathogenic *E. coli* that was absent in a healthy individual's stool (Bray, 1945). The ability for an *E. coli* strain to cause illness lies in its genetic makeup. A conserved genomic sequence is common between the commensal and pathogenic strains, allowing both to survive and colonize the host gut. In addition to a

core genomic sequence, pathogenic *E. coli* strains contain multiple gene clusters acquired through the integration of plasmids, bacteriophages, transposons, and horizontal gene transfer of pathogenicity islands (PAIs) that enhance their ability to target and infect specific tissues (Garmendia et al., 2005).

There are several virotypes of *E. coli* strains that have acquired the virulent genetic elements, which allow them to cause a broad spectrum of diseases. These pathogenic *E. coli* can cause intestinal diarrheal disease as well as extraintestinal disease such as urinary tract infections and sepsis/meningitis. The enterovirulent *E. coli* is further categorized into 6 subgroups: enteropathogenic *E. coli* (EPEC), enterohemorrhagic *E. coli* (EHEC), enterotoxigenic *E. coli* (ETEC), enteroaggregative *E. coli* (EAEC), enteroinvasive *E. coli* (EIEC), and diffusely adherent E. coli (DAEC) (Kaper et al., 2004). They all have specific virulence attributes that allow them to colonize sites apart from the lower bowel enterocytes.

Amongst the virotypes of *E. coli*, EPEC and EHEC contribute to most of the foodborne illnesses. They cause acute gastroenteritis in humans. EPEC is a major cause of infantile diarrhea in developing countries and malnutrition in the pediatric population worldwide, killing close to 2 million children each year (www.who.int). EHEC causes a wide spectrum of illnesses ranging from mild diarrhea to severe diseases such as hemolytic uremic syndrome, which causes kidney failure and even death. EHEC is distinguished from EPEC by the production of Shiga toxin (Mainil et al., 1987). Shiga toxin travels in the bloodstream from the colon to the kidney and damages renal endothelial cells, eventually leading to hemolytic uremic syndrome (Obrig, 1997).

Once inside the human body, EPEC targets the small intestine, while EHEC colonizes the colon (Kaper et al., 2004). They adhere to intestinal epithelial cells and subvert cytoskeletal processes to produce a characteristic intestinal histopathological feature known as 'attaching and effacing' (A/E) lesion (Moon et al., 1983). Bacteria adhere to the enterocytes and modulate the host cell cytoskeleton by accumulating polymerized actin beneath them. Microvilli are effaced, and pedestal-like structures arise from the epithelial cells, where bacteria are intimately attached. Although they remain extracellular, EPEC/EHEC are able to subvert host cell processes such as cytoskeletal assembly, endocytic trafficking, and cell cycle progression, which results in bacteria colonization of the mucosal site, evasion of host defenses, replication, and host cell damage. *E. coli's* ability to induce the A/E lesion and modulate host cellular processes are attributed to its 35kB chromosomal pathogenicity island called the locus of enterocyte effacement (LEE) (McDaniel et al., 1995) and the LEE-encoded multicomponent organelle called the type III secretion system (Jarvis et al., 1995).

1.3 Type III secretion system

1.3.1 The apparatus

Many Gram-negative pathogenic bacteria utilize a complex system called the Type III secretion system (T3SS) to mediate interaction with human, animal, and plant hosts. *E. coli*'s T3SS is assembled from the products of approximately 20-25 genes from the LEE pathogenicity island (Kaper, 1998). Electron microscopy studies show that the T3SS is a syringe-like apparatus that resembles the flagellar hook-basal body complex

(Kubori et al., 1998; Sekiya et al., 2001; Tamano et al., 2000). The T3SS apparatus is a multiprotein complex that spans the inner and outer bacterial membrane and extends from the surface forming a needle-like structure. T3SSs appear to have specifically evolved to direct translocation of effector proteins from the bacterial to the host cell cytoplasm. These effector proteins hijack host signal transduction pathways by mimicking host cellular proteins (Galan and Collmer, 1999).

The T3SS needle complex is composed of two distinct parts: (i) a cylindrical basal body that spans both the inner and outer membrane of the bacteria and (ii) a hollow extracellular needle-like structure. The needle complex basal body is composed of three parts: (i) an inner membrane ring, (ii) an outer membrane ring, and (iii) an inner-outer ring connector. In *E. coli*, along with EscV as the main component, EscR, EscS, EscT, and EscU constitute the inner membrane ring structure (Garmendia et al., 2005; Gauthier et al., 2003). EscC is a homo-multimeric secretin that forms the outer membrane ring structure and is involved in the transport of proteins across the outer membrane (Hueck, 1998). The lipoprotein EscJ in the periplasm connects the inner and outer membrane ring structures (Crepin et al., 2005). The needle structure is composed of polymerized EscF subunits (Wilson et al., 2001). Unique to EPEC/EHEC, the needle structure is extended by the hollow EspA filament (Daniell et al., 2001b; Knutton et al., 1998; Sekiya et al., 2001). The hollow tube formed by EspA has an outer diameter of 120 Å and a central channel diameter of 25 Å (Daniell et al., 2003). The effector proteins are translocated into the host cell through a pore in the plasma membrane formed by EspB and EspD via

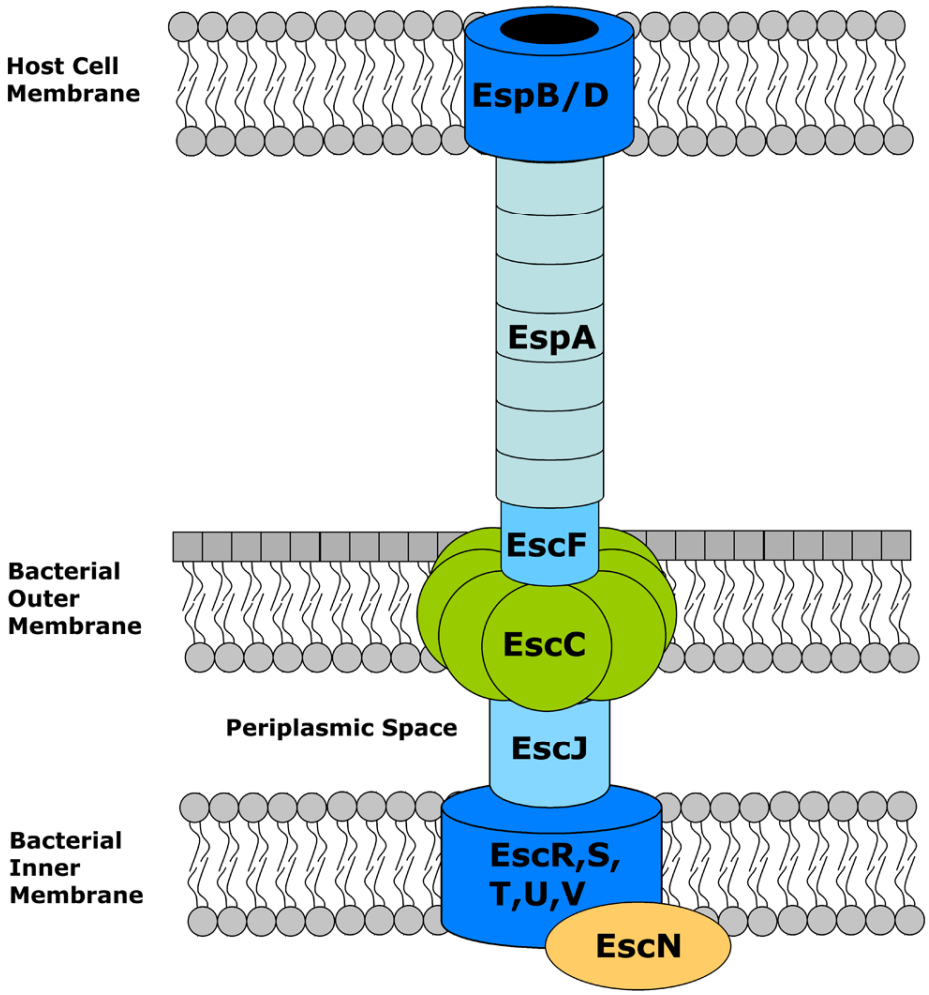


Figure 1.1: Schematic representation of the EHEC/EPEC type three secretion system

The type III secretion system spans both the inner and outer bacterial membrane. The secretin EscC multimerizes to form the outer membrane ring structure. The inner membrane ring structure is constituted of EscR,S,T,U, and V. The lipoprotein EscJ connects the outer to the inner membrane ring structures. The EscF polymerizes to form the hollow needle structure. The EspA filament extends the needle structure. The EspB and EspD create a pore in the host cell membrane and allow the EspA filament to establish a connection between the bacterium and the host cell. In the bacterial cytoplasmic side, the ATPase EscN provides energy to the system.

the narrow channel of the EscF/EspA filament (Buttner and Bonas, 2002; Daniell et al., 2003; Ide et al., 2001; Kresse et al., 1999; Lai et al., 1997). EspB and EspD have also been speculated to function as the capping protein of the EspA filament (Daniell et al., 2001a; Delahay et al., 1999; Fivaz and van der Goot, 1999; Hartland et al., 2000). On the bacterial cytoplasmic side, an essential component of the T3SS is the ATPase EscN. EscN resembles the catalytic β subunit of the F_0F_1-ATPases, sharing amino acid sequence similarity and providing energy to the system (Dreyfus et al., 1993; Eichelberg et al., 1994; Fan and Macnab, 1996; Woestyn et al., 1994) (Figure 1.1).

1.3.2 The chaperones

Type III secretion systems utilize a precise regulatory mechanism in directing the translocation of a small number of effectors substrates to the host cell. Type III secretion chaperones constitute a class of auxiliary proteins that are important for the efficient translocation and secretion of the effectors. In the absence of chaperones, bacterial cognate effectors are secreted less efficiently though the T3SS or not secreted at all. T3SS chaperones are small, acidic homodimers that bind to N-terminal regions of one or a few effectors at a time (Feldman and Cornelis, 2003). Although they have low amino acid sequence similarity, recent studies have revealed structural homology between the different chaperones (Lilic et al., 2006; Stebbins and Galan, 2003). Five of the EHEC/EPEC chaperones have been characterized thus far: CesAB, CesD, CesD2, CesF and CesT. The chaperones bind to a specific effector in the bacterial cytosol and promote the translocation of the effector while remaining in the bacterium.

The T3SS chaperones act as stabilizing, anti-folding, and T3SS targeting factors. First, chaperones have been shown to stabilize their cognate effectors in the bacterial cytosol prior to secretion. For example, in *Yersinia*, the YopE effector is unstable and rapidly degraded in the absence of its chaperone SycE (Frithz-Lindsten et al., 1995). Given the secretion channel is between 2 to 3 nm and the size of effectors ranges from 15 to 80 kDa or more, the effectors must be in a non-globular state to be secreted (Blocker et al., 2001; Kubori et al., 1998; Sekiya et al., 2001). Structural studies have shown that chaperones keep their cognate effectors in an extended, non-globular conformation (Birtalan et al., 2002; Phan et al., 2004; Schubot et al., 2005; Stebbins and Galan, 2001). The chaperone maintains its effector in a secretion-competent state. Lastly, the targeting of effector to the T3SS is highly governed by the chaperones. The chaperones direct the effectors to the T3SS apparatus. The removal of an effector's chaperone binding domain inhibits secretion by the pathogenic T3SS, but not the flagellar T3SS (Lee and Galan, 2004; Lilic et al., 2006). In addition, the chaperones prevent the premature interaction between effectors and the T3SS components (Woestyn et al., 1996).

1.3.3 The effector proteins

Pathogenic *E. coli* first adheres to epithelial cells through an unidentified adhesin. The type III secretion system is subsequently activated to construct the needle complex that injects the effector molecules into the host. Several effector molecules have been identified and characterized. The LEE-encoded bacterial outer membrane protein, intimin, facilitates the attachment of bacteria to the host cell (Frankel et al., 1994). It is

exported to the periplasm by the general secretory pathway. The LEE-encoded type III secreted protein Tir (transmembrane intimin receptor) is inserted into the host cell membrane and serves as the receptor for intimin. Tir's cytoplasmic domain is involved in the recruitment of actin-rich pedestal structures beneath the adherent bacteria in a Cdc42-independent manner (Kalman et al., 1999). The LEE-encoded effectors are Tir, EspF, EspG, EspH, Map, SepZ, and EspB. Little is known about the role of EspH in infection. It localizes to the host cell membrane and inhibits the formation of filopodia while promotes the formation of actin pedestal (Tu et al., 2003). SepZ's translocation is not associated with a specific phenotype, and its function remains unknown (Kanack et al., 2005).

Map, or mitochondrion-associated protein, affects mitochondrial function, causes filopodia formation, disrupts intestinal barrier integrity, and alters tight junction structure (Dean and Kenny, 2004; Kenny et al., 2002; Kenny and Jepson, 2000). Once translocated into the host cytoplasm, Map is targeted to mitochondria via its N-terminal mitochondrial targeting sequence (Kenny and Jepson, 2000). Map disrupts mitochondrial membrane potential, resulting in mitochondrial damage, and it mediates small GTPase Cdc42-dependent filopodia formation, which in turn inhibit Tir-Intimin pedestal formation (Kenny et al., 2002).

EspF causes apoptosis and induces redistribution of occludin, a tight junction associated protein that disrupts the host intestinal barrier function (McNamara et al., 2001). EspF contains a proline rich C-terminal region with three proline rich repeats in EPEC and four in EHEC, indicative of a role in protein-protein interaction (McNamara and Donnenberg, 1998; Viswanathan et al., 2004). Similar to Map, EspF contains N-

terminal mitochondrial targeting sequence and plays a role in mitochondrial membrane permeabilization. However, independent of mitochondrial targeting, Map and EspF independently disrupt intestinal barrier integrity and are essential for intestinal barrier dysfunction (Dean and Kenny, 2004). Map induces the release of cytochrome c from mitochondria into the cytosol and the cleavage of caspase-9 and caspase-3, suggesting an important role of EspF in triggering the mitochondrial death pathway (Nagai et al., 2005; Nougayrede and Donnenberg, 2004).

EspG induces membrane ruffling through the destabilization of microtubules and triggers actin stress fiber formation (Elliott et al., 2001). EspG is homologous to the *Shigella* effector VirA and able to complement *Shigella* VirA mutants during infection. VirA interacts with tubulin and helps trigger host microtubule destabilization (Elliott et al., 2001; Yoshida et al., 2002). Similarly to VirA, EspG has been shown to interact with tubulin and cause microtubule depolymerization, which consequently causes actin stress fiber formation by an undefined mechanism (Hardwidge et al., 2005).

EspB is localized in both the host cell membrane and cytosol. In addition to EspB's role in the formation of the translocation pore and the capping of the EspA filament, the cytosolic portion of EspB also modulates the host cytoskeleton by interacting with α-catenin. The N-terminal region of EspB recruits α-catenin to the EHEC/EPEC adherence site, and this interaction is important for A/E lesion formation. Cells transfected with the N-terminal region or full-length EspB demonstrate a decrease in F-actin accumulation at the site of bacterial adherence (Kodama et al., 2002; Wolff et al., 1998).

Several non-LEE-encoded effectors have been identified to be substrates of the T3SS. There are four that have been characterized: Cif, EspI (NleA), EspFU/TccP, and EspJ. These effectors are carried on small pathogenicity islets or prophages (Campellone et al., 2004; Dahan et al., 2005; Garmendia et al., 2004; Marches et al., 2003; Mundy et al., 2004b). EspI is encoded by prophage CP-933P (Chaudhuri et al., 2004) and co-localized with golgi (Gruenheid et al., 2004; Mundy et al., 2004a). Both EspFU/TccP (Tir-cytoskeleton coupling protein) and EspJ are located in prophage CP-933U. EspFU/TccP is a Nck-like protein that stimulates actin polymerization by activating N-WASP in the presence of Arp2/3 (Garmendia et al., 2004). EspJ has been shown to inhibit macrophage opsono-phagocytosis by both EHEC and EPEC (Marches et al., 2008). Lastly, Cif (cycle inhibiting factor), a cyclomodulin, will be described in the next section.

1.4 Cycle inhibiting factor

One substrate of the EPEC and EHEC T3SS has been shown to inhibit cell cycle progression. The cycle inhibiting factor (Cif) is an effector protein encoded by a lambdoid prophage that is incorporated into the bacterial genome (Marches et al., 2003); Loukiadis et al., 2008). The 32 kDa Cif is a modular protein composed of an N-terminal secretion and translocation signal and a C-terminal domain with effector activity in the host cell (Charpentier and Oswald, 2004). Its chaperone has yet to be identified. Upon translocation into the host cell, Cif induces a cytopathic effect (CPE) characterized by cell cycle arrest at the G_2/M transition associated with the formation of stress fibers and

focal adhesion recruitment in HeLa cells and intestinal epithelial cells (Caco-2, IEC-6) (Nougayrede et al., 2001; Taieb et al., 2006). The inhibition of stress fiber formation by Rho inhibitors does not prevent accumulation of G_2-arrested cells, implying that the cell cycle inhibition induced by Cif is not a product of cytoskeletal alterations (Nougayrede et al., 2001). Moreover, the lipid mediated delivery of purified Cif protein into HeLa cells results in cells displaying CPE phenotype, suggesting Cif is necessary and sufficient for the CPE (Taieb et al., 2006).

Some EHEC/EPEC strains have been identified to carry mutated *cif* with a truncated C-terminus or no *cif* at all. For example, the prototype O157:H7 EHEC strain does not contain *cif*. HeLa cells infected with strains lacking *cif* or carrying the *cif* C-terminal mutant (with the last 10 amino acids deleted) do not display Cif-induced phenotype, which indicates that the last 10 amino acids are essential for Cif's cytopathic activity. When *cif* mutant strains are complemented with a *cif*-containing plasmid, they are able to restore the Cif-induced phenotype in the host cell (Marches et al., 2003). Although the *cif* gene is not present in all EHEC and EPEC strains, recent work has revealed that the *cif*-encoding lambdoid prophage from *E. coli* E22 is capable of propagating to new strains by lysogenic conversion *in vitro* (Loukiadis et al., 2008).

Interestingly, Cif-infected epithelial cells share phenotypic homologies to cells treated with cytolethal distending toxin (CDT). CDT is a secreted tripartite holotoxin composed of three subunits: CDT A, B, and C. CDT-B is the active subunit that is shown to possess DNaseI activity and shares structural homology with DNaseI (Lara-Tejero and Galan, 2000; Nesic et al., 2004; Stebbins, 2004). Both CDT and Cif-induced cell cycle arrest occur with the sustained inhibitory phosphorylation of the mitosis inducer CDK1

(Comayras et al., 1997; Cortes-Bratti et al., 1999; Marches et al., 2003; Whitehouse et al., 1998). However, Cif has no sequence homology to CDT-B or DNaseI, and there is no evidence that it can induce DNA double-strand breaks. Unlike CDT, Cif does not activate the upstream effectors of the G_2 checkpoint, the kinases ATM and ATR (Taieb et al., 2006). Both ATM and ATR function in cell cycle checkpoint pathways activated by DNA damage from double-strand breaks and DNA replication defects, respectively (reviewed in (McGowan and Russell, 2004)). Therefore Cif is thought to hijack the mitotic checkpoint without inducing DNA damage in an ATM/ATR independent pathway. The Cif pathway leading to cell cycle inhibition remains to be elucidated.

1.5 Cell cycle regulation

The eukaryotic cell replication is an intricate process that is divided into four distinct phases: gap phase 1 (G_1), DNA synthesis (S), gap phase 2 (G_2), and mitosis (M). DNA is replicated in the S phase, and cell division occurs in the M phase. The S and M phases are separated by G_1 and G_2. The G_1 phase describes the period after cytokinesis and before the S phase. The G_2 phase occurs after the S phase and lasts until the cell enters mitosis. The two gap phases serve an important function besides allowing the cells to grow between the S and M phases; they monitor the external and internal environments for suitable cell division conditions. The G_1 checkpoint (or the restriction point) ensures that the cell is large enough to divide with proper nutrients to support the daughter cells. The G_2 checkpoint verifies that DNA replication is complete and free of DNA damage.

If the conditions are unfavorable, the G_1/S and G_2/M checkpoints will prevent cell cycle progression at the specific point until the checkpoint requirements have been met.

1.5.1 Proteolysis in cell cycle regulation

Eukaryotic cells have evolved a complex control system that regulates the progression through each stage of the cell cycle. The central components of this regulation system consist of the cyclin-dependent protein kinases (Cdks). The activity of Cdks is regulated by several mechanisms such as association with regulator cyclins, interaction with Cdk inhibitors, phosphorylation and other post-translational modifications (Sherr, 1996; Sherr and Roberts, 1999). The amount of Cdk regulators and inhibitors oscillate at various stages of the cell cycle as a result of proteolysis (Murray, 2004). Most regulated proteolysis can be attributed to the ubiquitin-proteasome system. The proteins are marked for degradation by ubiquitylation and then targeted to the 26S proteasome complex (Pickart and Cohen, 2004; Weissman, 2001).

The ubiquitylation of a protein requires at least three enzymes: a ubiquitin-activating enzyme (E1), a ubiquitin-conjugating enzyme (E2), and a ubiquitin ligase (E3) (reviewed in (Passmore and Barford, 2004)). Ubiquitin is activated at its COOH-terminus via a thioester linkage to E1. Activated ubiquitin is then transferred to E2. E2 interacts with E3, whose function is to bring E2 and the substrate together (reviewed in (Scheffner et al., 1995)). E2 can either transfer ubiquitin directly to the lysine residue of the substrate, as in the case of the SCF complex (Deshaies, 1999), or it can transfer ubiquitin to E3, as observed in the HECT-domain E3s (Kee et al., 2005). The E3

components are primary responsible for substrate recognition. Two major types of E3 ligases are involved in cell cycle proteolysis: SCF (Skp1/Cullin/F-box protein) and APC/C (anaphase-promoting complex or cyclosome) (Deshaies, 1999; Harper et al., 2002; Murray, 2004; Nakayama et al., 2001; Seol et al., 1999). APC/C is activated during mitosis and G_1 to degrade proteins that impede mitotic progression and are deleterious to G_1 progression if accumulated (Bai et al., 1996; Cohen-Fix et al., 1996; Funabiki et al., 1996; King et al., 1995; Sudakin et al., 1995). The SCF complex was initially thought to be involved only in G_1-S transition, but there has been increasing evidence of the SCF complex involvement in other cell cycle phases (Bai et al., 1996; Feldman et al., 1997; Skowyra et al., 1997).

The multisubunit SCF complex is comprised of the three invariable subunits Skp1, Cullin-1 (Cul1), and RING H2 protein Rbx1 (also know as Roc1 or Hrt1) as well as an interchangeable component, the F-box protein (Figure 1.2). Cul1 serves as the rigid elongated scaffold that interacts with both Skp1 and Rbx1 subunits (Wu et al., 2003a; Zheng et al., 2002b). Cul1's N-terminus binds the Skp1 adaptor protein and the C-terminus binds Rbx1. Cul1 and Rbx1 form the catalytic core that recruits the to the SCF complex. Skp1 is the adaptor protein that binds to different F-box proteins through the F-box motif. The F-box protein is responsible for substrate recognition, and it orients the substrate for ubiquitylation by E2 of the SCF complex. The SCF complex is able to interact with a distinct set of substrates mainly due to the variable F-box proteins (Cardozo and Pagano, 2004). Skp2 is a well-characterized F-box protein of the SCF complex that participates in both G_1-S and G_2-M transition. Skp2 contains a binding site that recognizes specific amino acid sequences of phosphorylated

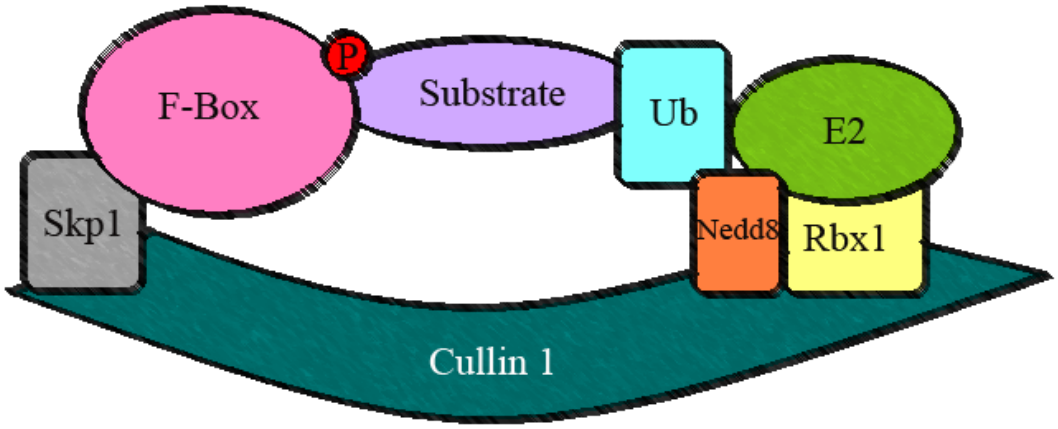


Figure 1.2: A depiction of the SCF complex with its substrate

The SCF complex has Cullin as its base. Cullin is a rigid elongated scaffold with Rbx1 attached to its C-terminal domain. Nedd8 is conjugated to a conserved lysine residue of Cullin1. Nedd8 and Rbx1 assist the recruitment of E2 and ubiquitin. They also position E2 in close proximity to the substrate. The F-box protein is anchored to the N-terminal domain of Cullin1 via the adapter Skp1. The F-box protein recognizes the phosphorylated substrate and orients the substrate to E2 for ubiquitylation.

substrates (Tsvetkov et al., 1999). SCF complex activation requires the covalent conjugation of the small ubiquitin-like protein Nedd8 to the conserved lysine residue of the Cul1 (Hori et al., 1999). The neddylation of Cul1 increases the E3 ligase activity through Nedd8's facilitated recruitment of the E2 (Furukawa et al., 2000; Morimoto et al., 2000; Read et al., 2000; Wu et al., 2000) and the displacement of the SCF inhibitor CAND1 (Liu et al., 2002; Zheng et al., 2002a).

The SCF complex catalyzes polyubiquitylation of specific cell cycle regulators. These polyubiquitylated proteins are immediately recognized and degraded by the 26S proteasome. Removal of the cell cycle regulators triggers several critical events in the cycle. For example, the Cdk inhibitors, p27 and p21, suppress G_1/S-Cdk and S-Cdk activities. Accumulation of the two proteins prevents cells from entering the S phase. The SCF complex plays a crucial role in the degradation of the p27 and p21 to promoter cell progression into S phase (Bornstein et al., 2003; Carrano et al., 1999; Sutterluty et al., 1999; Tsvetkov et al., 1999; Yu et al., 1998). The SCF complex is involved in the G_2 to M phase transition by acting on Wee1, a kinase that fine-tunes the Cdk activation during entry to mitosis (Kellogg, 2003; Michael and Newport, 1998; Moshe et al., 2004). Cdk1 must be dephosphorylated by Cdc25 phosphatases in order to promote entrance to M phase. Wee1 has an inhibitory effect on Cdk1, therefore, inactivation of Wee1 shifts the balance to Cdk1 dephosphorylation. The SCF complex's effect is not limited to cell cycle regulation. Approximately 70 F-box proteins have been identified, which provides the SCF's versatility targeting proteins outside of the cell cycle as well (Jin et al., 2004).

1.6 Aim of the thesis

Regulation of the eukaryotic cell cycle involves a complex meshwork of proteins that instruct the cell to grow and replicate or to senesce. Cif is a T3SS cyclomodulin from pathogenic *Escherichia coli* that is capable of deregulating the host's cell cycle. The mechanisms underlying Cif mediated cell cycle arrest and recruitment of focal adhesions has yet to be established. Cif has no sequence homology to any known

proteins. In order to elucidate Cif's enzymatic mechanism during host cell infection, structure-function studies of Cif were performed using x-ray crystallography in combination with biochemical, cellular, and genetic assays. The structural investigation of Cif could provide necessary information on its mechanism of action as well as shed some light on the intricacies of cell cycle regulation. The first section of the thesis describes Cif's 3D structure and function. The subsequent part focuses on the identification of Cif's host cell target.

CHAPTER TWO:

MATERIALS AND METHODS

2.1 Mutagenesis of the putative catalytic triad

Amino acid substitutions were introduced into the *cif* gene by *in vitro* site-directed mutagenesis using oligos containing specific base changes. The mutant plasmids were amplified by PCR using Pfu Turbo DNA polymerase (Stratagene). The template plasmid was removed by digestion with DpnI prior to transformation. The Cif full-length point mutants were sequence verified and transformed into specified *E. coli* strains.

2.2 Infection assay

293T cells were synchronized at G_1/S phase by double thymidine block (O'Conner and Jackman, 1995). 293T cells were infected with *E. coli* strains transformed with either wild type or mutant *cif* carrying plasmid at a multiplicity of infection (MOI) of 100, under the arabinose inducible promoter. After 30 minutes post addition of the bacteria to 293T, 0.2% of arabinose was added to the medium to induce Cif expression in the bacteria. The bacterial infection stage required the medium to be free of FBS. 4 hours post infection, 293T cells were washed with 1X PBS, added with fresh DMEM containing 10% FBS, 100μg/ml gentamycin (Gibco), and incubated for an additional 24 hours. 293T cells were harvested and fixed in 70% ethanol. The fixed cells were stained

with propidium iodide and sorted by FACScalibur flow cytometer (BD Biosciences). Data from at least 10,000 cells were analyzed by Cell Quest Pro (BD Biosciences).

The infection assay using HeLa cells was performed by Gregory Jubelin and Eric Oswald as followed. Bacterial strains were cultured overnight in Luria-Bertani (LB) broth then subcultured 1:100 in interaction medium (DMEM with 25mM Hepes and 5% fetal calf serum (FCS)) for 3 hours at 37°C. HeLa cells (ATCC CCL-2), synchronized at the G_1/S phase transition by a double thymidine block (O'Conner and Jackman, 1995), were washed and infected in interaction medium with MOI of 100:1. Following 90 minutes infection, the cells were washed and cultivated for the indicated times in DMEM medium supplemented with 10% FCS and 200 μg/ml gentamicin.

2.3 Secretion Assay

Overnight static culture of transformed *E. coli* grown in DMEM was used to inoculate DMEM at a ratio of 1:50. 0.2% arabinose was added to the medium to induce the expression of Cif Flag tagged protein. The bacteria were grown for 6 hours at 37°C, 5% CO_2, statically. The bacteria were spun down and the supernatant was passed though a 0.45um filter. The type III translocated proteins in the supernatant were precipitated with 10% trichloroacetic acid and then incubated on ice for 1 hour. The precipitated proteins were washed twice with 100% acetone and resuspended in 3X SDS loading buffer. The bacterial pellet fraction and the precipitated supernatant proteins were separated by SDS-PAGE and analyzed by immunoblotting. Cif was detected by mouse anti-Flag (Sigma).

2.4 BioPORTER® assay

cif WT and point mutants were cloned into a modified pET21a vector and transformed into BL21 (DE3). They were grown and purified using a Ni-NTA gravity column (QIAGEN) in a similar method as pGEX4T3-Cif. The bacteria were lysed in lysis buffer C (50mM Tris pH 8.0, 500mM NaCl, and 5mM imidazole pH 8.0), washed in wash buffer C (50mM Tris, pH 8.0, 500mM NaCl, and 30mM imidazole, pH 8.0), and eluted in elution buffer C (50mM Tris pH8.0, 500mM NaCl, and 250mM imidazole pH 8.0). Following the 3C removal of the hexahistidine tag, Cif was purified by Source 15Q sepharose and Superdex 200 gel filtration chromatography. The purified Cif was dialyzed into PBS. BioPORTER® (Genlantis) assays were performed by Gregory Jubelin and Eric Oswald with purified Cif proteins (WT and mutants) as described previously (Taieb et al., 2006). When indicated, E64 (Sigma) was used at 70µM and preincubated 10 min with purified Cif (WT or mutant) before mixing with BioPORTER®.

2.5 Stress fiber formation and cell cycle analysis

This section was performed by Gregory Jubelin and Eric Oswald. Study of cell morphology and visualization of cytoskeleton was done as described (Taieb et al., 2006); briefly, HeLa cells in chamber slides were fixed for 15 min in 4% formaldehyde, permeabilized in 0.1% Triton X-100 and then incubated with rhodamine-phalloidin (Molecular Probes) to stain F-actin and with DAPI (Sigma) to stain DNA. Images were acquired with a DMRB fluorescence microscope equipped with a DFC300FX digital

camera (Leica). Cell cycle distribution analyses were performed using a FACScalibur flow cytometer (Becton Dickinson) as described previously (Sert et al., 1999). Data from at least 10,000 cells were analyzed using FloJo software v8.5 (Tree Star).

2.6 *In vitro* protease assay

The BODIPY FL casein proteolytic assay was performed as suggested by the manufacturer (Invitrogen). Five or ten micrograms of Cif, active site mutants, various N-terminal deletion mutants, and protease controls (papain and subtilisin) were added to BODIPY FL casein and incubated at room temperature, protected from light, for at least 1 hour in a digestion buffer consisting of 10mM MES, pH 6.5. For the examination of Cif loop mutant, Cif WT, loop mutant, and protease control papain were added to BODIPY-FL casein in different pH. The detection of cleavage products was followed by fluorescence (excitation/emission: 485/530nm). HeLa extract (200µg) was also included in the assay buffer in some trials as indicated.

2.7 Auto-radiolabelling assay

Ten microgram of Cif WT or mutants was incubated with 2µl of ^{14}C-acetyl-CoA (50μCi/ml, GE Healthcare) and acetylation buffer (50mM Hepes pH 8.0, 10% glycerol, 1mM DTT, and 1mM PMSF) to a total volume of 30µl at 30°C for 1 hour. The reaction was terminated by spotting the samples onto p81 filter papers (Millipore). The filter papers were dried completely before washing 4 times with 125mL wash buffer (50mM

$NaHCO_3/Na_2CO_3$ pH 8.0) for 10 minutes. 4 mL of Readysafe® scintillation fluid was added to the dried filter paper in a scintillation vial. The ^{14}C-labeling was measured by scintillation counting. 50-100μg of HeLa cell extract was included in some assays.

2.8 Acetyltransferase assays and mass spectrometry analysis

Each of the purified recombinant Cif WT (5mg, Cif1-282 theoretical molecular weight 32122 Da), Cif C109S (5mg, Cif1-282-His theoretical molecular weight 33441 Da), and Cif triple mutant (C109S/H165A/D187A, 2mg, Cif1-282 theoretical molecular weight 31995 Da) was incubated at 30°C for 60 minutes in acetylation buffer (20mM Tris pH 8.0, 200mM NaCl, 2mM DTT, and 200μM acetyl-CoA (Sigma)). To check if Nedd8 was modified by Cif, purified co-expressed Cif WT and Nedd8 complex (5mg complex, Nedd8 theoretical molecular weight 9073 Da) was incubated at 30°C for 60 minutes in the acetylation buffer. The reaction was stopped by flash freezing the sample in liquid nitrogen. Proteins were analyzed by electrospray ionization mass spectrometry (ESMS). Briefly, samples were desalted using C-4 ZipTip (Millipore) and eluted from the ZipTip using 50% acetonitrile/0.2% formic acid. The samples were introduced into a Q-Tof Micro (Waters Corporation) using the nanospray technique. The data acquisition and spectrum processing were performed using the MassLynx 4.0 software provided by the manufacturer. 50-100 spectra were acquired and combined. The mass of the protein was determined from the combined spectrum using the MaxEnt 1 algorithm and/or by software, which calculated the mass after two multiply charged ions differing by one charge were manually identified.

2.9 Nedd8 binding assays

293T cells were transfected with pFLAG-CMV-Nedd8 using GenePORTER2 Transfection Reagent (Genlantis) following the manufacturer's protocol. 48 hours post transfection, 293T cells were lysed in RIPA buffer (50mM Tris pH 8.0, 250mM NaCl, 1.0% NP-40, 0.5% DOC, 0.1% SDS) in 3 freeze thaw cycles. Cif C109S was cloned into the pGEX4T3 vector and expressed in BL21(D3E) cells. First, GST-Cif C109S was purified from the bacterial lysate by glutathione sepharose (GE Healthcare). 100μg of purified GST-Cif C109S was bound to pre-equilibrated Glutathione Sepharose 4B (GE Healthcare) in equilibrating buffer (20mM Hepes pH 7.5, 5mM $MgCl_2$, 100mM NaCl, and 2mM DTT) at 4°C for an hour on a rotator. Either 100μg or 250μg of 293T lysate was added to the GST-Cif C109S bounded glutathione beads and further incubated at 4°C for 2 hours on a rotator. The glutathione beads were washed once with equilibrating buffer, twice with wash buffer (20mM Hepes pH 7.5, 300mM NaCl, 5mM $MgCl_2$, 0.1% NP-40, and 2mM DTT), and lastly once with equilibrating buffer. 40μl of 3X SDS loading buffer were added to the protein bound beads, and the beads were boiled at 95°C for 5 minutes. The samples containing only the supernatant were separated by SDS-PAGE and analyzed by immunoblotting. FLAG-Nedd8 was detected by mouse anti-FLAG monoclonal antibody (Sigma).

2.10 Bacterial co-expression of Cif and Nedd8

The *nedd8* gene was subcloned from a plasmid containing the *nedd8* full-length cDNA clone (RZPD German Resource Center for Genome Research). *cif* 23-282 and *nedd8* were cloned into a modified CDFDuet-1 vector that contained two multiple cloning sites (MCS) (Novagen). The MCS1 of the engineered CDFDuet-1 has an N-terminal hexahistidine with a 3C cleavage site that allows for N-terminal tag removal. *cif* 23-282 was cloned into the MCS1 (SalI, NotI), and *nedd8* was cloned into MCS2 (HindIII, XhoI). pDuet-Cif23-282-Nedd8 was transformed into BL21(D3E) cells. Cif23-282 and Nedd8 were expressed in a manner similar to pET21a-Cif100-282-His. The bacteria were lysed by passage through an Emulsiflex-C5 (Avestin, Inc.). His-Cif23-282 and Nedd8 were co-purified by hexahistidine tag binding to chelating sepharose and finally by gel filtration chromatography (buffer: 20mM Tris pH 8.0, 200mM NaCl, and 2mM DDT) and analyzed by SDS-PAGE. Both pDuet-Cif23-99-Nedd8 and pDuet-Cif100-282-Nedd8 were purified similarly to that of pDuet-Cif23-282-Nedd8.

2.11 Autoproteolysis assay

To test the AvrPphB activation model, 50μg of Cif WT or Cif C109S was incubated with or without 5mM PMSF or 2μM E64 in buffer (20mM Tris pH 8.0, 200mM NaCl, and 2mM DT) for 12 hours at 37°C. The samples were analyzed by SDS-PAGE and visualized by coomassie blue staining. To test the AvrRpt2 activation model,

Cif WT, Cif C109S, and Cif 100-282 were tagged at both N- and C-termini with hexahistidine. 2μg of Cif protein was incubated with 2μl of rabbit reticulocyte lysate (Promega) or 200μg of yeast extract in 20mM Tris pH 8.0, 50mM NaCl, and 2mM $MgCl_2$ for 3 hours at 30°C. The samples were analyzed by immunoblotting using anti-His_6-HRP (1:1,000, Sigma).

2.12 SDS-PAGE analysis and immunoblotting

Protein samples were boiled at 95°C in 3X SDS-PAGE loading dye. Samples were separated by 15-18% SDS-PAGE gel. For immunoblotting analysis, the protein bands were transferred from the SDS-PAGE gel to nitrocellulose membrane either overnight at 60 milliamp or 1 hour at 100 volt in Tris-glycine buffer (25mM Tris, 190mM glycine, 0.1% SDS, 20% methanol, pH 8.5). The membrane was first blocked with 1X TBS-T (20mM Tris pH 8.0, 150mM NaCl, and 0.05% Tween-20) plus 5% milk for 30 minutes at room temperature and then washed 3 times, 10 minutes each with 1X TBS-T. Primary antibody, diluted in 1X TBS-T plus 5% milk at the concentration suggested by the manufacturer, was added to the membrane and incubated for 1 hour in room temperature with shaking. The membrane was then washed 3 times with 1X TBS-T for 10 minutes each time. The secondary antibody was diluted in 1X TBS-T with 5% milk and added to the membrane for 1 hour in room temperature. The membrane was washed 3 times with 1X TBS-T for 10 minutes each time. The ECL reagent (GE Healthcare) was added to the membrane for signal detection.

2.13 Purification of APPBP1/UBA3, Ubc12, Cullin-1/Rbx1, and Nedd8

Bacterial expression plasmids of Cullin-1/Rbx1, Ubc12, APPBP1/UBA3, and Nedd8 were generous gifts from B. Schulman. Cullin-1 and Rbx1 were purified according to experimental method described in (Li et al., 2005). APPBP1-UBA3 was purified as described in (Huang and Schulman, 2005).

pGEX4T3-Ubc12 was grown and harvested in a similar method as the modified pGEX4T3-Cif. GST-Ubc12 was first purified by glutathione sepharose gravity column using the same method and reagents as GST-Cif purification. GST-Ubc12 did not bind to glutathione sepharose efficiently, thus the flow through and wash fractions were reloaded onto the column to maximize the yield. Thrombin (1:50 wt/wt) was added to GST-Ubc12 for the GST tag removal and incubated for at least 16 hours at 4°C with buffer T (50mM Tris pH 8.0, 200mM NaCl, 10mM DTT, and 5mM $CaCl_2$). It was important to dialyze out the glutathione due to its inhibitory effect on the thrombin cleavage. The cleaved Ubc12 was diluted to 50mM NaCl and loaded onto Source 15S cationic exchange chromatography to remove thrombin. Ubc12 and uncleaved GST-Ubc12 eluted into the flow through and wash fractions. These fractions were then loaded onto Source 15Q anionic exchange chromatography. While the cleaved GST tag and GST-Ubc12 bound to the column, the Ubc12 eluted into the flow through and wash fractions. Ubc12 was further purified using gel filtration chromatography and stored in a buffer containing 25mM Tris pH 8.0, 150mM NaCl, and 5mM DTT.

The mature conjugable Nedd8 (without the C-terminus following the Gly-Gly dipeptide) was encoded in a pGEX4T3 plasmid. GST-Nedd8 was grown, harvested, and

purified with a glutathione sepharose gravity column in a similar experimental protocol as GST-Cif. GST-Nedd8 was diluted to 50mM NaCl from 200mM NaCl and loaded onto Source 15Q anionic exchange chromatography. GST-Nedd8 was bound to the column and then eluted with increasing salt concentration. Thrombin (1:100 wt/wt) was added to GST-Nedd8 for the GST tag removal and incubated in the elution buffer with 5mM $CaCl_2$ at 4°C for at least 16 hours. The sample was further purified by gel filtration chromatography. The sample eluted in two peaks, with one peak containing GST-Nedd8 and the other containing only Nedd8. Nedd8 was stored in 20mM Tris pH 8.0, 200mM NaCl, and 2mM DTT buffer.

2.14 Reconstitution of Neddylated Cullin-1

For a 40µl reaction, 20ng APPBP1/UBA3, 10µg Ubc12, 80pmol (~8.14µg) Cul1/Rbx1, and 15µg Nedd8 are incubated in the neddylation buffer (40mM Tris pH 7.5, 4mM ATP, 12mM $MgCl_2$, and 50mM NaCl) at 37°C for 1 hour. 3X SDS-PAGE loading dye was added to stop the reaction. To test if Cif could disrupt the E1's or E2's activity, 10µg of various Cif constructs and mutants were incubated with the reaction. The samples were separated on a 4-20% Tris-HCL Ready Gel (Bio-Rad) then immunoblotted with rabbit anti-Nedd8 (1:500, Alexis) or mouse anti-Nedd8 (1:5,000, Sigma). Cif cross-reacted with rabbit anti-Nedd8.

To purify the neddylated Cul1/Rbx1 complex, the neddylation reaction was purified by GST pulldown. The GST tag was attached to Rbx1. The neddylated Cul1/Rbx1 complex was incubated with glutathione resin at 4°C, rotating for 1 hour in equilibrating buffer (50mM Tris pH 8.0, 50mM NaCl, and 2mM DTT). The resin was

washed twice with wash buffer (50mM Tris pH 8.0, 200mM NaCl, and 1mM DTT). The elution buffer (50mM Tris pH 8.0, 50mM NaCl, 2mM DTT, and 5mM glutathione) was added to GST-Rbx1/Cul1 bounded resin and incubated while rotating for 4 hours at 4°C. The elution was dialyzed to remove the glutathione (buffer: 20mM Tris pH 8.0, 50mM NaCl, and 1mM DTT) and stored at –80°C.

2.15 Deneddylation assay

Neddylated Cullin-1/Rbx1 was incubated with either purified Cif WT or Cif C109S (20mM Tris pH 8.0, 50mM NaCl, and 1mM DTT) at 37°C for one hour. The samples were separated by SDS-PAGE and analyzed by immunoblotting. Nedd8 was detected by mouse anti-Nedd8 (Sigma).

Nedd8-AMC, Ubiquitin-AMC, or SUMO-AMC (BIOMOL International, LP) at a concentration of 0.5μM was incubated with either 100nM of purified Cif WT, Cif C109S, or various N-terminally truncated Cif in 120μM reaction buffer (50mM Hepes pH 7.5 and 1mM DTT) at room temperature, for 30 minutes in the dark. First the reaction buffer was mixed with Cif before the AMC modified substrate was added to the mixture. The samples were analyzed by spectrophotometer at excitation 380nm and emission 460nm at 1-minute interval for 30 minutes.

CHAPTER THREE:

BIOCHEMISTRY AND CRYSTALLOGRAPHY OF CIF

3.1 Cloning and protein expression

3.1.1 Non-template gene synthesis

The *cif* gene is only present in some EHEC or EPEC genomes, which are not available commercially. Therefore, the 849 base pair long *cif* gene was initially synthesized using assembly PCR (Stemmer et al., 1995). The method requires the presence of a large number of overlapping oligodeoxyribonucleotides that, when assembled, form templates for the gene of interest. The oligos are then subjected to several rounds of PCR to assemble the full-length gene (Figure 3.1A).

The *cif* gene oligos were designed using a web-based algorithm DNAWorks. DNAWorks has several user-defined parameters that allow the user to design a PCR-based gene according to specifications. In designing *cif*, the melting temperature of the designed oligos was set to 58°C and the codon frequency was optimized for the *E. coli* expression system. DNAWorks allows the user to input specific 5' and 3' flanking sequences for directional cloning and to exclude specific restriction sites from the synthesized gene. First, the program reverse translated Cif's amino acid sequence into nucleotide sequence. The sequence was then divided into sections and optimized to

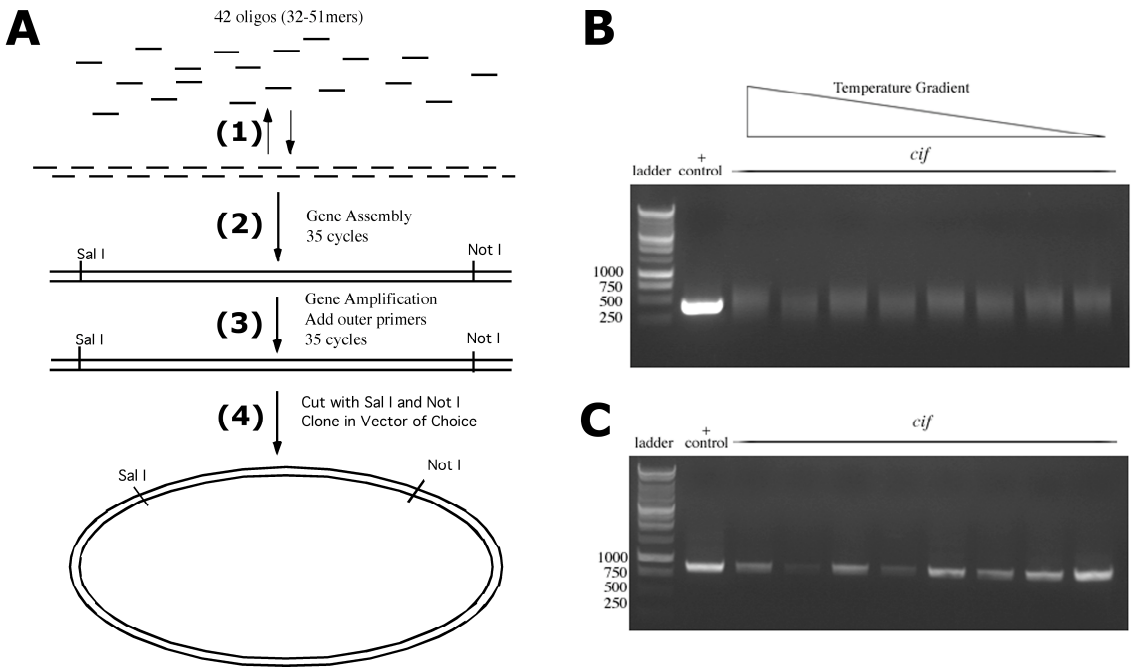


Figure 3.1: Non-template gene synthesis

(A) PCR assembly protocol for synthesis of the *cif* gene.

(1) 42 overlapping oligos, when assembled, form the template for *cif*.

(2) Gene Assembly: the oligos are mixed and subjected to 35 cycles of PCR.

(3) Gene Amplification: the PCR product from the gene assembly step is further amplified with outer primers.

(4) The synthesized *cif* is digested with SalI and NotI and cloned into the vector of choice.

(B) Gene assembly PCR. The 42 oligos were combined in a PCR mixture and subjected to temperature gradient PCR reaction. The *cif* gene appears as a smeary band.

(C) Gene amplification PCR. The PCR products from the gene assembly step are amplified with outer primers. The *cif* gene is shown as one distinct band.

generate oligos that are highly homogeneous in melting temperatures, and the tendency for hairpin formation was also minimized. The synthesized oligos were combined and assembled using a two-step PCR method (Hoover and Lubkowski, 2002).

The *cif* gene having the same amino acid sequence found in EPEC and EHEC but with a different DNA sequence was successfully synthesized by assembly PCR. A total of 42 oligos, ranging in lengths from 32 to 51 bases, were generated. BamHI, HindIII, NotI, and SalI restriction sites were specifically excluded from the DNA sequence for future subcloning purposes. A SalI sequence was included in the 5' flanking sequence, and a NotI sequence was included in the 3' flanking sequence. The mixed oligos were diluted to 5 μM and subjected to the first round of PCR (gene assembly step). Various DNA polymerases, PCR protocols, and PCR machines were tested before the gene was successfully synthesized. The working combination required the BIO-RAD iCycler with temperature gradient function. The reaction was carried out with one unit of high-fidelity Taq polymerase (Invitrogen) and 0.2 mM dNTP (Fisher Scientific) with a temperature gradient during the annealing step. The gene assembly's PCR protocol is initiated with a 'hot start' step and 5 minutes of Taq activation at 95°C. This step was followed by 35 cycles of 30 seconds denaturation at 95°C, 30 seconds annealation with a temperature gradient ranging from 50°C to 60°C, and 1.5 minutes DNA extension at 72°C. Lastly, the PCR reaction was incubated at 72°C for 10 minutes. The PCR product from the gene assembly step appears as smeared bands on the 1% agarose gel (Figure 3.1B).

Gene amplification occurs when outer primers are added to the gene assembly PCR products and are subjected to a second round of PCR (Figure 3.1C). The PCR

protocol for gene amplification was the same as for gene assembly, except the annealing temperature was set to a uniform 58°C. The PCR product resulted in a single band on the agarose gel. The *cif* full-length PCR product was purified from the gel and subcloned into pCR2.1-TOPO vector (Invitrogen). The clones were sequence verified for the DNAWorks-based *cif* sequence. *cif* was subcloned into modified *E. coli* expression vectors pET21a (Novagen) and pGEX4T3 (GE Healthcare) using SalI and NotI sites and then transformed into BL21(DE3) cells. Engineered pET21a has an N-terminal hexahistidine with a rhinovirus 3C protease site that allows for tag removal. The modified pGEX4T3 has an engineered 3C cleavage site for glutathione S-transferase (GST) removal and two hexahistidine tags at the N-terminus.

3.1.2 Protein expression and purification

Ten milliliter overnight cultures of the pGEX4T3-Cif full-length protein were used to inoculate each liter of LB media (EMD Chemicals Inc) containing 100μg/ml ampicillin. Cultures were grown at 37°C to an $O.D_{600}$ of 0.8, then induced with 1mM IPTG and grown for an additional 16 hours at 20°C. Bacterial cells were harvested and resuspended in 10ml lysis buffer A (50mM Tris pH 8.0, 200mM NaCl, 10mM DTT, 1mM PMSF) per liter of original culture and then lysed using an Emulsiflux-C5 (Avestin). Immediately following lysis, 1mM PMSF was added to prevent proteolytic digestion. To digest DNA and decrease viscosity, 0.1mg/ml DNaseI and 5mM $MgCl_2$ were added to the lysate. After 10 minutes incubation on ice, 5mM EDTA was added to

terminate the reaction. The lysate was spun at 16,000 rpm for 45 minutes in order to separate soluble proteins, including Cif, from cell debris and inclusion bodies.

Cif was first purified using a glutathione sepharose gravity column (GE Healthcare). The bacterial lysate was loaded onto the pre-equilibrated column (in lysis buffer A), washed once with 5 column volumes of lysis buffer A and once with 5 column volumes of wash buffer A (50mM Tris pH 8.0, 500mM NaCl, and 10mM DTT). The protein was then eluted with 5 column volumes of elution buffer A (50mM Tris pH 8.0, 200mM NaCl, 10mM DTT, and 5mM glutathione). After the initial purification step, the overall yield of GST-Cif was approximately 1 gram of protein per 2 liters of bacterial culture. The protein sample was dialyzed into 3C digestion buffer (50mM Tris pH 8.0, 200mM NaCl, 5mM DTT, 1mM EDTA) and digested overnight with 3C protease at a 50 to 1 ratio (wt:wt) in order to remove the GST tag. Cif was further purified by anionic exchange chromatography using a 25mL Source 15Q column (GE Healthcare) by FPLC (Åkta). Cif was diluted to 100mM NaCl prior to column loading. Cif eluted off the column at 200mM NaCl. Fractions containing Cif were pooled together and concentrated with Centricon-70 (Millipore) to 45mg/ml. In the final purification step, Cif was injected onto a Superdex 75 gel filtration column and exchanged into crystallization buffer (20mM Tris pH 8.0, 200mM NaCl, and 2mM DTT). The protein concentration was measured by Bradford assay and visualized by coomassie staining on a 15% SDS-PAGE gel.

To determine potential structural domains amenable for crystallization, various Cif constructs were created based on secondary structure predictions and limited proteolysis experiments. Using the PHD secondary structure prediction (Rost et al.,

1994), Cif14-282, Cif57-282 and Cif66-282 were cloned into a modified pGEX4T3 vector and a modified pET21a vector. The purified Cif full-length protein was subjected to limited proteolysis using the non-specific serine protease subtilisin. Subtilisin removes flexible protein regions that may impair crystallization efforts. Subtilisin was added to Cif at a 1 to 10% ratio (wt:wt) and incubated on ice for 20 minutes (Figure 3.2). The reaction was initiated by 5mM $CaCl_2$ and stopped with the addition of 100mM PMSF. The digested samples were separated on a 15% SDS-PAGE gel and transferred onto a PVDF membrane. Protein bands that persisted in the presence of increasing subtilisin were sequenced by Edman degradation at The Rockefeller University Proteomics Resource Center. The subtilisin digested products were identified as fragments starting at residues 31, 93, and 100. Cif31-282, Cif93-282, and Cif100-282 were subcloned into a modified pGEX4T3 vector and a modified pET21a vector. All constructs were transformed into BL21(DE3) cells (Table I).

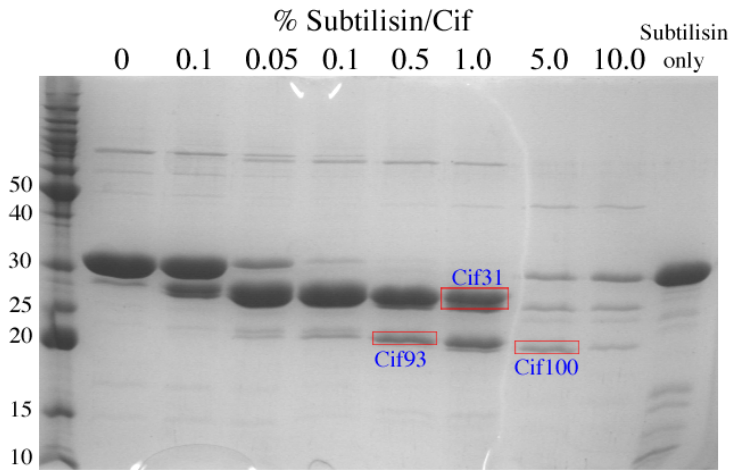
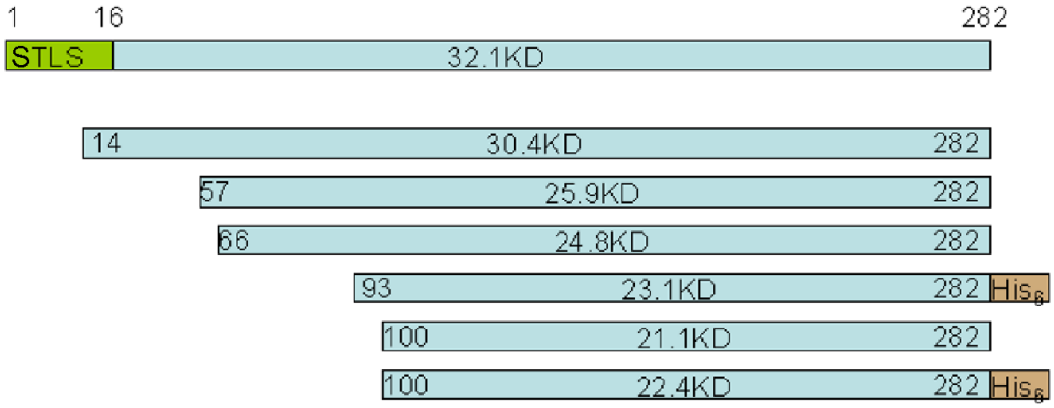


Figure 3.2: Protease footprinting delineates the crystallizable Cif domain
Purified Cif1-282 was incubated with an increasing concentration of subtilisin at 4°C for 20 minutes. The fragments (boxed in red) persisted throughout the proteolytic digestion of Cif. These samples were submitted for N-terminal sequencing and identified as fragments starting at residues 31, 93, and 100.

Table I: Cif constructs, solubility, purification yields and crystallization



Construct	*Vector*	*Expression*	*Solubility*	*Yield*	*Crystals?*
Cif Full length	pGEX4T3	37°C	yes	23mg/L	no
PHD Secondary Structure Prediction:					
Cif 14-282	pGEX4T3	37°C	yes	30mg/L	no
Cif 57-282	pGEX4T3	37°C	insoluble	N/A	N/A
Cif 66-282	pGEX4T3	37°C	insoluble	N/A	N/A
Subt Digest:					
Cif 31-282	pGEX4T3	37°C	yes	30mg/L	no
Cif 93-282 His	pET21a	20°C	insoluble	20mg/L*	no
Cif 100-282	pET21a	20°C	insoluble	25mg/L*	no
Cif 100-282 His	pET21a	20°C	insoluble	25mg/L*	yes

* After refolding

Various Cif constructs were made based on PHD secondary structure prediction and limited proteolysis. Cif100-282-His yielded well diffracted crystals.

As described in Chapter Two, the insoluble constructs were refolded to yield highly soluble protein. Yield indicates quantity of Cif protein purified per liter of *E. coli* culture broth. These data are a representative snap shot of a much larger number of trials with constructs and variables.

N/A = crystallization experiment has not been performed for the given construct.

STLS = secretion and translocation signal (Charpentier and Oswald 2004).

Crystallization trials were performed with the following constructs: Cif full-length, 14-282, 31-282, 93-282-His, 100-282, and 100-282-His. Only Cif100-282-His yielded well-diffracting crystals. pET21a-Cif100-282-His was grown at 37°C to $O.D_{600}$ of 0.8 then induced with 1mM IPTG and grown for an additional 4 hours. The bacteria were harvested and resuspended in 15ml lysis buffer B (8M urea, 10mM Tris pH 8.0, and 0.1M $NaH_2PO_4 \bullet H_2O$) per liter of original culture. Cells were subsequently lysed by sonication. Cif100-282-His was denatured under 8M urea. The lysate was spun at 16,000 rpm for 45 minutes to separate denatured protein from cell debris. The supernatant contained Cif. Cif100-282-His was purified using a nickel fast flow chelating resin gravity column (GE Healthcare). The denatured protein was loaded onto the pre-equilibrated column (in lysis buffer B), washed once with 5 column volumes of lysis buffer B and once with 5 column volumes of wash buffer B (8M urea, 10mM Tris pH 8.0, 0.1M $NaH_2PO_4 \bullet H_2O$, and 30mM Imidazole pH 8.0). The protein was then eluted with 5 column volumes of elution buffer B (8M urea, 10mM Tris pH 8.0, 0.1M $NaH_2PO_4 \bullet H_2O$, and 500mM Imidazole pH 8.0).

Cif100-282-His was refolded into the native state by gradual removal of the denaturing agent through dialysis. Prior to refolding, Cif was diluted in lysis buffer B to 1.5mg/ml, which recovered approximately 95% of the refolded protein. The renaturation step required a total of 4 buffer changes of 4L renaturation buffer (20mM Tris pH 8.0, 200mM NaCl, 5% glycerol, 2mM EDTA, and 5mM DTT) during 24 hours. The last buffer exchange removes the glycerol from the renaturation buffer. 3C protease was then added to the protein sample in a 1 to 50 ratio (wt:wt) in order to remove the N-terminal

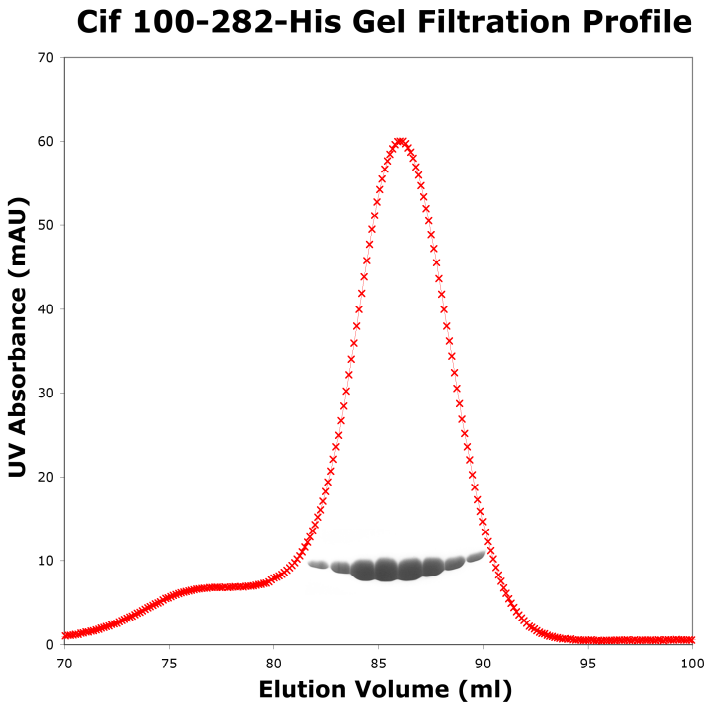


Figure 3.3: Cif 100-282-His gel filtration chromatography

Cif 100-282-His eluted as one peak on a Superdex 200 gel filtration column. The eluted fractions were analyzed by SDS-PAGE gel electrophoresis and visualized by coomassie staining as shown underneath the peak.

hexahistidine tag. Refolded Cif100-282-His was purified by anionic exchange and gel filtration chromatography, as previously described (Figure 3.3).

In order to solve the crystal structure using single-wavelength anomalous dispersion (SAD), seleno-methionine (SeMet) was incorporated into Cif. The Cif100-282-His containing plasmid was transformed into the methionine auxotroph strain 834(DE3). An overnight culture in LB medium was grown and spun down. The bacterial pellet was resuspended in minimal media with seleno-methionine and grown in a shaker. When bacteria reached OD_{600} of 0.6, 100ml of bacteria was used to inoculate

every liter of minimal media. Protein expression in bacteria was induced with 1mM IPTG when OD_{600} reached 0.4. Lastly, the bacteria were grown for an additional 3.5 hours and then harvested. SeMet Cif100-282-His was purified and crystallized using the same protocols as for the native Cif100-282-His protein. The purified and crystallized His-tagged Cif protein was 22.4 kDa and began at residue Arg-100.

3.2 The crystallization and structure determination of the Cif protein

A random factorial sparse-matrix screen was used to methodically test over 2000 crystallization conditions for purified Cif100-282-His (25mg/ml). Cif crystals were first observed in a 96-well sitting drop plate in Nextel's Anionic Screen (QIAGEN). These crystals were reproduced and optimized in 24-well plates using the hanging drop vapor diffusion method and a 1:1 volume ratio of protein to well solution. All reagents and proteins were equilibrated at 4°C prior to screening. Crystallization trials were performed on Cif100-282, in the absence of a C-terminal hexahistidine tag. It was slightly less soluble than the His-tagged construct and did not crystallize under the same experimental conditions.

Small rectangular shaped crystals of Cif100-282-His appear after 2 days at 4°C in 1.85-2.4M sodium malonate and 0.1M sodium acetate anhydrous pH 4.2-5.2 or tri-sodium citrate dihydrate pH 3.8. The native crystal used for data collection grew in 1.85M sodium malonate and tri-sodium citrate dihydrate pH 3.8. The SeMet crystal grew in 1.85M sodium malonate and 0.1M anhydrous sodium acetate pH 5.0. Crystal nucleation was mainly dependent on sodium malonate concentration; the increase in

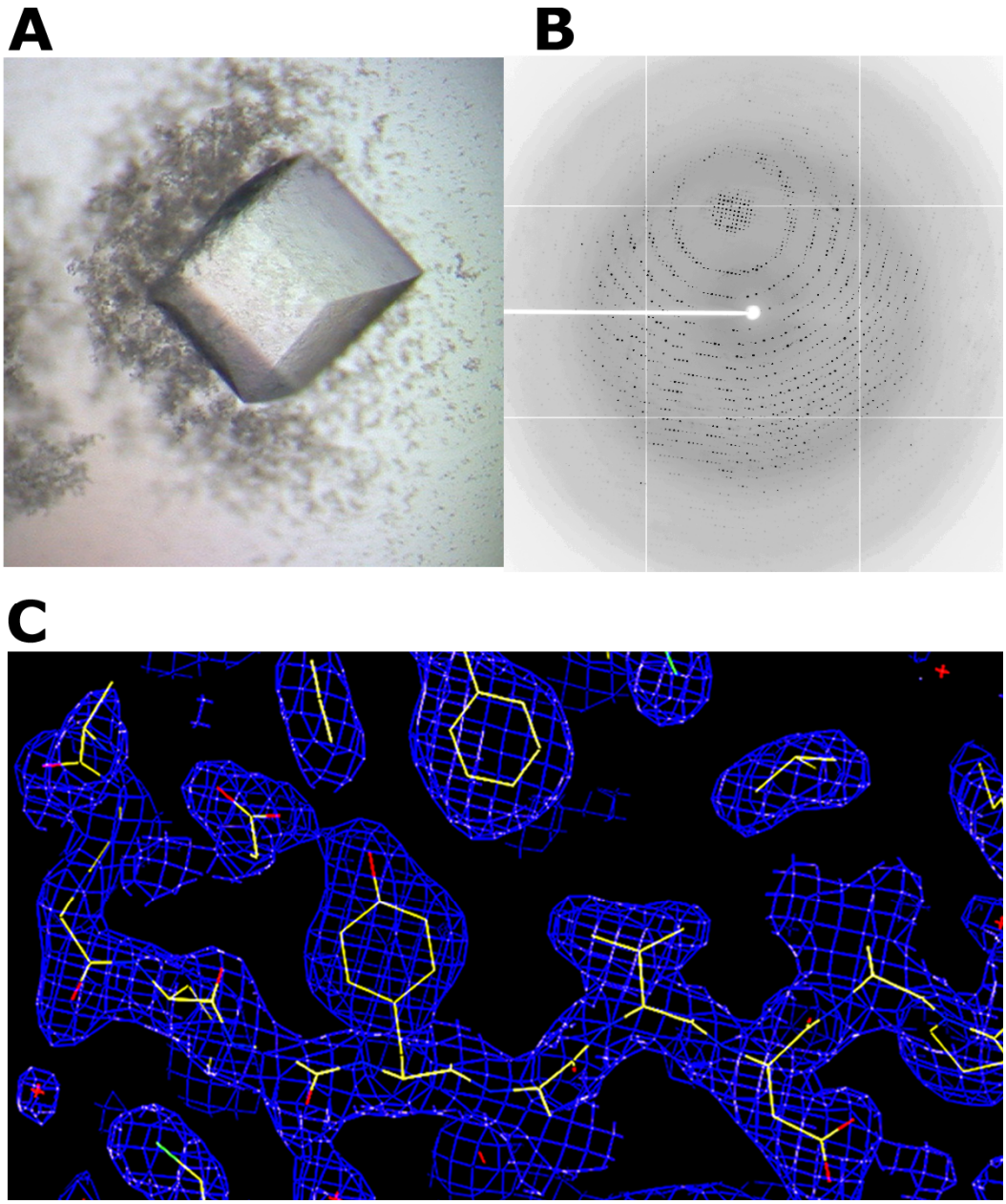


Figure 3.4: Cif 100-282-His crystal, diffraction pattern, and electron density

(A) Cif 100-282-His crystal. The Cif100-282-His crystal appeared after 2 days and grew to full size in 2-3 weeks at 4°C. The crystal grew in 1.85M sodium malonate and 0.1M tri-sodium-citrate dihydrate pH 3.8.

(B) The native Cif protein diffracts to 1.7Å.

(C) The 2.0 Å resolution of 2Fo-Fc electron density map (1.3σ contour level) from single-wavelength anomalous diffraction (SAD) phasing using Solve/Resolve with the final model superimposed.

concentration resulted in greater nucleation. Cif crystals grew up to 0.4mm in all dimensions after 2 weeks. The crystals had sharp edges but were not very birefringent (Figure 3.4A). Cif crystallized in the $p4_222$ space group, with the unit cell parameters a=b=111.515Å and c=106.808Å and angles $\alpha=\beta=\gamma=90°$. There were two Cif molecules per asymmetric unit.

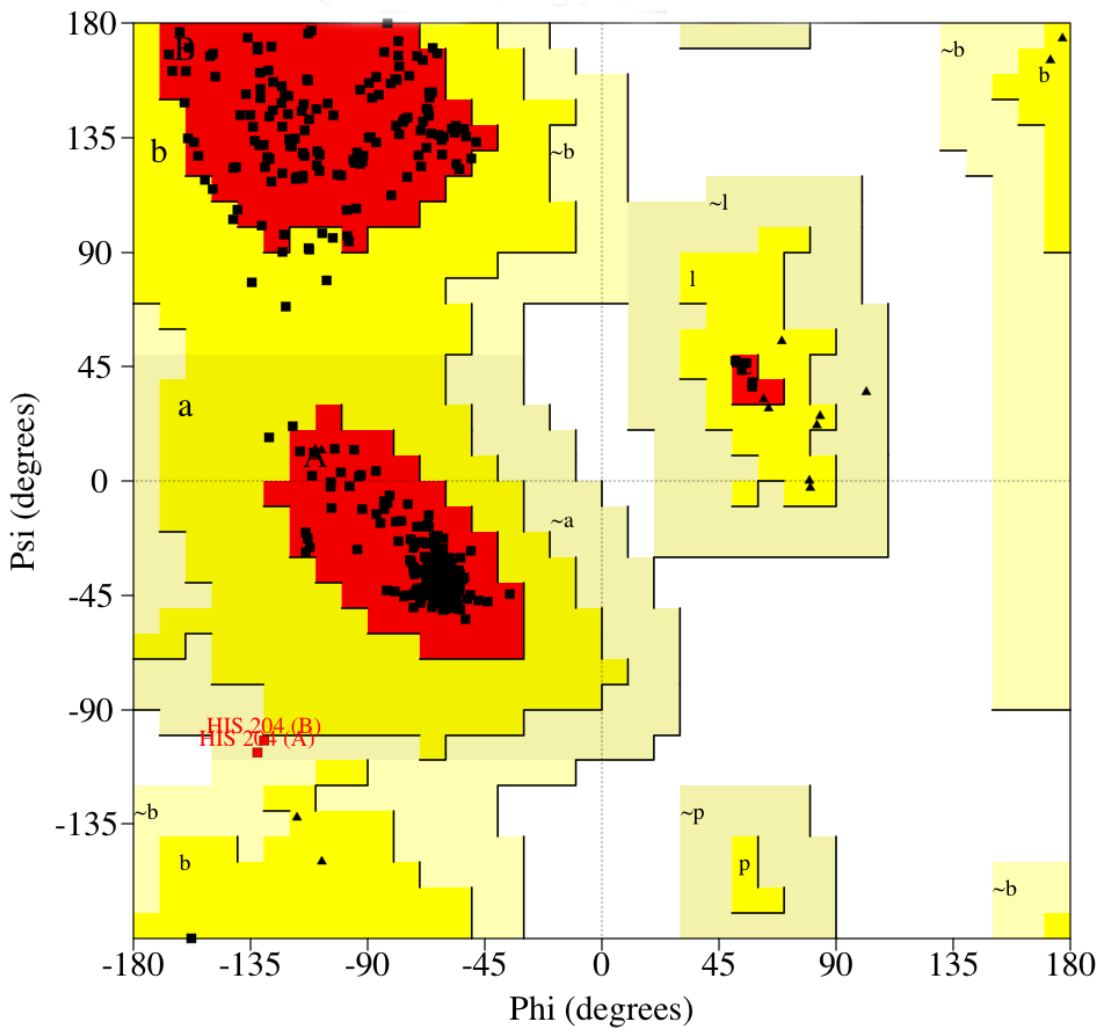
Diffraction data for the native and SeMet Cif100-282-His crystals were collected at X29 and X9A beamlines, respectively, at the National Synchrotron Light Source at Brookhaven National Laboratories. The SeMet and native Cif crystals were soaked for 10 seconds in a cryoprotectant solution consisting of 2.6M sodium malonate pH 7.0. The SeMet Cif crystal was flash-frozen at 100K by transferring into a gaseous nitrogen stream on the site of collection. The native Cif was flash-frozen in liquid propane and stored in liquid nitrogen before mounting on the goniometer. A fluorescence scan on the SeMet Cif crystal was used to locate the peak wavelength for the selenium absorption edge. The SeMet Cif data set was collected at the peak wavelength which maximizing the anomalous signal. The SeMet Cif protein diffracted to 2.0Å, and native Cif diffracted to 1.7Å (Figure 3.4B).

Data were processed with HKL2000 (Otwinowski and Minor., 1997). Phases were determined by SAD using the program SOLVE, and improved by RESOLVE (Terwilliger and Berendzen, 1999). SOLVE located a total of 10 selenium sites, with 5 from each of the two Cif molecules in the asymmetric unit. The initial model was automatically built using Arp/Warp 6.1 (Morris et al., 2002; Perrakis et al., 1999; Perrakis et al., 1997) and further refined using 1.7Å resolution data with REFMAC5 (Murshudov et al., 1997; Murshudov et al., 1999; Winn et al., 2001; Winn et al., 2003)

Table II: Data collection and refinement statistics

	SeMet SAD	Native
Data Collection		
Space group	$P4_22$	$P4_22$
Cell dimensions		
a, *c* (Å)	111.515, 106.808	111.768, 107.518
Wavelength (Å)	0.979	1.000
Resolution (Å)	50-2.0 (2.07-2.00)	50-1.70 (1.76-1.70)
R_{sym}	4.7 (17.3)	4.5 (39)
I/sI	35.9 (12.7)	28.7 (3.76)
Completeness (%)	99.3 (100)	98.86 (99.7)
Redundancy	8.5	4.7
Refinement		
Resolution (Å)		50-1.70
No. reflections		70,819
R_{work} / R_{free}		0.172/0.209
Average B-factor		26.6
R.m.s deviations		
Bond lengths (Å)		0.02
Bond angles (°)		1.710

$R_{sym} = \sum_h \sum_i | I_{h,i} - I_h | / \sum_h \sum_i I_{h,i}$, for the intensity of *i* observations of reflection *h*. $R_{work} = \sum | F_P - F_{calc} | / \sum F_P$; F_{calc} = model structure factor and 5% data omitted for R_{free}. Bond and angle deviations are from ideal values; B-factor deviations are between bonded atoms. The SeMet SAD data set was used for phasing, and the structure was subsequently refined to 1.7Å resolution using the Native data set. Values in parentheses are for the highest resolution shell.



Plot statistics

Residues in most favoured regions [A,B,L]	318	95.2%
Residues in additional allowed regions [a,b,l,p]	14	4.2%
Residues in generously allowed regions [~a,~b,~l,~p]	2	0.6%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	334	100.0%
Number of end-residues (excl. Gly and Pro)	791	
Number of glycine residues (shown as triangles)	18	
Number of proline residues	10	
Total number of residues	1153	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Figure 3.5: Cif Ramachandran Plot

Cif100-282-His Ramachandran plot is generated by PROCHECK (Laskowski et al., 1993).

from the CCP4 suite of programs (1994; Potterton et al., 2003; Saff, 1997) (Figure 3.4C). The final model was completed through iterative cycles of manual building with the program O (Jones et al., 1991) and refinement with REFMAC5 (Table II). Images were generated using Pymol (DeLano, 2002), Swiss-PdbViewer (Guex and Peitsch, 1997), and PDB Sum (EMBL-EBI). The R_{work}/R_{free} values of the final model are 17.2% and 20.9%, respectively. The average B-factor value is 26.6%. Out of 334 residues, PROCHECK (Laskowski et al., 1993) identified 95.2% of residues are in the most favored regions. 4.2% of residues are in the additionally allowed regions, and 0.6% of residues are in the generously allowed regions (Figure 3.5). The overall G-factor is 0.11. The contact and buried interface area of the two Cif molecules in the asymmetric unit were calculated with the CCP4 suite programs AREAIMOL and CONTACT (Lee and Richards, 1971; Saff, 1997). The first seven residues (100-106) and the last five histidines from the hexahistidine tag were not visible in the electron density map and therefore were omitted from the model.

CHAPTER FOUR:

CIF STRUCTURAL ANALYSIS

4.1 Overall structure of Cif

Cif(100-282) consists of a single domain with a mixed α/β fold. Cif has a central, four-stranded anti-parallel β-sheet, with a B4, B1, B2, B3 topology, surrounded on both sides by six α-helices (H1-H6) and connecting loops (Figure 4.1A). The B1-B3 β-strands are continuous in sequence and are connected by short turns. B4 is separated from B3 by helices H3-H5 and lies parallel to the B1 β-strand. The N-terminus of the crystallized construct (starting with Proline 107) forms an α-helix (H1) that packs with its axis nearly parallel to the strands in the β-sheet. The C-terminal helix (H6) points away from the core. The β-sheet is sandwiched by the α-helices with H1 and H6 at one side and H2-H5 on the other.

Two Cif molecules are present as a dimer in the asymmetric unit of the crystal (Figure 4.2A), forming an interface that buries approximately 318Å^2. This protein-protein interface is mainly comprised of N-terminal (residues 107-129) and C-terminal (residues 270-288) regions of one Cif molecule with the corresponding regions of the second Cif molecule. The dimer interface has a total of 4 hydrogen bonds and 119 hydrophobic/van der Waals contacts (Figure 4.2B). One hydrogen bond is found in the C-terminal region and three in the loop following the N-terminal helix. The protein-protein interface at the C-terminal region involves Cif's C-terminal helix H6 as well as the linker residues (GAAALE)

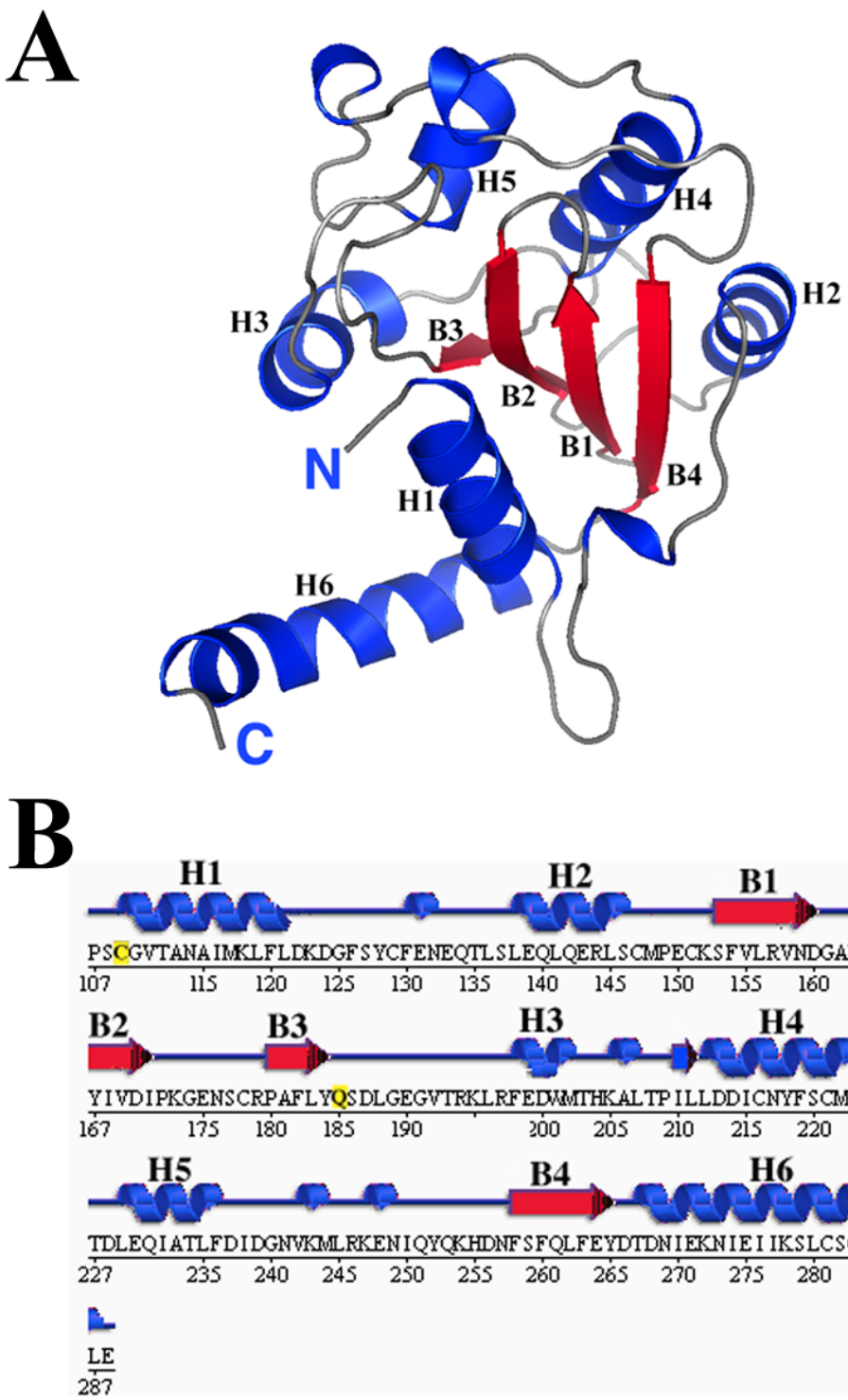


Figure 4.1: Overall structure of Cif

(A) A ribbon diagram of a Cif100-282-His monomer.

(B) Sequence and secondary structure of Cif100-282-His. The putative catalytic triad residues are highlighted in yellow. Figure generated by PDB Sum (EMBL-EBI)

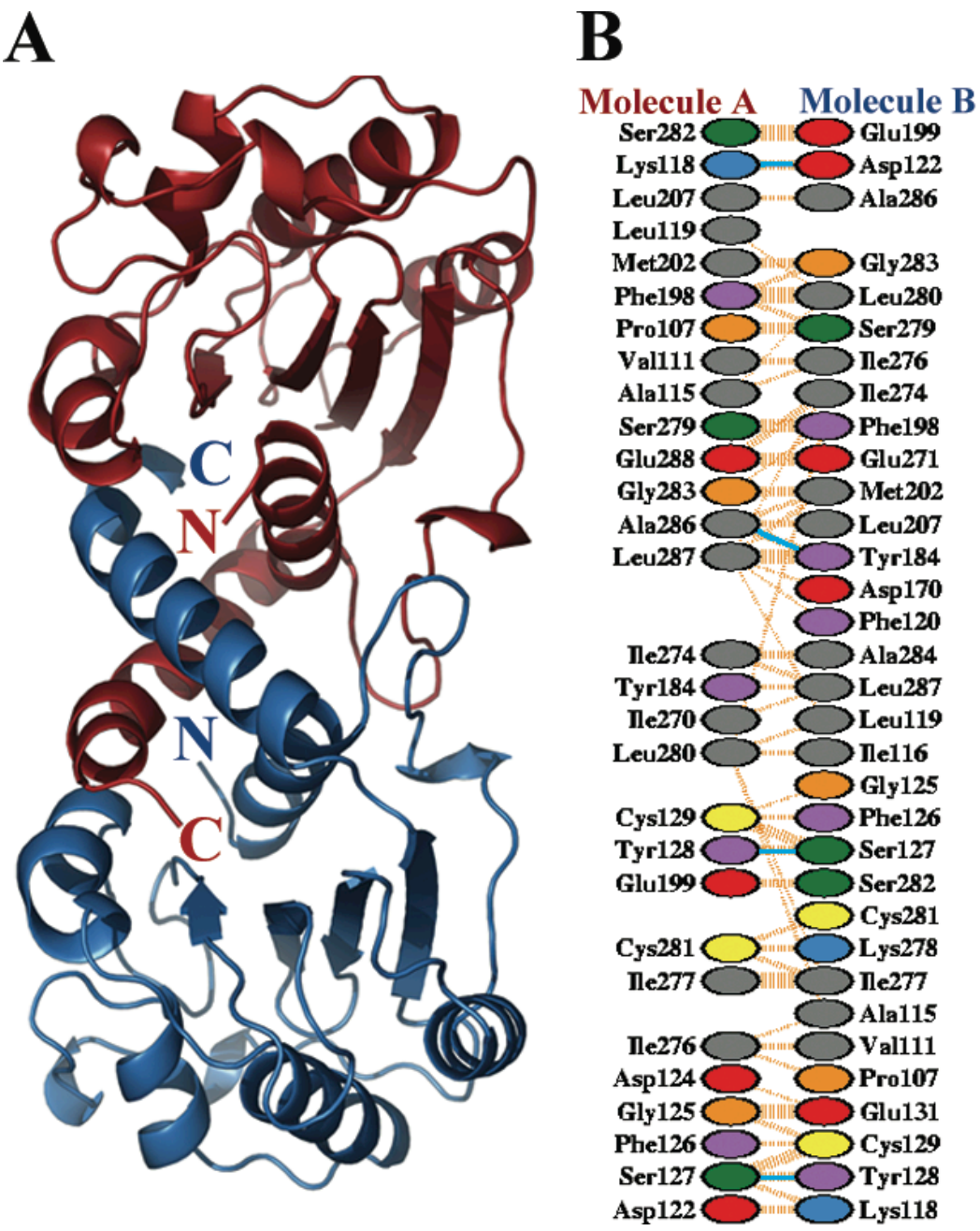


Figure 4.2: Structure of two Cif molecules

(A) Two Cif molecules in the asymmetric unit shown as ribbon diagram.

(B) The contact interface residues between the N- and C-termini in two Cif molecules of the asymmetric unit. The orange dashed lines represent non-bonded contacts, and the blue solid lines represent hydrogen bonds. Figure generated by PDB Sum (EMBL-EBI).

between Cif and the first histidine of the hexahistidine tag. This region is ordered and participates in dimer formation. The linker residue Ala286 from Molecule A hydrogen bonds to the phenol group hydroxyl of Tyr184 from Molecule B. The C-terminal helix from the first Cif monomer lies against the C-terminal helix of the second Cif molecule (Figure B). The second area of dimer interaction is found in the loop residues 122-129 of both Cif molecules. The side chain oxygen atom of Ser127 forms a hydrogen bond with the main chain NH group of Tyr128 in each monomer. The third hydrogen bond found in the N-terminal region is between the side-chain amine nitrogen (NZ) of Lys118 from Molecule A to the carboxylate oxygen atom (OD2) of Asp122 from Molecule B.

The observed Cif dimerization could be attributed to a crystal packing artifact or to the protein-protein interactions in the C-terminal α-helix and the N-terminal region. Since Cif's C-terminal helix is essential for activity, its functional role could be to bring two Cif molecules together as a dimer. Probing how this helix contributes to protein solubility and dimer formation could yield interesting functional information. Experiments are described in Chapter Five.

4.2 Cif is a structural member of the cysteine protease superfamily

A search of known structures in the Protein Data Bank revealed that Cif shares structural homology to *Pseudomonas syringae* AvrPphB (PDB ID code 1ukf) with a Z-score of 4.0 and root-mean-square deviation (RMSD) of 3.8 over 88 residues (Holm and Sander, 1995) (Table III). However, the proteins bear no sequence similarity. AvrPphB is a member of an enzyme superfamily of that consists of cysteine proteases,

acetyltransferases, and transglutaminases (Andreeva et al., 2004). While there is considerable divergence across this superfamily in the overall fold, an anti-parallel β-sheet along with an N-terminal helix which packs against the β-strands, are conserved (Rawlings and Barrett, 1994) (Figure 4.3). The catalytic core fold of Cif also closely follows the topology of other members of the cysteine protease superfamily. The conserved fold in Cif begins at its N-terminal helix (H1) and includes the central four β-strands (with B4, B1, B2, and B3 linking topology). Several of the catalytic residues in these proteins are located on these elements as discussed below.

Table III: Dali Result

WARNING pairs with Z<2.0 are structurally dissimilar
SUMMARY: PDB/chain identifiers and structural alignment statistics
NR. STRID1 STRID2 Z RMSD LALI LSEQ2 %IDE REVERS PERMUT NFRAG TOPO PROTEIN
1: 5941-A 1ukf-A 3.9 3.6 87 188 8 0 0 11 S HYDROLASE avirulence protein avrpphB
2: 5941-A 1cv8 3.8 3.3 86 173 9 0 0 11 S CYSTEINE PROTEASE staphopain
3: 5941-A 1x9yA 3.0 3.5 86 346 7 0 0 12 S HYDROLASE cysteine proteinase fragment
4: 5941-A 1dki-A 2.9 3.3 89 303 9 0 0 12 S TOXIN pyrogenic exotoxin b zymogen fragment
5: 5941-A 3gcb 2.8 4.1 100 458 10 0 0 12 S HYDROLASE gal6 (yeast bleomycin hydrolase)
6: 5941-A 1cs8-A 2.6 3.9 89 316 15 0 0 12 S HYDROLASE human procathepsin l Mutant

Cif's structure is compared to two members of the cysteine protease superfamily AvrPphB and arylamine N-acetyltransferase (NAT). AvrPphB, with its high Z-score and low RMSD, exhibits the closest fold to Cif. It is categorized as a member of the papain-family cysteine proteases or clan CA proteases. These cysteine proteases all have an invariant catalytic triad consisting of cysteine, histidine, and aspartate/asparagine that is essential for proteolytic activity (Barrett and Rawlings, 2001). The cysteine protease AvrPphB is a T3SS avirulence protein from a plant pathogen *P. syringae* that triggers a disease resistance response in its host plants (Mansfield et al., 1994). AvrPphB's catalytic triad consists of cysteine 98, histidine 212, and aspartate 227 (Shao et al., 2002). Like other members of the cysteine protease superfamily, NAT has a protease-like catalytic triad (Cys69-His107-Asp122) (Sinclair et al., 2000). NAT catalyses the acetylation of arylamines, hydrazines, and their N-hydroxylated metabolites by transferring the acetyl group from acetyl-CoA to the terminal nitrogen group of the substrate (Weber and Hein, 1985).

Initially, a potential catalytic triad (Cys109-His165-Asp187) was identified in the Cif crystal structure based on the spatial proximity between these three residues and the assumption that the third residue was either an aspartate or asparagine. Aligning this putative catalytic triad with that of other members of the superfamily resulted in a poor overall alignment in which the common core fold was misaligned (Figure 4.4A). When the core fold of Cif and the members of superfamily are structurally aligned, the catalytic triad of Cif is shifted by several angstroms. We then aligned only the side chains of Cys109 and His165 with the corresponding catalytic cysteine and histidine from the cysteine proteases (Figure 4.4B). This gave a far better structural alignment of the

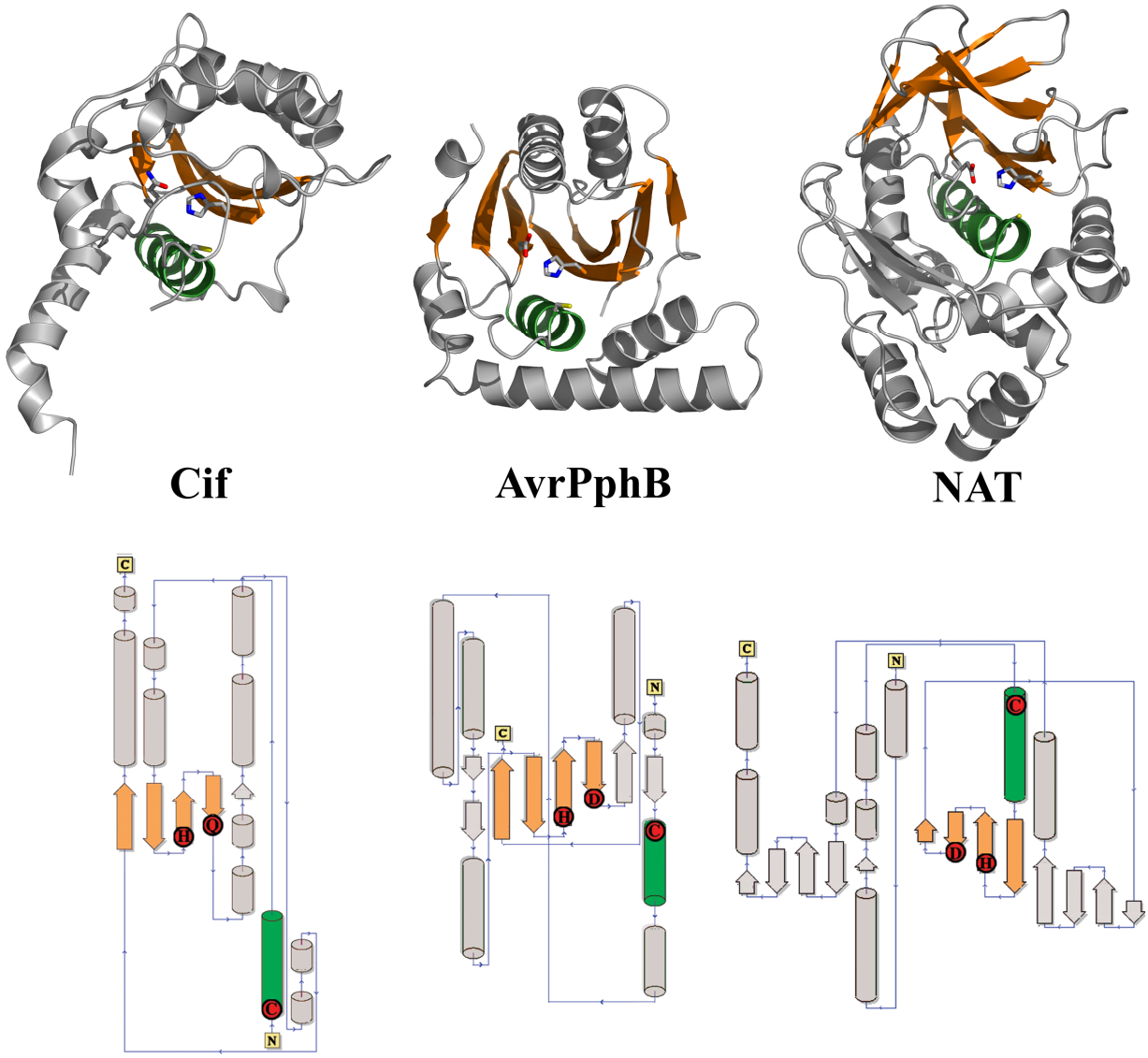


Figure 4.3: Structural comparison of Cif with cysteine protease superfamily members

Upper panel: Ribbon diagrams of Cif, AvrPphB (PDB ID code 1ukf), and NAT (PDB ID code 1e2t). The catalytic triad in the active site is rendered as a stick model. The structurally conserved α-helix is colored green, and the β-sheets are colored orange. Lower panel: Topology diagrams of the listed proteins. The catalytic triad residues of each protein are depicted as red circles.

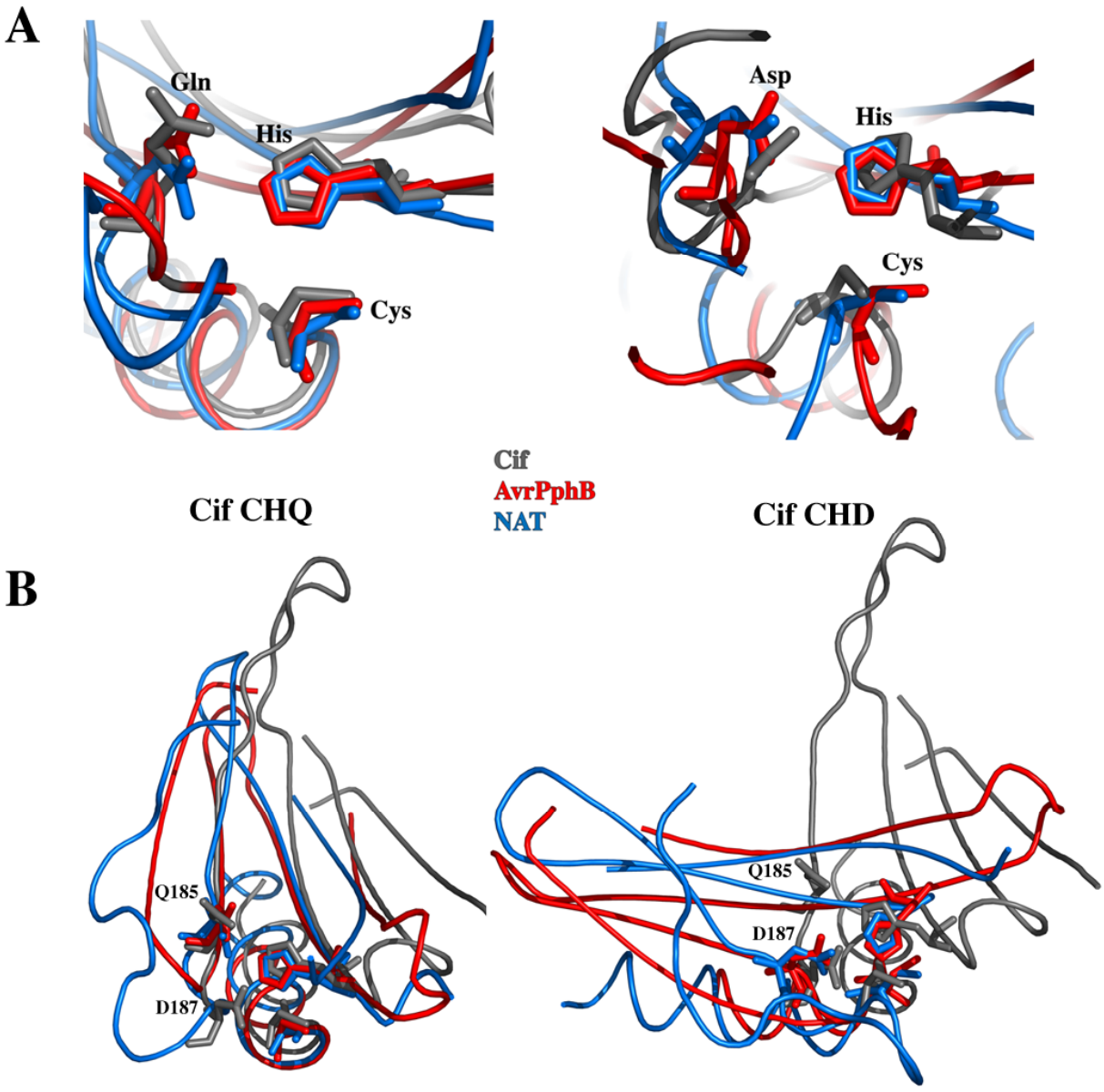


Figure 4.4: Structural comparison of Cif's proposed catalytic triads

(A) Superposition of the catalytic triads of Cif (grey), AvrPphB (red), and NAT (blue). Left panel: superposition using Cif Cys-His-Gln. Right panel: superposition using Cif Cys-His-Asp.

(B) The alignments of the catalytic cores of Cif (grey), AvrPphB (red), and NAT (blue) using two different putative Cif catalytic triad. The catalytic cores are rendered as cartoon loops with the catalytic triad residues in sticks. Left panel: Cif Cys-His-Gln as the reference. Right panel: Cif Cys-His-Asp as the reference.

common core, but Asp187 no longer aligned with the Asp/Asn of the other catalytic triads. However, glutamine 185 did; it falls in proximity to the Asp/Asn of the other proteins. In fact, the distance between bonding element in this alignment is more similar to other catalytic triads than with the Asp alignment (Figure 4.5). The amide oxygen (OE1) of Gln185 forms a hydrogen bond (2.81Å) with the protonated (ND1) side chain of His165, keeping the imidazolium ring in a favorable orientation in a very similar fashion to papain, staphopain, and AvrPphB (Shao et al., 2002). The distance between NE2 of His165 and SG of Cys109 averages 3.72Å, which corresponds to the distance observed in other members of the cysteine protease superfamily (Shao et al., 2002).

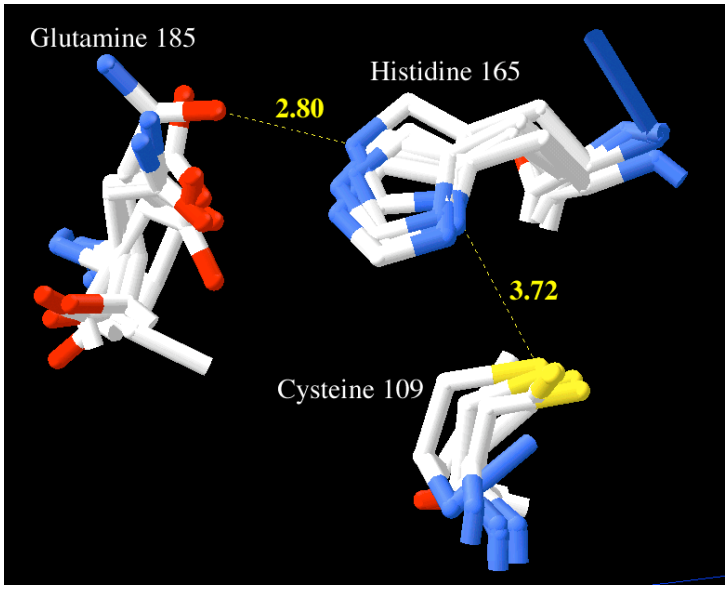


Figure 4.5: The bonding distances between Cif's catalytic residues

Cif's Cys-His-Gln catalytic triad is compared to AvrPphB/NAT's catalytic Cys-His-Asp and papain /staphopain's catalytic Cys-His-Asn. The amide oxygen (OE1) of Gln185 forms a hydrogen bond (2.81Å) with the protonated (ND1) side chain of His165. The distance between NE2 of His165 and SG of Cys109 averages 3.72A, which agrees well with other members of the cysteine protease superfamily.

In comparing the residues present at the molecular surface of Cif with those of other members of the cysteine protease superfamily (AvrPphB and NAT), it is observed that all three proteins have solvent exposed catalytic cysteine and histidine residues (Figure 4.6B). In AvrPphB and NAT, the catalytic aspartate is buried. Cif Asp187 is solvent exposed while Gln185 is buried, providing further supporting evidence for Gln185 being the likely third catalytic residue as opposed to Asp187. Although there is no predecessor for a glutamine acting as the third catalytic residue in cysteine proteases, our structural analysis suggests that Cif's Gln185 is indeed part of the catalytic triad.

In Cif, the nucleophile, Cys109, is located on the N-terminus of α-helix H1 and His165 is located on the N-terminus end of the B2 β-strand. The third catalytic residue Gln185 lies on the C-terminal end of the B3 β-strand. The relative positioning of the catalytic residues in AvrPphB and NAT is similar to that of Cif. The catalytic cysteine residue of both AvrPphB and NAT lies on the N-terminus of a helix positioned similarly with respect to the β-sheet as the helix H1 of Cif. The histidine residue is located on a β-strand, and the aspartate residue is positioned on or close to the β-strand. Like other papain-like cysteine proteases, Cif's β-sheet and the packing α-helix are separated by a groove containing the active site formed by residues Cys109 and His165, which are on either side of the active site cleft. We could not locate in Cif the Asn/Gln residue that could participate in the formation of a catalytically active structure known as the oxyanion hole. The oxyanion hole is an important feature in the cysteine protease active site that stabilizes the main-chain carbonyl group of the P1 residue of the substrate. Generally the oxyanion hole is formed by the main-chain amide of the catalytic cysteine

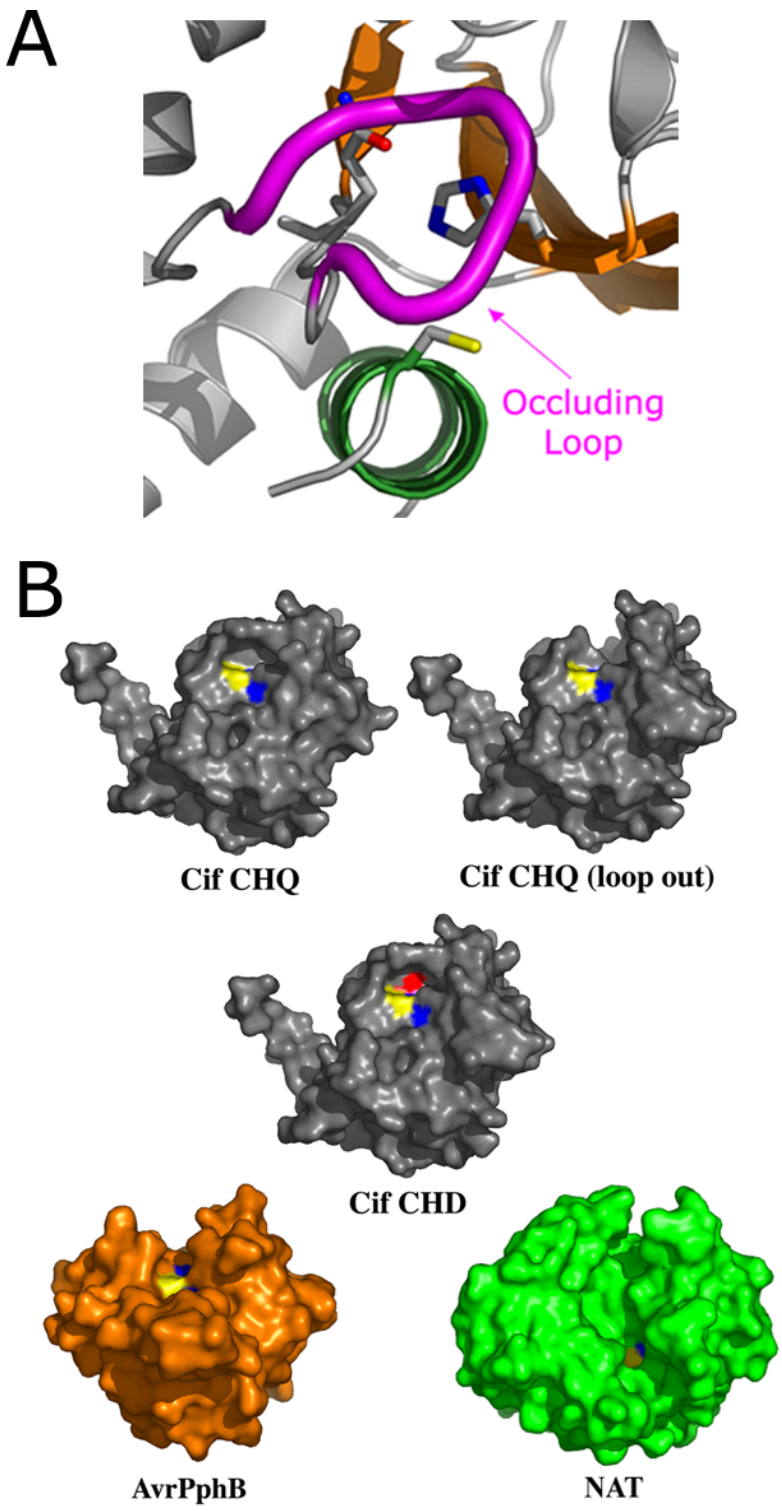


Figure 4.6: The Cif active site is occluded relative to other members of the cysteine protease superfamily

Figure 4.6: The Cif active site is occluded relative to other members of the cysteine protease superfamily

(A) The catalytic triad of Cif, with the occluding loop shown in magenta that blocks part of the active site.

(B) Molecular surface representation of Cif, AvrPphB, and NAT. The catalytic triad residues are colored as following: cysteine in yellow, histidine in blue, and aspartate in red. In Cif CHQ (loop out), *in silico* removal of the occluding loop (residues 189-195) creates an extensive cleft, reminiscent of the substrate binding grooves of AvrPphB and NAT. Cif CHQ is where Cif has cysteine, histidine, and glutamine as its putative catalytic triad residues. Cif CHD has cysteine, histidine, and aspartate as its putative catalytic triad residues.

and by the side-chain amide of the glutamine or asparagines residue (Rawlings et al., 2006). Based on the putative catalytic triad orientation, the oxyanion hole Asn/Gln residue would likely lie before proline 107, which is located outside of the visible electron density.

Cif's putative catalytic triad can be superimposed well with the catalytic triad of AvrPphB (Zhu et al., 2004), other cysteine proteases such as papain (PDB ID code 1ppn) (Kamphuis et al., 1984) staphopain (PDB ID code 1cv8) (Hofmann, 1993), and NAT from *Salmonella enterica* serovar Typhimurium (PDB ID code 1e2t) (Figure 4.4) (Sinclair et al., 2000). Such alignment yields an aligned common core, which further defines the structural homology between the cysteine protease superfamily.

A unique structural feature of Cif relative to other superfamily members is the presence of an occluding loop situated near the active site (Figure 4.6A). The occluding loop (residues 189-195) obstructs part of the active site, generating a narrow protein-binding cleft (Figure 4.6B). As a consequence, the loop possibly restricts active site accessibility to a polypeptide substrate. Such an occluding loop feature is not present in the structure of AvrPphB or NAT. These two proteins possess deep and extended substrate binding clefts. To visualize Cif without the occluding loop, we generated an image of Cif with the loop residues 189-195 deleted. As shown in Figure 4.6B, the removal of the loop created a deeper catalytic cleft, which would allow the substrate to bind more efficiently. The biological relevance of Cif's occluding loop is unknown. Whether substrate or co-factor binding would displace the loop or the loop acts as an "on-off" switch needs further investigation.

CHAPTER FIVE:

EXPERIMENTAL RESULTS BASED ON STRUCTURAL ANALYSIS OF CIF

5.1 C-terminal helix structural analysis

E. coli strains carrying mutations in the last 10 amino acids of Cif do not cause cytopathic effect in infected cultured cells. In the crystal structure, both the N- and C-terminal helices of Cif contribute substantially to dimer contacts by interacting with the corresponding helix in the second Cif molecule. As the C-terminal helix was previously shown to be essential for Cif activity (Marches et al., 2003; Charpentier and Oswald, 2004), this raises the possibility that its functional role could be to form an active dimer. To elucidate the biological relevance of Cif's C-terminal helix, two C-terminal truncated variants were constructed: (i) Cif1-265, which has its C-terminal helix completely deleted; and (ii) Cif1-272, which has the last 10 amino acids deleted, based on prior work demonstrating their importance for Cif's activity (Marches et al., 2003). Their solubility and gel filtration profiles were examined.

Both Cif1-272 and Cif1-265 were insoluble and had to be denatured, refolded, and purified under the same conditions as Cif100-282-His (Chapter Two). Cif 1-272 and Cif1-265 were loaded on to Superdex200 size exclusion column. The gel filtration profiles of the C-terminal truncated Cif proteins were compared to that of refolded Cif full-length construct under low salt buffer conditions (20mM Tris pH 8.0, 50mM NaCl,

and 2mM DTT) (Figure 5.1A,B). Experimental molecular weights (MW) were calculated based on a standard calibration curve. Under low salt conditions, Cif1-282 exhibited the properties of a dimer having an experimental MW of 69KD that was approximately double the size of a monomer's theoretical MW (32KD). However, these data do not exclude the possibility that Cif1-282 elutes as an elongated monomer. Based on their elution profiles, both Cif1-272 and Cif1-265 mutants appeared to be either highly elongated (Stroke's radius of 2.28 and 1.86, respectively) or they could also exist as higher molecular weight aggregates. The Cif N-terminus could possibly contribute to the elongation or oligomerization of Cif1-282 and the Cif C-terminal truncated mutants. Thus, a mutant with an N-terminal truncation, Cif100-282-His (the crystallized construct), was tested. Similar gel filtration characteristics between Cif100-282-His and Cif1-282 were found (Figure 5.1C,D). Under low salt conditions, refolded Cif100-282-His ran at an experimental MW of 54.9 KD which is approximately twice that of a monomer's theoretical MW (22KD). The dimerization of Cif could also be due to low ionic strength. Various Cif constructs were tested in high salt buffer conditions (20mM Tris pH 8.0, 200mM NaCl, and 2mM DTT). Cif1-282 and Cif100-282-His ran as monomer with an experimental MW of 39KD and 31.3KD respectively. Nevertheless, the C-terminal mutants both ran twice the size of their monomeric MW with Cif1-265 at 61.6KD and Cif1-272 at 58.7KD (Figure 5.1B).

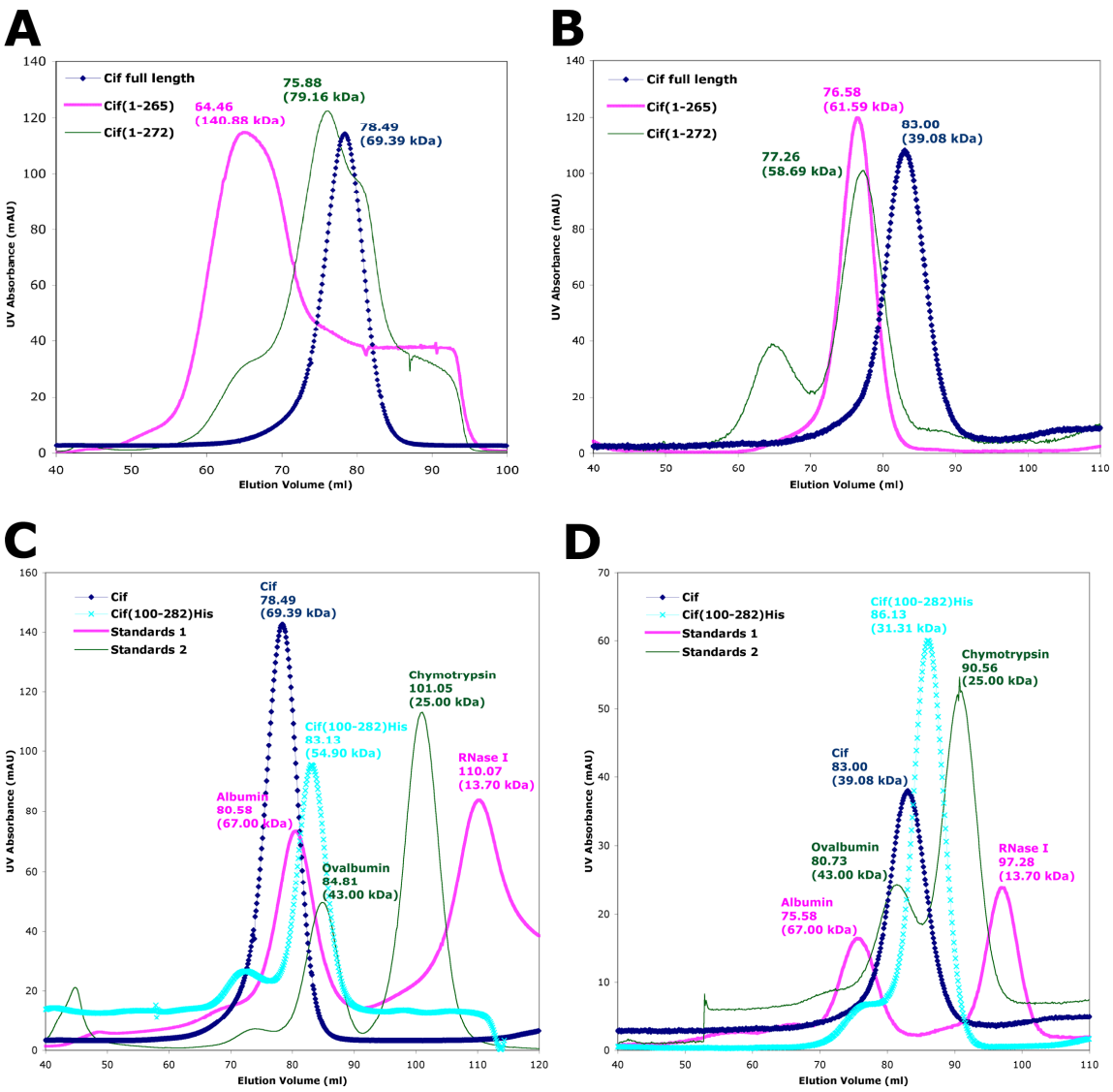


Figure 5.1: Gel filtration profiles of Cif and deletion mutants in different salt concentrations

The gel filtration profiles of Cif C-terminal deletion mutants in **(A)** 50mM and **(B)** 200mM salt concentration. Cif full-length and crystallized construct Cif(100-282)His gel filtration profile in **(C)** 50mM and **(D)** 200mM salt concentration. Cif and various standard marker proteins were run on Superdex 200 gel filtration column using an AKTA FPLC. A standard calibration curve based on the standard markers' elution volumes was used to determine the experimental molecular weights of Cif. The calculated molecular weights of globular proteins based on the elution volume are indicated in parentheses.

Gel filtration chromatography experiments indicated that both the full-length and the crystallized construct of Cif exhibited properties of a dimer but only in low salt conditions (Figure 5.1A). The C-terminal mutants were insoluble and following refolding, ran as higher molecular weight aggregates on size exclusion chromatography (Figure 5.1A). These data confirm the truncations in the C-terminus greatly affect the solubility and folding of the protein (Charpentier and Oswald, 2004) but do not clarify its role in dimer formation. Since dimer formation occurs only *in crystallo* or in low salt conditions, it is quite possible that the biologically relevant state of Cif is monomeric.

5.2 Analysis of catalytic triad mutations in Cif

5.2.1 Mutagenesis of the catalytic triad and infection assays

Based on structural analysis, Cif contains a putative catalytic triad consisting of Cys109, His165, and Gln185 or Asp187. To test whether these residues were critical to Cif's function, point mutations were constructed by site-directed mutagenesis of the native Cif and evaluated in an infection assay. Pathogenic *E. coli* strains with mutated *cif* or lacking the *cif* gene were complemented with *cif* wild type and *cif* mutant carrying plasmids to probe whether cell cycle arrest and enlargement of HeLa cells are induced.

EDL-933 and ED-31 are the two human EHEC strains, and C712-65 is a human EPEC strain. EDL-933 is the serotype O157:H7 without the *cif* gene in its genome. ED-31 belongs to serogroup O111 and carries a truncated, nonfunctional *cif* (Marches et al., 2003). C712-65 is the serotype O125:H6 that was not previously examined for *cif*'s presence. All three strains' genomic DNA were isolated and sequenced for *cif* gene

presence. EDL-933 and ED-31 did not have the full-length *cif* gene and agreed with Marches et al's data. C712-65 does contain the functional *cif* gene. C712-65 was shown to induce cell cycle arrest at the G2/M phase in infected cells.

E. coli strains EDL-933 and ED-31 were transformed via electroporation with an arabinose-inducible vector carrying either C-terminal flag epitope tagged Cif wild type or Cif point mutants (C109S, H165A, Q185, or D187A). These transformed strains were used to infect G_1/S synchronized 293T cells. 293T cells were synchronized at G_1/S phase by the double thymidine block (O'Connor et al., 1995). Cells infected with the Cif point mutant C109S and H165A had similar DNA-content profiles as the *E. coli* strains that carried either a non-functional Cif or Cif-null. On the other hand, Cif WT infected cells were arrested at the G_2/M phase of the cell cycle (Figure 5.2A). Though the D187 mutant in our analysis initially appeared to be non-functional, later experiments reveal that this point mutant is partially active. Mutations of Cif Q185 were not tested in the infection assay due to the protein's poor solubility in the bacterial recombinant protein expression system. Cif Q185 mutant would likely not be folded properly in the pathogenic bacteria due to its insolubility and not be translocated through the T3SS. Despite many efforts at substituting of Q185 for leucine, alanine, serine, asparagine, glutamate, and glycine, point mutations at this residue disrupted the folding of the protein. As shown in the molecular surface representation of Cif (Figure 4.6B), Q185 is buried in the interior of the protein. Mutation of this residue could highly compromise the protein's integrity, resulting in improper folding.

The infection assay was further refined using EPEC strain E22 mutated for *cif*. HeLa cells were infected with EPEC_*cif* mutant expressing plasmid-encoded Cif WT or

Cif mutants. As expected, Cif WT caused enlargement of cells, actin stress fibers formation, and cell cycle arrest at the G_2/M phase transition (Figure 5.2B). In contrast, cells infected with EPEC_*cif* expressing C109S or H165A mutants behaved like the control cells (Figure 5.2B), indicating that these residues are critical for Cif activity and/or translocation. In contrast, the EPEC_*cif* expressing D187A mutant revealed a partial effect, characterized by a heterogeneous monolayer including both enlarged cells harboring actin stress fibers and unaltered cells still cycling as evidenced by the presence of mitotic figures (Figure 5.2B). This partial effect was confirmed by FACS analysis of the DNA-content of infected cells. In cells infected with EPEC_*cif* expressing Cif D187A, the percentage of cells with 4N DNA-content (G_2-arrested) increased up to 40% compared to 18%, 15% and 14% found in control cells or cells infected with EPEC_*cif* expressing Cif C109S or Cif H165A respectively, but remained lower than that induced by Cif WT (40% with Cif D187A and 57% with Cif WT).

5.2.2 Cif point mutants solubility and translocation

To check whether the absence or reduced activity of the catalytic triad mutants was due to poor protein expression or solubility, the strains were treated to activate T3SS and examined for protein expression and secretion as described in Chapter Two. We showed that Cif WT, C109S, and D187A were expressed and secreted from bacteria into the medium by immunoblotting detection. H165A was poorly expressed in bacteria, which suggesting the histidine mutant is unstable (Figure 5.3A).

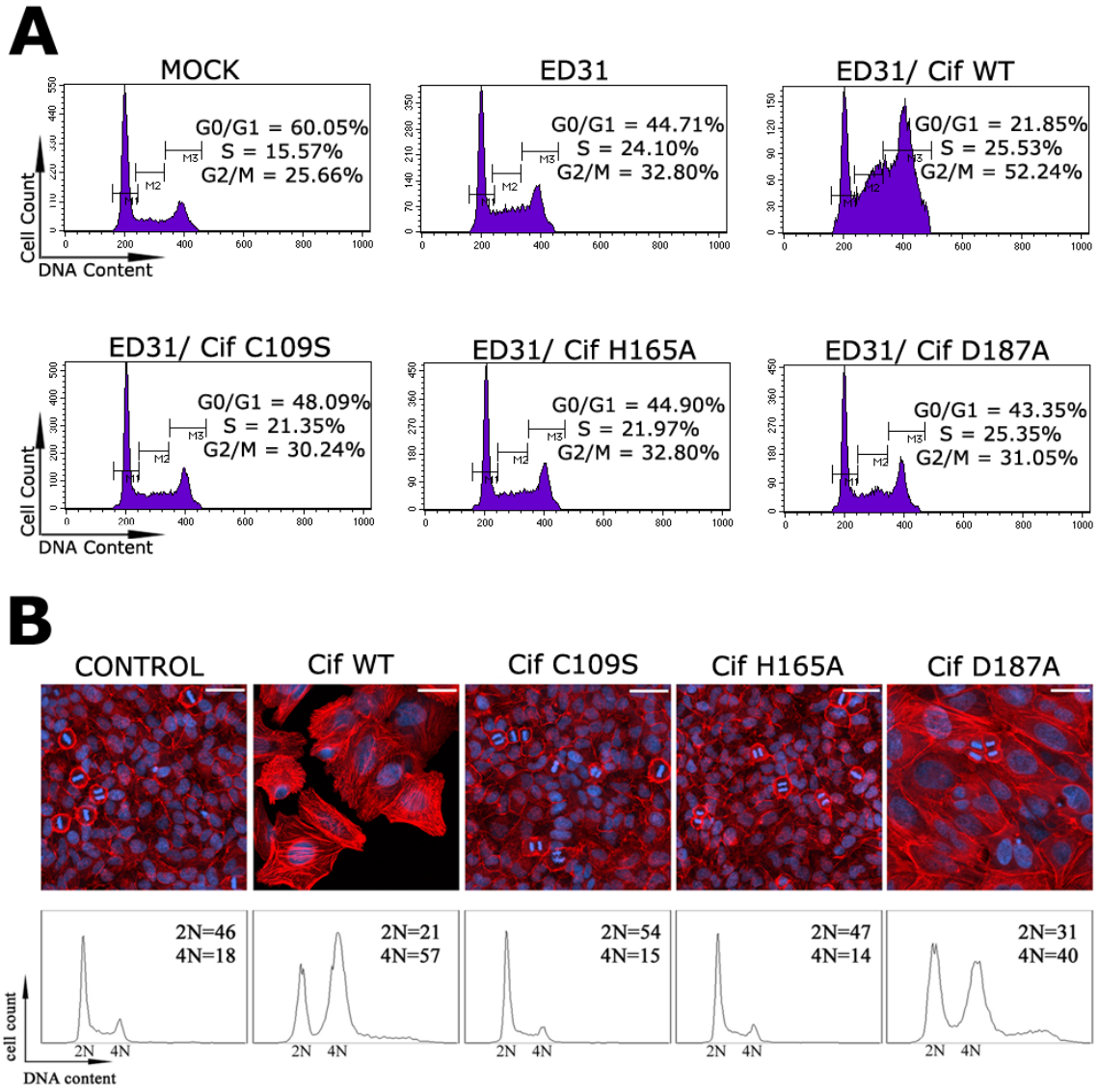


Figure 5.2: The cysteine and histidine residues of Cif catalytic triad are critical for stress fiber formation and cell cycle arrest

(A) 293T cells were synchronized at G1/S phase by double thymidine block and were infected for 4 hours. Cell cycle analysis by flow cytometery of propidium iodide stained 293T cells was performed 24 hours post-infection.

(B) Cell cycle distribution analyses of HeLa cells infected with E22Δ*cif* complemented with plasmid carrying either Cif WT or catalytic triad mutants. Upper panels: F-actin was labeled with phalloidin-rhodamine (red) and DNA with DAPI (blue) 72h post-infection. Bars correspond to 50 μm. Lower panels: cell cycle distribution according to DNA-content was analyzed by flow cytometry 48h post-infection. Percentages of 2N and 4N populations are indicated. These experiments were performed by Gregory Jubelin and Eric Oswald.

However, our secretion assay did not distinguish between flagellum secreted protein and the T3SS specific translocated proteins. We monitored the injection of the Cif mutated forms using the TEM1 β-lactamase fusion assay (Charpentier and Oswald, 2004). Wild type Cif and its mutants were fused with TEM1 β-lactamase and expressed in EPEC_*cif* strain. Translocation levels were determined by adding CCF2 as a substrate for the intracellular TEM-1 β-lactamase and by immunoblotting. All three point mutants were translocated but each to a different degree. The C109S mutant was translocated at a level similar to that of Cif WT (Figure 5.3B). In contrast, the D187A and H165A mutants were translocated into infected HeLa cells with less efficiency. We observed a two-fold and a ten-fold reduction in translocation for D187A and H165A, respectively. The significantly lower translocation level of the H156A mutant suggests that the mutant is poorly expressed and/or highly unstable in the bacterium and is quickly degraded. Immunoblotting of the H165A mutant translocation also supports this observation (Figure 5.3). These results confirmed that the cysteine residue of the putative catalytic triad is critical for Cif's activity. The C109S mutant is translocated as the wild type Cif but fails to inhibit host cell cycle progression. However, the reduced translocation of D187A and H165A mutants prevent proper analysis of the role these residues play in Cif's function using the infection model.

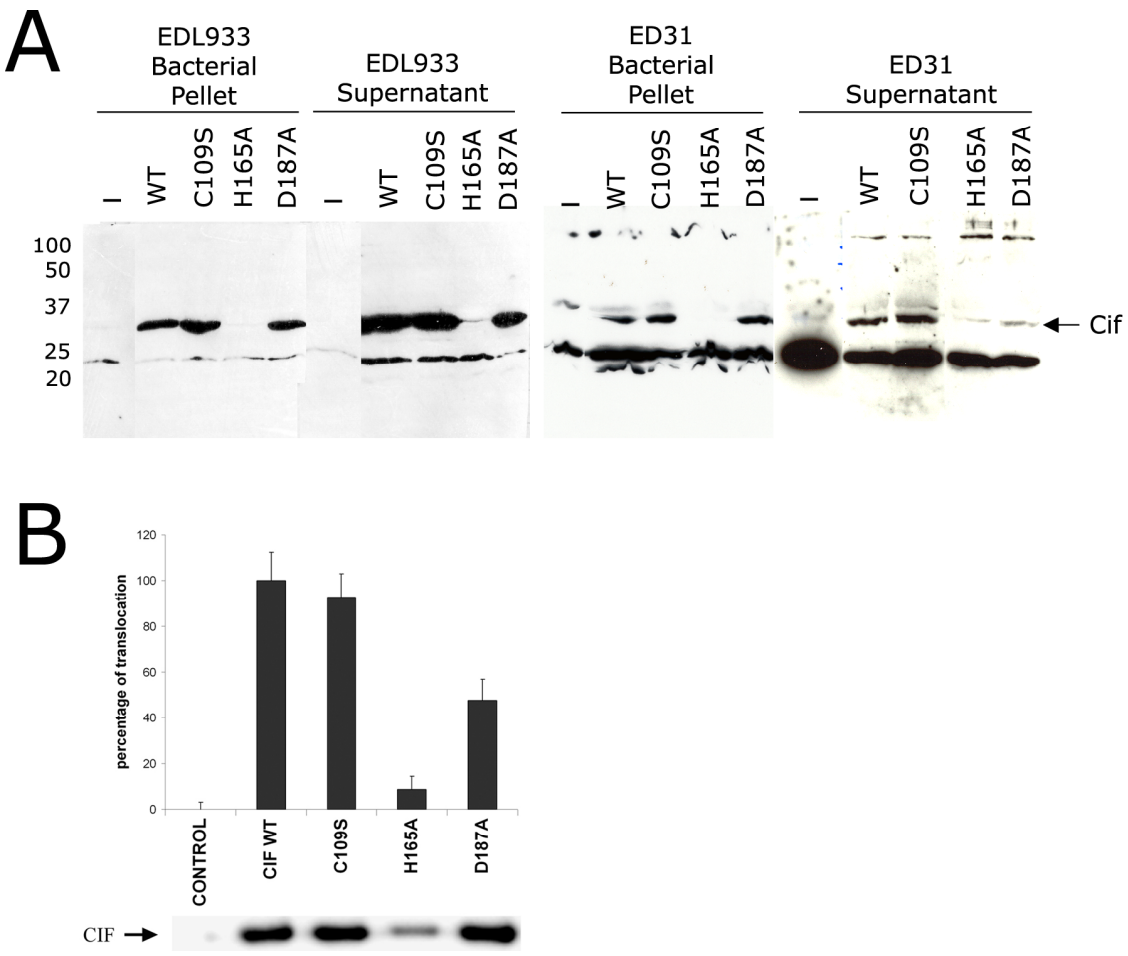


Figure 5.3: The secretion and translocation of Cif WT and catalytic triad mutants

Figure 5.3: The secretion and translocation of Cif WT and catalytic triad mutants

(A) ED31 and EDL933 complemented with plasmid carrying either Cif WT or catalytic triad mutants were grown in DMEM to activate type III secretion system. The Cif protein was expressed with a C-terminal FLAG-tag. The pellet fraction contained the bacteria, and the supernatant fraction contained the secreted proteins. The fractions were resolved by SDS-PAGE and probed for Cif using anti-FLAG antibody. The images were recomposed from the same membrane. An unknown band below Cif non-specifically reacts in all fractions with the anti-FLAG antibody.

(B) HeLa cells were exposed 90 min to E22Δ*cif* hosting plasmid expressing ß-lactamase TEM-1 protein (control) or plasmid expressing translational fusion between TEM-1 and Cif WT, mutant C109S, H165A, or D187A. Upper panel: translocation levels were determined using CCF2-AM as a substrate for intracellular TEM-1 enzyme as described previously (Charpentier and Oswald, 2004). Results are indicated as percentages compared to the translocation level of Cif WT, which was set to 100%. The error bars represent the standard error of the mean from three independent experiments. Lower panel: the level of Cif proteins accumulated in bacteria just before the translocation assay was detected by western-blotting analysis using anti-Cif antibodies. These experiments were performed by Gregory Jubelin and Eric Oswald.

5.2.3 Lipid mediated delivery of purified Cif into HeLa cells

To bypass the T3SS, translocation heterogeneity, and the need to infect cells with bacteria, Cif WT and mutant proteins were purified and introduced into HeLa cells by BioPORTER® (Genlantis), a lipid-mediated protein delivery agent (Taieb et al., 2006). BioPORTER® non-covalently encapsulates protein to create a protective vehicle. The mixture of BioPORTER® and protein is added directly to the cells, where the BioPORTER®/protein attaches itself to the plasma membrane. The BioPORTER® fuses with the lipid membrane and releases the protein into the cytoplasm, or the BioPORTER®/protein mixture is endocytosed into the cell, where the protein is released into the cytosol.

Treatment of HeLa cells with a mixture of BioPORTER® and purified Cif WT protein resulted in cell enlargement and formation of actin stress fibers, whereas BioPORTER® and Cif D187A induced a partial effect with both enlarged cells harboring actin stress fibers and unaltered cells (Figure 5.4). In contrast, treatment with a mixture of BioPORTER® and Cif C109S or H165A mutant did not affect cell morphology (Figure 5.4). Use of this alternative delivery method confirmed the inactivity of the C109S mutant and the partial phenotype induced by the D187A mutant shown in the infection model. In addition, these results also demonstrated the essential role of the H165 residue since its mutation abolished the phenotype upon lipid delivery. Taken together, these results clearly indicate that the mutation of the cysteine and histidine residues in the putative catalytic triad completely abrogate cell cycle arrest normally

induced by Cif and D187A to a lesser extent. These results provide strong evidence that Cif's catalytic triad is critical for cytotoxic activity.

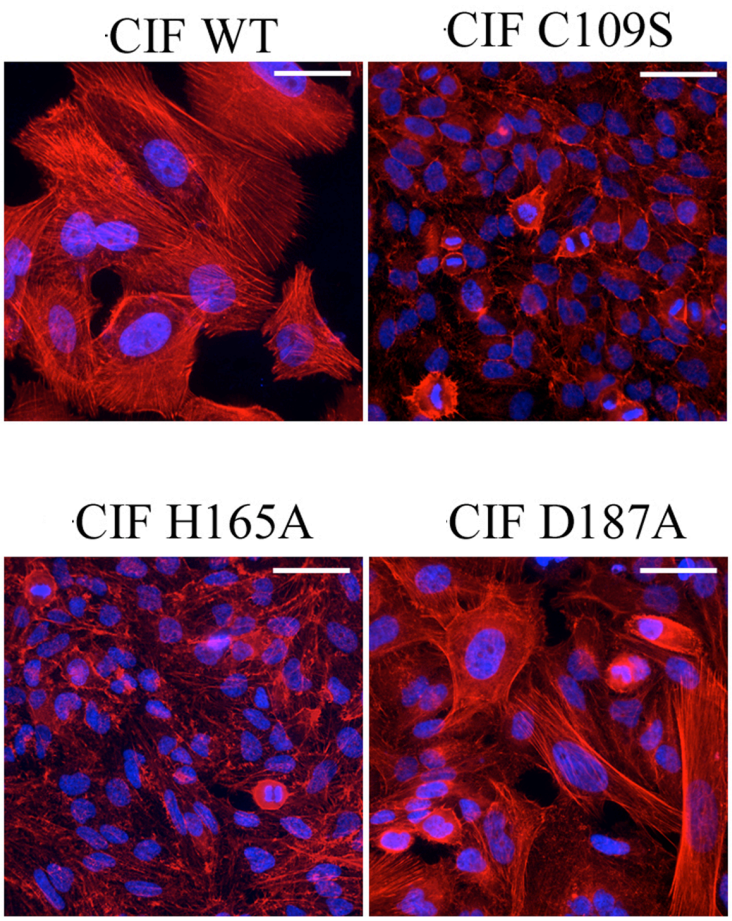


Figure 5.4: Purified Cif C109S and H165A mutants do not cause cytopathic effect HeLa cells were exposed 4 hours to the indicated proteins in combination with a lipidic delivery agent (BioPORTER®) and further incubated 72 hours. F-actin was stained with phalloidin-rhodamine and DNA with DAPI (blue). Bars represent 50 µm. These experiments were performed by Gregory Jubelin and Eric Oswald.

5.3 Cysteine protease experiments

5.3.1 Autoproteolytic activity test

Cif is structurally homologous to the cysteine protease AvrPphB. AvrPphB triggers the hypersensitive response in plants by cleaving Arabidopsis's protein kinase PBS1. But first AvrPphB must go through an autoproteolytic process to convert the 35KD precursor into a 28KD protein. The conversion is dependent on the invariant catalytic residues (Cys98, His212, and Asp227) (Shao et al., 2002). Since AvrPphB possesses autoproteolytic activity, Cif WT and catalytic triad point mutants were assayed for such activity. 10μg of Cif WT and mutants were incubated in the presence and absence of 2μM E64 and 5mM PMSF at 37°C for 24 hours. E64 is a non-competitive, irreversible inhibitor of cysteine proteases, and PMSF inhibits a broad spectrum of serine proteases and some cysteine proteases such as papain. Experimental data indicated that Cif does not self-cleave at pH 8.0. Both Cif WT and mutants appear as full-length on the SDS-PAGE gel (Figure 5.5A).

Cif could have autocatalytic activity that resembles *Pseudomonas syringae* effector AvrRpt2. AvrRpt2's predicted secondary sequence structure is similar to the catalytic core of staphopain (Axtell et al., 2003), which is also identified as a structural homologue of Cif by Dali (Table III). Unlike AvrPphB, AvrRpt2 undergoes auto-proteolytic processing after it is translocated into the host cell. AvrRpt2 requires an eukaryotic host factor known as cyclophilin ROC1 for the self-cleavage activity (Coaker et al., 2005). It is possible that Cif requires a host factor(s) for self-processing. Cif 's

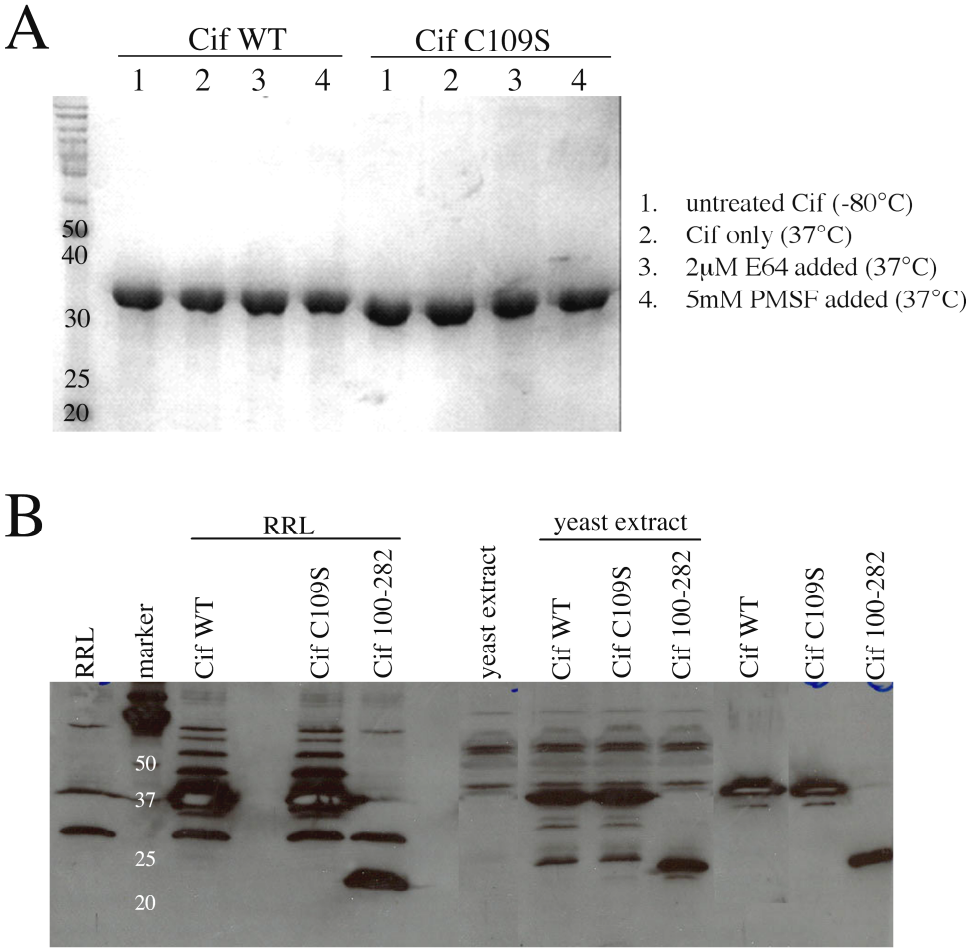


Figure 5.5: Cif does not have autoproteolytic activity *in vitro*

(A) The first 4 lanes correspond to wild type Cif. The 4 terminal lanes correspond to Cif C109S mutants. The samples were incubated at 37°C for 18 hours in the presence and absence of protease inhibitors. Wild type Cif does not show any autoproteolytic activity (lane WT #2)

(B) Various N-terminal His_6-tagged Cif constructs were incubated with either rabbit reticulocyte lysate (RRL) or yeast extract.

autoproteolytic activity was tested in the presence of yeast extracts and rabbit reticulocyte lysate. Cif WT and mutants with N-terminal hexahistidine tag were incubated with either yeast extract or rabbit reticulocyte lysate at 37°C for one hour and analyzed by immunoblotting. The immunoblotting detection using anti-His(6) (Invitrogen) shows no visible cleavage between Cif WT and mutants (Figure 5.5B).

Under the conditions tested, Cif did not exhibit autoproteolytic activity in the presence or absence of eukaryotic proteins. Like *Yersinia's* T3SS effector YopT, Cif may not need processing to be fully functional without processing. YopT is a cysteine protease with an invariant catalytic triad cysteine, histidine, and aspartate that cleaves RhoA, Cdc42, and Rac at their C-termini, thus releasing them from the membrane, and disrupting the host cytoskeleton (Shao et al., 2002; Sorg et al., 2001; Zumbihl et al., 1999)

5.3.2 *In vitro* studies of Cif

The structural analysis of Cif suggested that it is a member of the cysteine protease superfamily. We therefore attempted to demonstrate protease activity of purified recombinant Cif *in vitro*. Cif, catalytic triad mutants and various N-terminal deletion mutants were tested with a non-specific protease substrate. Casein derivative that is labeled with the green-fluorescent BODIPY FL (Molecular Probes' EnzChek Protease Assay Kit) was used as a substrate to test Cif protease activity. Proteolytic cleavage is detected by the increase in fluorescence resulting from protease cleavage of the intramolecularly quenched substrate and releases the highly fluorescent BODIPY FL

dye-labeled peptide. The proteolytic assay was performed as suggested by the manufacturer. Briefly, five or ten micrograms of Cif, active site mutants, various N-terminal deletion mutants, and protease controls (papain and subtilisin) were added to BODIPY FL casein and incubated at room temperature, protected from light, for at least 1 hour in a digestion buffer consisting of 10mM MES, pH 6.5. The detection of cleavage products was followed by fluorescence spectrophotometry (excitation/emission: 485/530nm). HeLa extract (200µg) was also included in the assay buffer in some trials as indicated (Figure 5.6A). Variations in pH (pH 4.2-9.0), salt concentration (0-200mM NaCl), and additives (1mM ATP, 5mM $MgCl_2$, or 2mM $CaCl_2$) were screened for optimization of the activity buffer. However, none of the assay conditions showed evidence of protease activity in Cif. It is possible that Cif is highly specific for its intracellular substrate and does not cleave general protease substrate like casein. As with clostridial neurotoxins and anthrax lethal factor (Pellizzari et al., 1996; Vitale et al., 2000), Cif could interact with specific structural motifs to recognize its substrate and achieve full enzymatic activity.

To test whether the occluding loop could restrict accessibility to the active site, we generated a Cif mutant lacking residues 189-195 by site-directed mutagenesis with primers that had this region deleted. Cif Δ189-195 did not increase protease activity in the *in vitro* assay (Figure 5.6B). Treatment of HeLa cells with a mixture of BioPORTER® and the Cif Δ189-195 mutant protein did not induce stress fibers or cell cycle arrest (Figure 5.7A), although the poor solubility and stability of this construct makes any interpretation of the results difficult, as the deletion likely compromised the protein at the structural level.

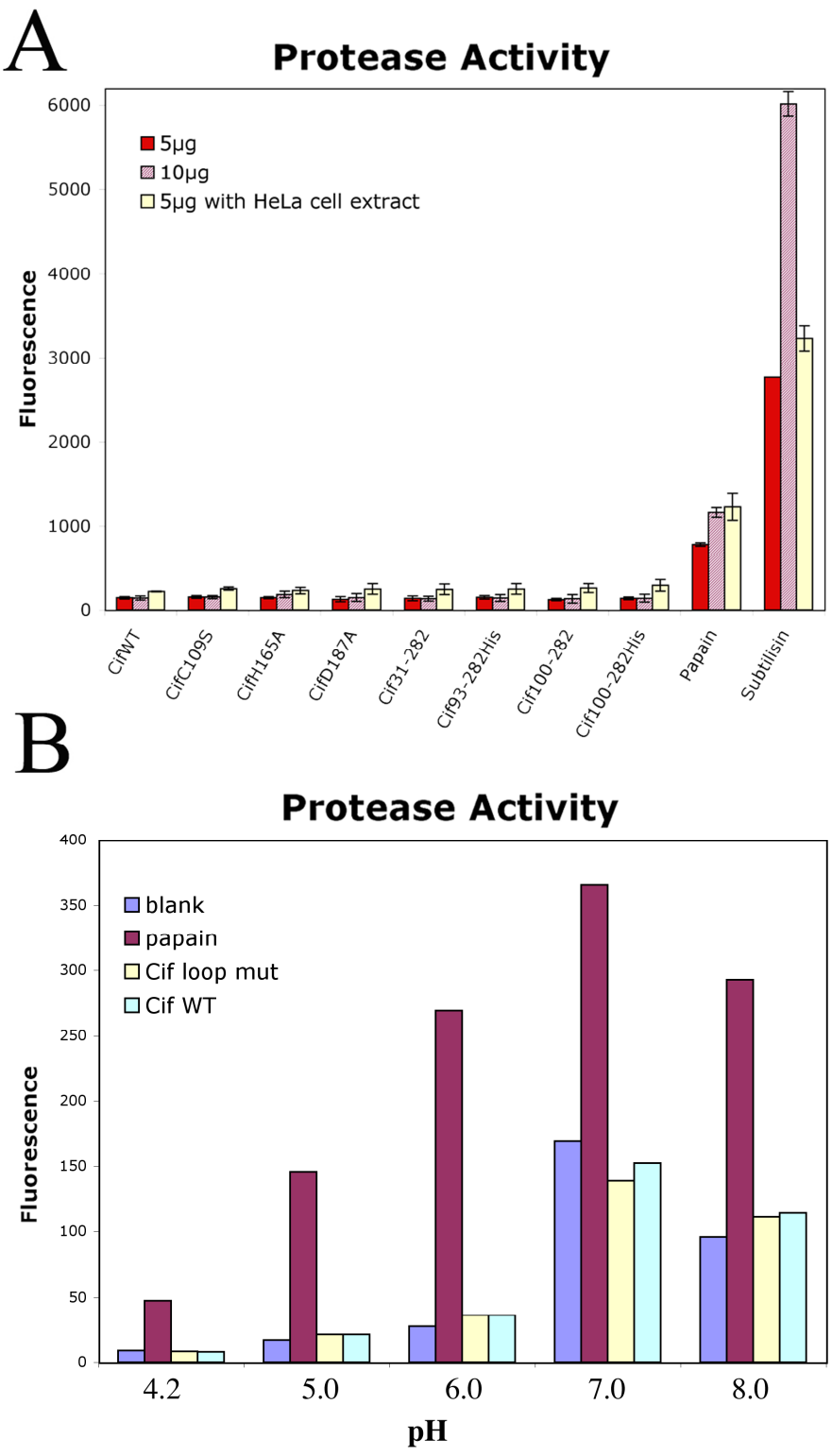


Figure 5.6: Cif does not cleave general protease substrate casein *in vitro*

Figure 5.6: Cif does not cleave general protease substrate casein *in vitro*

(A) Cif does not possess general protease activity *in vitro*. Five or ten micrograms of purified Cif constructs were incubated with BODIPY FL conjugated casein for one hour at room temperature with or without HeLa cell extract in the buffer. The fluorescence of BODIPY FL-labeled protease hydrolysis products was measured (excitation/emission - 485/530nm). Five or ten micrograms of cysteine (papain) and serine (subtilisin) proteases were included as positive controls. The error bars represent the standard error of the mean from three independent experiments.

(B) Cif loop mutant Δ189-195 was incubated with BODIPY FL casein at various pHs.

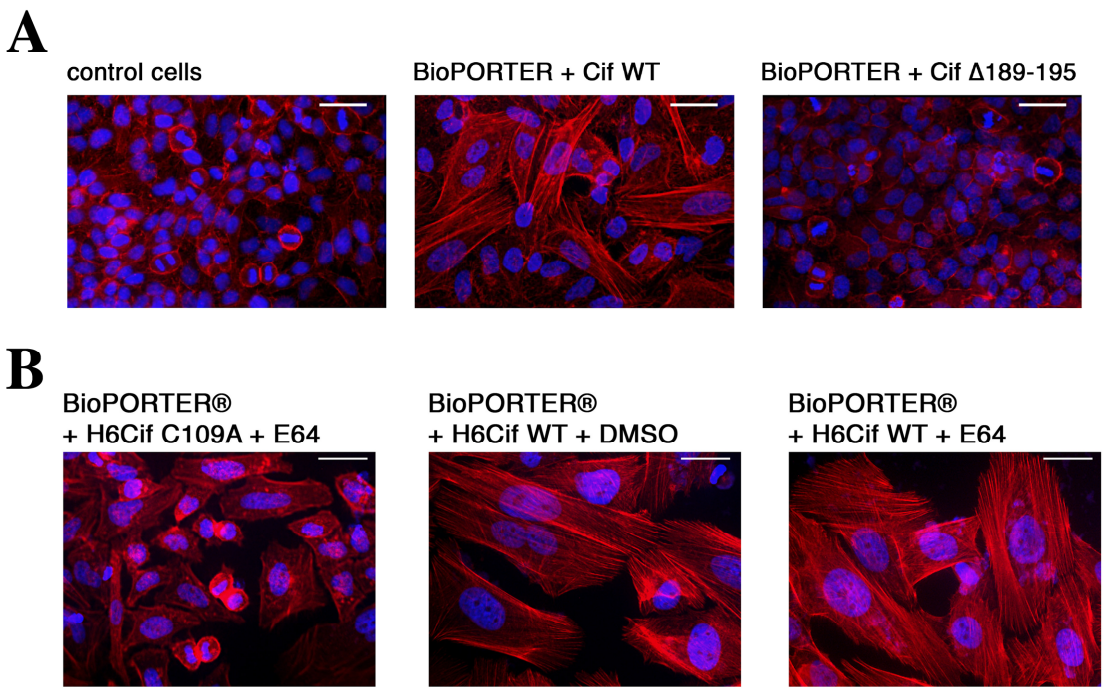


Figure 5.7: Cysteine protease activity is not required for the Cif induced phenotype

Cellular morphology of HeLa cells treated with BioPORTER® reagent containing either mock (DMSO), wild type, or mutant Cif. F-actin was stained with phalloidin-rhodamine (red) and DNA with DAPI (blue). Bars represent 20 µm. These experiments were performed by Gregory Jubelin and Eric Oswald.

(A) Effect of Cif Δ189-195 loop mutant. **(B)** Cif pre-incubated with 70µM of the irreversible cysteine protease inhibitor E64.

In a further attempt to assess protease activity of Cif, we used the cysteine protease inhibitor E64. Addition of 70μM E64 in the mix of BioPORTER® and Cif WT assay did not impair Cif activity (Figure 5.7B). This latter experiment suggests that cysteine protease activity is not required for the Cif induced phenotype. E64 has been shown to be highly specific for cysteine proteases. It is possible that Cif possesses a cysteine protease like catalytic triad but functions like a non-protease member, such as NAT of the cysteine protease superfamily.

5.4 Acetyltransferase activity analysis

In addition to having similarities with cysteine proteases, Cif also shares structural homology with arylamine N-acetyltransferases. As NAT is not inhibited even with high concentrations of E-64 or other cysteine protease inhibitors (Rodrigues-Lima et al., 2001), this leaves open the possibility that acetyltransferase activity for Cif is responsible for the cytopathic phenotype. For example, the *Yersinia* T3SS virulence effector YopJ, long thought only to be a cysteine protease, has been shown to act as an acetyltransferase that modulates host activity (Mukherjee et al., 2006; Orth et al., 2000).

To investigate the acetyltransferase activity of Cif, in the absence of a putative substrate, we first examined Cif for autoacetylation activity using a radiological assay. Cif WT or catalytic triad mutants were incubated with ^{14}C-acetyl-CoA, and the labeling of Cif was measured using a scintillation counter after retention on a nitrocellulose filter (Chapter Two). 50-100μg of HeLa cell extract was included in some assays to see if

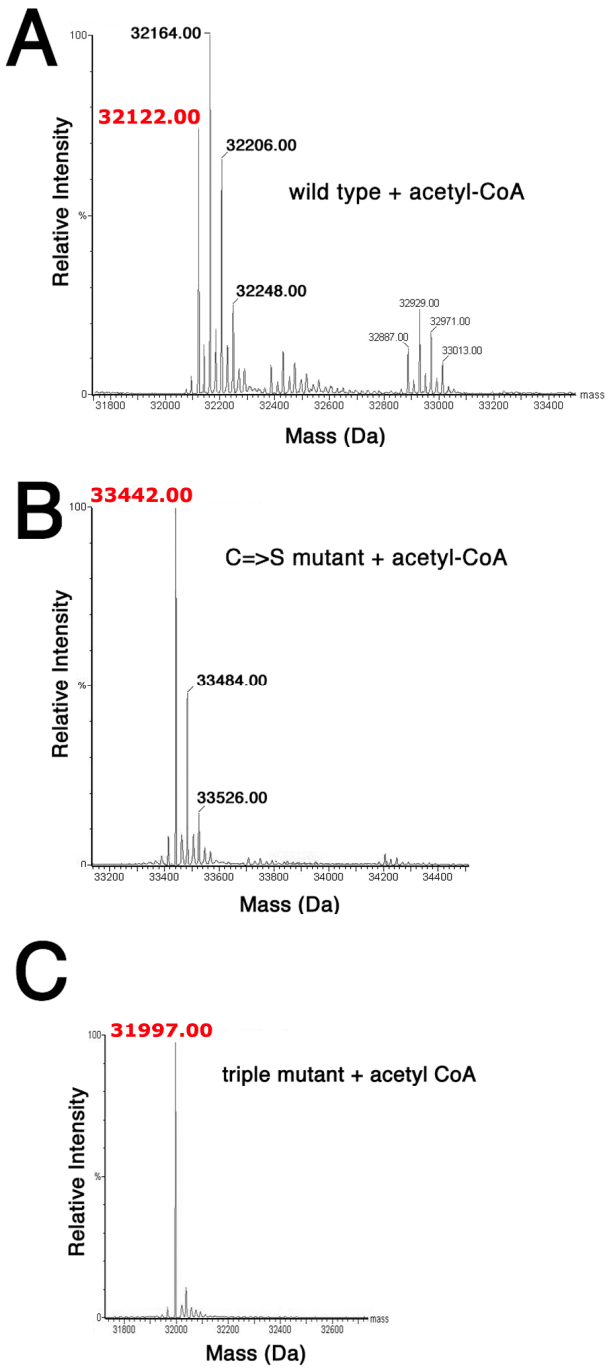


Figure 5.8: Cif possesses acetyltransferase activity *in vitro*

Electrospray ionization mass spectra of wild type Cif and catalytic site mutants incubated with acetyl-CoA. The mass colored in red in each graph corresponds to the expected Cif mass (non-acetylated). The other peaks of Cif are separated by 42 Da and interpreted as acetylated species. (**A**) Wild-type Cif (**B**) Cif (C109S) (**C**) Cif triple mutant of the active site.

(i) Cif required host cellular factors for activity and (ii) whether Cif could cause a detectable amount of acetylation in host proteins. Under these assay conditions, we were unable to detect any radiolabeled Cif (data not shown).

We then used mass spectrometry, a more sensitive technique to probe Cif autoacetylation activity. Purified Cif WT protein and C109S mutant were incubated with acetyl-CoA and then analyzed by mass spectrometry (Figure 5.8). In the wild type Cif protein spectrum, there were a series of peaks separated by 42 Da, that were interpreted as multiple acetylation events, as well as a much reduced peak at a mass corresponding to unmodified Cif (Figure 5.8A). In contrast, the catalytic triad mutant Cif C109S showed partial acetylation activity. The main peak in its spectrum was the expected, unaltered mass. However, there was a peak that was 42 Da larger than the expected mass, but at much lower level than observed in the wild type protein. No evidence of poly-acetylation was observed (Figure 5.8B). It is possible that a single point mutation from cysteine to serine could still possess residual activity *in vitro*. We therefore generated a triple mutant of Cif with two of the putative catalytic residues mutated (C109S/H165A/D187A) and tested its auto-acetylation activity *in vitro*. The mass spectrum for the triple mutant had one distinct peak corresponding to Cif's expected mass and showed no substantial peaks at 42 Da intervals (Figure 5.8C). Both mutants behaved identical to the wild type protein in solution in terms of stability and solubility (data not shown). These mass spectrometry results suggest that Cif possesses acetyltransferase activity and that mutations in the active site reduce (Cif C109S) or abrogate (triple mutant) this auto-acetylation activity.

CHAPTER SIX:

IDENTIFICATION OF CIF HOST TARGET

6.1 Yeast two-hybrid screening

Identifying host interaction partners for Cif should be beneficial in elucidating the biochemical mechanism of Cif induced cell cycle arrest. Cif WT and an inactive Cif with a mutated catalytic cysteine residue (C109S) were used for yeast two-hybrid (Y2H) screening. The screen relies on the substrate binding to the catalytic domain of the protease or acetyltransferase. If Cif is a cysteine protease or an acetyltransferase, then the active Cif may cleave or acetylate the substrate before a stable interaction with the substrate is formed. Thus, an inactive Cif can function as a trap to stably bind the substrate without cleaving or acetylating it. Cif WT and Cif C109S mutant were cloned into vector pB27 (LexA, C-terminal fusion) and used as bait for Y2H screening. Y2H screenings of both human colon and placenta cDNA libraries were performed and analyzed by HybriGenics using modified techniques as previously described (Fromont-Racine et al., 1997). The yeast two-hybrid screen (HybriGenics), identified Nedd8 as the host cell interaction partner for Cif (Table IV).

6.2 Cif binds the small ubiquitin-like molecule Nedd8

A cDNA clone encoding Nedd8 (neural precursor cell expressed, developmentally down-regulated 8) was isolated as a positive interaction partner of Cif. Nedd8 was

Table IV: Yeast two-hybrid hits

Bait/ cDNA library	Clone Name	Gene Name	Start	Stop	Frame	Orientation
Cif C109S/ Colon	hgx1199 1aA-41	hNEDD8	-94	No Data	IF	Sense
	hgx1199 1aA-28	hNEDD8	-88	277	IF	Sense
	hgx1199 1aA-42	hNEDD8	3	294	IF	Sense
	hgx1199 1aA-47	hNEDD8	3	294	IF	Sense
	hgx1199 1aA-26	hNEDD8	3	294	IF	Sense
	hgx1199 1aA-43	hNEDD8	3	294	IF	Sense
	hgx1199 1aA-23	hNEDD8	4	No Data	OOF1	Sense
Cif WT/ Placenta	hgx1084v1_pB27A-19	hNEDD8	3	No Data	IF	Sense
	hgx1084v1_pB27A-189	hNEDD8	3	572	IF	Sense
	hgx1084v1_pB27A-255	hNEDD8	3	525	IF	Sense
	hgx1084v1_pB27A-301	hNEDD8	3	570	IF	Sense
	hgx1084v1_pB27A-311	hNEDD8	3	570	IF	Sense
	hgx1084v1_pB27A-169	hNEDD8	3	572	IF	Sense
	hgx1084v1_pB27A-303	hNEDD8	3	570	IF	Sense
	hgx1084v1_pB27A-308	hNEDD8	3	573	IF	Sense
	hgx1084v1_pB27A-109	hNEDD8	3	571	IF	Sense
	hgx1084v1_pB27A-280	hNEDD8	3	No Data	IF	Sense
Cif WT/ Colon	hgx1030v1_pB27A-20	hNEDD8	3	294	IF	Sense
	hgx1030v1_pB27A-2	hNEDD8	3	294	IF	Sense
	hgx1030v1_pB27A-30	hNEDD8	3	294	IF	Sense

identified in three Y2H screenings with different Cif constructs, first as a lower confidence hit using the wild type protein in both human colon and placenta cDNA libraries and then as a high confidence hit using the C109S mutant in the human colon cDNA library (Table IV).

Nedd8 is a highly conserved 81 amino acid protein that is 60% identical and 80% homologous to ubiquitin (Kumar et al., 1993). As with ubiquitin activation, the C-terminus of Nedd8 requires cleavage by a specific hydrolase before it can be covalently conjugated to its substrates, the cullin family of proteins and enhance their activity

(Lammer et al., 1998; Liakopoulos et al., 1998; Osaka et al., 1998). Cullin family proteins are subunits of RING based E3 ubiquitin ligase complexes, which regulate the degradation of broad range of cellular proteins. For example, the Cullin-1 associated SCF E3 ligase is the key regulator in G_1 to S transition in the cell cycle, signaling cascades, and developmental processes (Bai et al., 1996; Dealy et al., 1999; Feldman et al., 1997; Jiang and Struhl, 1998; Kamura et al., 1999; Skowyra et al., 1997; Yaron et al., 1998).

To biochemically verify the genetic interaction of Cif with Nedd8 that was observed in the yeast two-hybrid screen, we first performed a glutathione S-transferase (GST) pulldown experiment (Chapter Two). Initially, purified GST-Cif C109S or GST alone, acting as a control, were immobilized on a glutathione-sepharose matrix. The proteins were then incubated with 293T cell lysate containing transient expressed Nedd8 tagged with the FLAG epitope. While the GST control did not pull out any FLAG-tagged Nedd8, the GST-Cif C109S bound FLAG-tagged Nedd8 and what were interpret to be various Nedd8 modified proteins from the 293T extract observed at higher molecular weights (Figure 6.1A). Consistent with the yeast two-hybrid results, the C109S mutant pulled down Nedd8 more efficiently than wild type Cif under the same experimental conditions.

Co-expression of recombinant Cif C109S and Nedd8 protein in bacteria resulted in the formation of a stable complex *in vitro*. With bacterially co-expressed hexahistidine-tagged Cif C109S and Nedd8, a protein complex was co-purified by nickel-chelating affinity chromatography, and this complex was co-eluted using gel filtration chromatography (Figure 6.1B). The complex is stable in salt concentrations ranging from

0 to 200mM NaCl, but it dissociates into separate components in 500mM NaCl. The co-expression of Cif WT and Nedd8 also yields a complex, however pulls out Nedd8 less efficiently. Furthermore, when refolding both urea denatured Cif and Nedd8, a stable protein complex was formed that co-eluted off the gel filtration column (Figure 6.1C). Interestingly mixing individually purified Cif with Nedd8 does not yield a protein complex (Figure 6.1D). Cif and Nedd8 instead elute separately.

These results suggest that Cif and/or Nedd8 may need to be in a particular conformational state prior to interaction with one another. The gradual refolding of the two proteins together could generate intermediates in a protein conformation that is suited for complex formation. On the other hand, the refolding of these proteins could potentially pose a problem by forcing Cif and Nedd8 into biologically irrelevant conformations, consequently resulting in an artificial complex. However, the co-expression of Cif and Nedd8 in a biological system, with chaperones assisting in proper folding of the proteins, reduces the likelihood of such an event. In fact, in a bacterial co-expression system, the formation of a Cif-Nedd8 complex could be facilitated by bacterial chaperones. The Cif-Nedd8 interaction is possibly mediated by an eukaryotic factor in the mammalian cell.

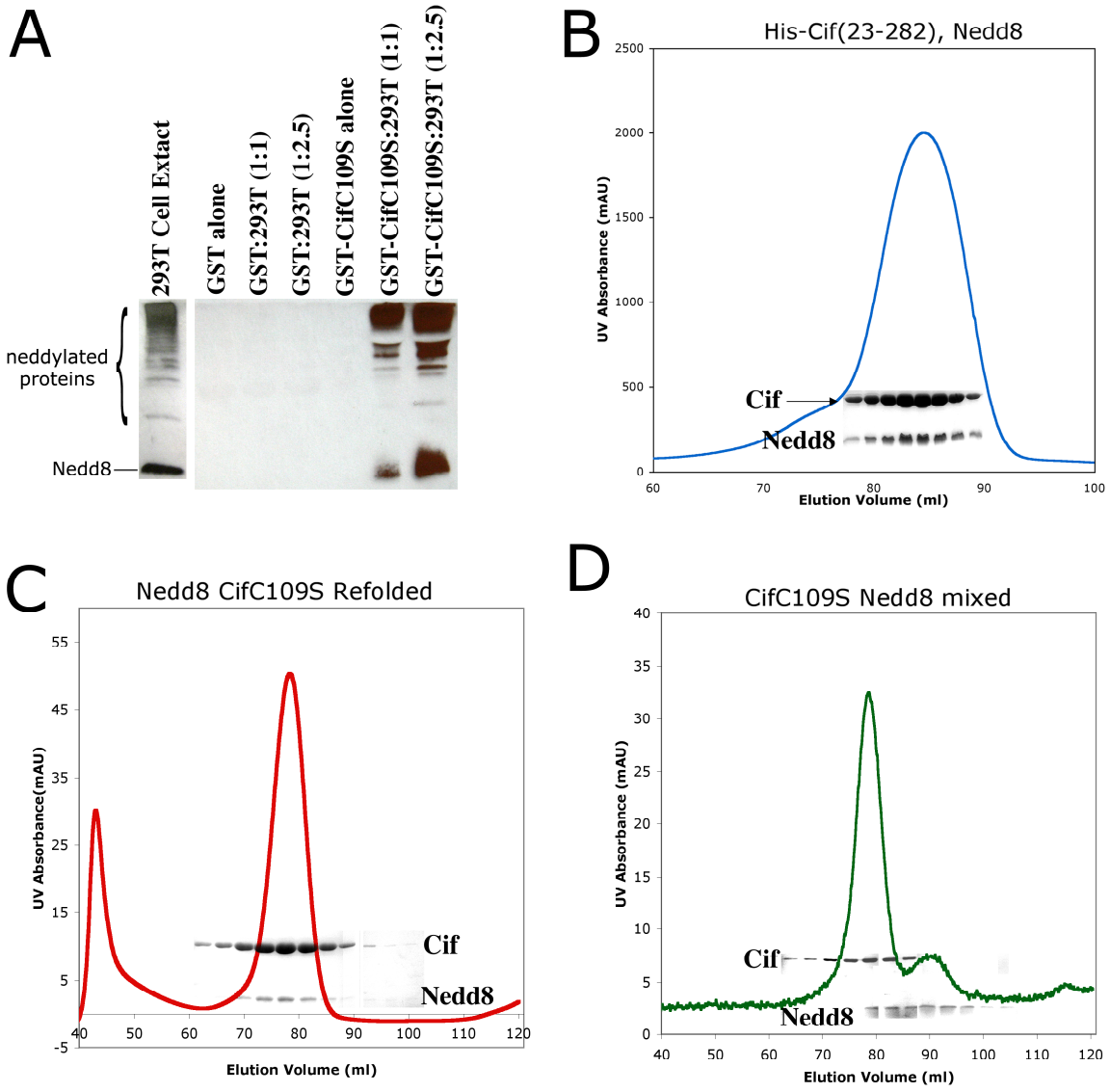


Figure 6.1: Cif interacts with Nedd8 *in vitro*

Figure 6.1: Cif interacts with Nedd8 *in vitro*

(**A**) A GST-Cif C109S pull-down experiment using 293T cell extract expressing FLAG-Nedd8, immunoblotted with anti-FLAG antibody. The ratio of GST/GST-Cif C109S sample to 293T cell extract is indicated. The GST control did not pull out FLAG-Nedd8 from the 293T extract. Using glutathione beads, GST-Cif C109S specifically pulled out FLAG-Nedd8, and higher molecular weight neddylated proteins, proportional to the amount of cell extract incubated in the assay.

(**B**) The gel filtration profile of His-Cif(23-282) in complex with Nedd8. The peak in the chromatograph corresponds to the complex of His-Cif(23-282) and Nedd8. Protein contents from these fractions were visualized using SDS-PAGE gels stained with coomassie blue.

(C) The gel filtration profile of CifC109S in complex with Nedd8 after refolding. The peak in the chromatograph corresponds to the complex of CifC109S and Nedd8. The peak at 45ml elution volumeProtein contents from these fractions were visualized using SDS-PAGE gels stained with coomassie blue.

(D) The gel filtration profile of CifC109S mixing with Nedd8. The peaks in the chromatograph correspond to CifC109S and Nedd8. Protein contents from these fractions were visualized using SDS-PAGE gels stained with coomassie blue.

6.3 Cif N-terminal domain is involved in Nedd8 interaction

Once the binding interaction between Cif and Nedd8 was established, we sought to determine which Cif domain interacted with Nedd8. The N-terminal hexahistidine tagged Cif23-99 or Cif100-282 and Nedd8 were co-expressed in bacteria from a bicistronic plasmid (Chapter Two). The proteins were loaded onto nickel-chelating sepharose to immobilize the hexahistidine tagged Cif and eluted with imidazole wash. The bacterial co-expression trials revealed that the crystallized, protease/acetyltransferase homology domain (100-282) does not bind Nedd8 on its own, whereas an N-terminal peptide spanning residues 23-99 was able to pull down Nedd8 and co-elute on a gel filtration column (Figure 6.3A). This was an interesting finding, suggesting that the N-terminal domain of Cif was critical to the Nedd8 interaction. To ascertain the importance of the N-terminal domain in the cytopathic phenotype, we delivered an N-terminal deletion construct (100-282-His) into HeLa cells using BioPORTER®. This Cif construct was unable to induce cell cycle arrest or stress fiber formation and was indistinguishable from untreated, control cells (Figure 6.3B). Cif's C-terminal domain, containing the active catalytic triad, is therefore not sufficient to induce the cytopathic effect. These findings suggest that Cif's N-terminal domain is functionally important for substrate binding and/or recognition, whereas the C-terminal domain is critical for enzymatic activity.

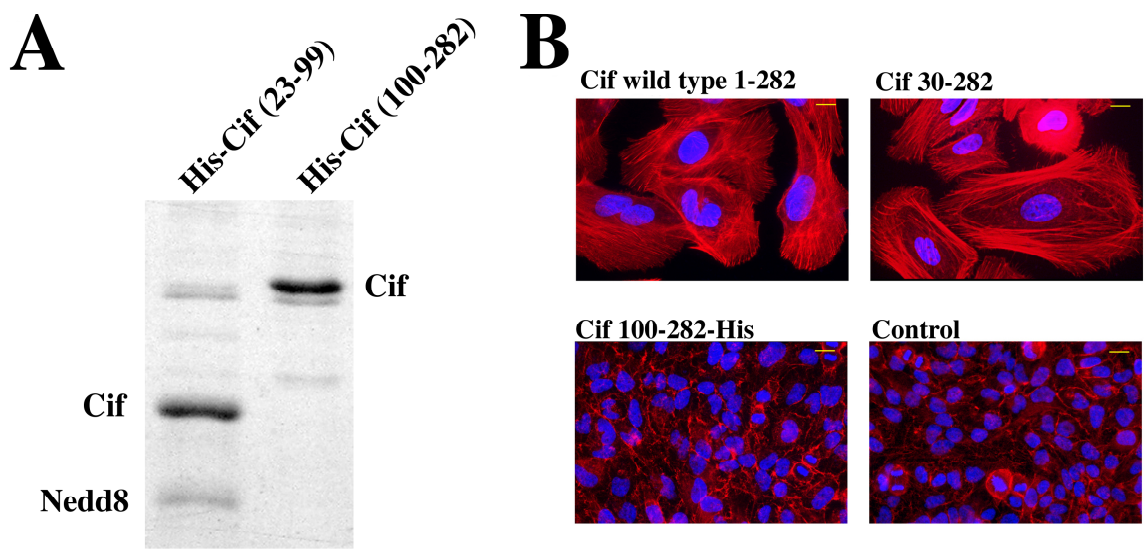


Figure 6.2: The N-terminal domain is required for Cif activity and binds the host protein Nedd8

(A) His-Cif(23-99)/His-Cif(100-282) pulldown on nickel-chelating sepharose of Nedd8 visualized using SDS-PAGE gel stained with coomassie blue. His-Cif(23-99)/His-Cif(100-282) were bacterially co-expressed with Nedd8.

(**B**) The N-terminal domain (30-99) of Cif is required for CPE induction. HeLa cells were exposed 4 hours with the indicated purified Cif protein or PBS (control) in combination with the lipid delivery agent (BioPORTER®) and were further incubated 72 hours. F-actin was stained with phalloidin-rhodamine (red) and DNA with DAPI (blue). Bars correspond to 20 µm. These experiments were performed by Gregory Jubelin and Eric Oswald.

6.4 Cif does not cleave/acetylate Nedd8 or neddylated Cullin-1

The yeast two-hybrid results, the GST pull-down and co-expression experiments, and the established importance of Cif's N-terminal domain for the cytopathic phenotype, all provided evidence for Nedd8 playing an important role in Cif's ability to deregulate the eukaryotic cell cycle. Cif could modulate the cell cycle by sequestering, inhibiting, or activating Nedd8. Thus, we first examine whether Cif could proteolytically cleave Nedd8. Extensive incubations of Cif and GST tagged Nedd8 did not detect any such degradation (data not shown). Cif could also function as a deneddylase, a specific protease, similar to DEN1/NEDP1 that hydrolyzes Nedd8 at a C-terminal Gly-Gly dipeptide, transforming it into a mature conjugable protein (Gan-Erdene et al., 2003; Hochstrasser, 1998; Mendoza et al., 2003; Wu et al., 2003b). To assess whether Cif cleaves Nedd8, we used the substrate Nedd8-AMC, with a C-terminal amidomethylcoumarin, in order to monitor the cleavage reaction (Figure 6.4A). Since Nedd8 shares structural and sequence homology to ubiquitin and SUMO (a ubiquitin-like protein), Cif's activity on ubiquitin-AMC and SUMO-AMC were also examined (Figure 6.4B). Recombinant Cif did not cleave the AMC moiety off any of the ubiquitin-like proteins in this assay while the deneddylase NEDP1 actively cleaved Nedd8 at the C-terminus, releasing the AMC moiety into solution. NEDP1 did not cleave ubiquitin or SUMO indicating this protease exhibits high substrate specificity.

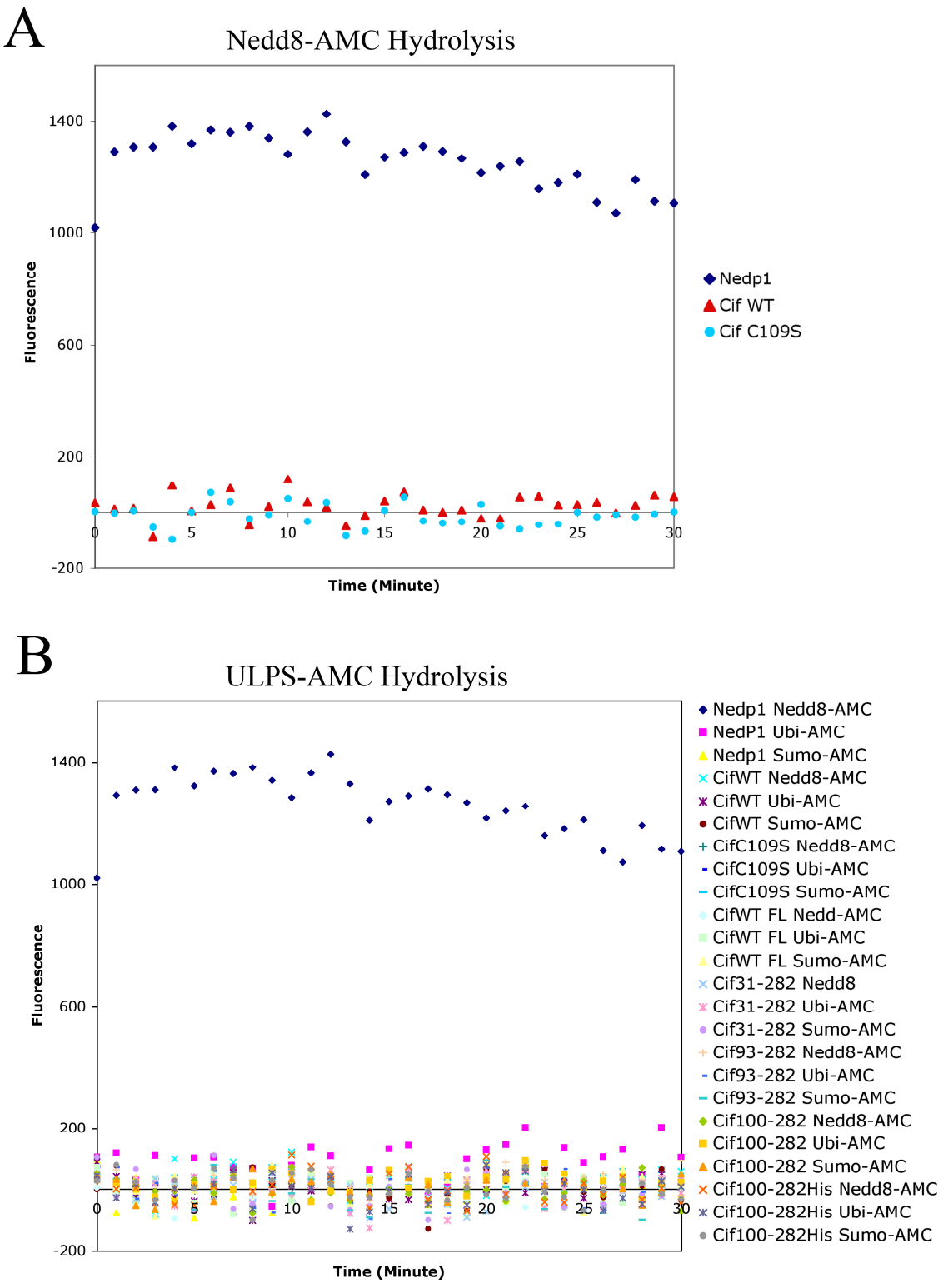


Figure 6.3: Cif does not hydrolyze ubiquitin-like proteins at their C-terminus

Figure 6.3: Cif does not hydrolyze ubiquitin-like proteins at their C-terminus

(A) The ability of Cif WT, Cif C109S, and NEDP1 to perform Nedd8-AMC hydrolysis. Both CifWT and Cif C109S did not cleave Nedd8 at the C-terminus. The deneddylase enzyme, NEDP1, was able to hydrolyze Nedd8-AMC at the C-terminus, yielding a fluorescence signal.

(B) The ability of various Cif constructs to perform ubiquitin-like protein-AMC (ULPs) hydrolysis. The Cif proteins were incubated with ubiquitin-AMC, Nedd8-AMC, or SUMO-AMC. Cif did not cleave any of the ULPs at the C-terminus. NEDP1 was highly specific for Nedd8.

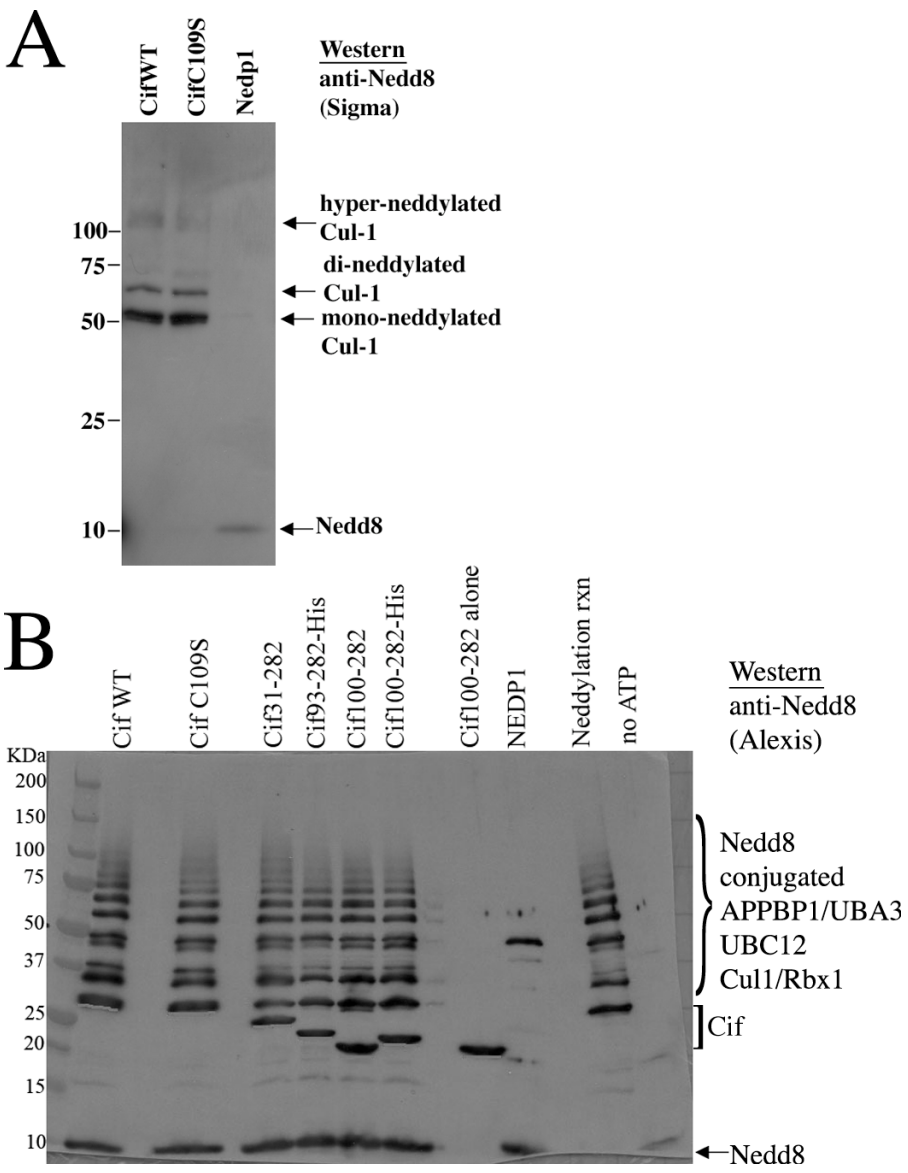


Figure 6.4: Cif does not possess *in vitro* deneddylation activity for Cul1, APPBP1/UBA3, or Ubc12

(A) Deneddylation of neddylated Cul1. Cif WT, Cif C109S, or NEDP1 was incubated with neddylated Cullin-1 at 37°C for 2 hours. The samples were separated by SDS-PAGE and immunoblotted with anti-Nedd8 antibody. Unlike the NEDP1 control, neither Cif WT nor Cif C109S reduces the amount of neddylated Cul1 in the reaction.

(B) Cif was incubated in the neddylation reaction that included APPBP1/UBA3, Ubc12, and Cul1. Cif did not have an inhibitory effect on E1 activating or E2 conjugating activity, whereas NEDP1 successfully inhibited Cul1 neddylation. Cif proteins cross-reacted with the rabbit anti-Nedd8 antibody (Alexis).

Another possible mode of action for Cif is that it only cleaves Nedd8 from Nedd8 conjugated proteins. We reconstituted one well-studied neddylated substrate, Cul1 of the SCF complex, *in vitro* to assay for such activity. The neddylation of Cul1 is a complex reaction that requires the activities of various enzymes: the E1 activating enzyme APPBP11/UBA3, the E2 conjugating enzyme Ubc12, and the SCF subunit Rbx1. The reconstitution of neddylated Cul1 is described in Chapter Two. Neddylated Cul1 in the SCF complex was incubated with either Cif WT or Cif C109S. We did not observe a decrease in the amount of neddylated Cul1 in the reaction containing Cif WT, whereas the control hosting deneddylase NEDP1 was very active in doing so (Figure 6.5A). To see whether Cif has an inhibitory effect on the neddylation of Cul1 by acting on either the E1 or E2 enzyme, Cif was incubated in the Cul1 neddylation assay. As shown in Figure 6.5B, Cul1 was poly-neddylated despite the presence of Cif. Cif does not inhibit E1's Nedd8 activating activity or E2's conjugating activity on Cul1. Meanwhile, NEDP1 successfully prevents the Nedd8 conjugation of Cul1. Finally, given our results that Cif possesses acetyltransferase activity, we examined whether Cif acetylates Nedd8. Cif, acetyl-CoA, and Nedd8 were co-incubated, and the mixture was sent for mass spectrometry analysis. The mass spectrum showed no evidence of Nedd8 acetylation (Figure 6.6).

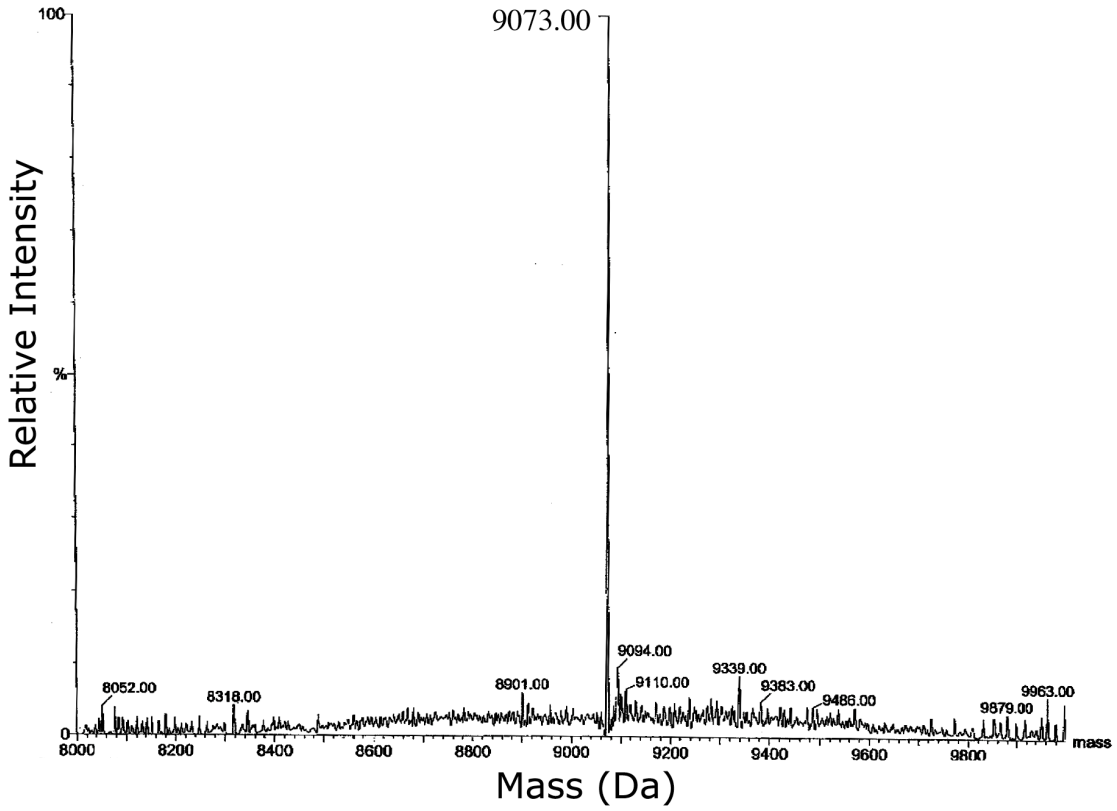


Figure 6.5: Cif does not acetylate Nedd8 *in vitro*

Electrospray ionization mass spectrum of Nedd8 after incubation with Cif WT and acetyl-CoA.

Even though Cif does not cleave or acetylate Nedd8 under our assay conditions, Cif binds Nedd8 with its N-terminal domain, not its catalytic C-terminal domain. It is therefore possible that Cif does not directly act on Nedd8, but instead on Nedd8 associated proteins. This hypothesis was tested by incubating Cif with neddylated Cul1. No change in Cul1's neddylation pattern was observed. This does not eliminate the possibility that Cif interacts with other members of the cullin family (Cullin 2, 3, 4a/b, 5, 6, 7, 8, and Parc) or other yet to be identified Nedd8 modified proteins.

CHAPTER SEVEN:

CONCLUSIONS

Bacterial pathogens utilize a wide array of virulence factors to modulate host cell biochemistry. In particular, the protein substrates of T3SS possess remarkable properties that alter the host cytoskeleton, signal transduction pathways, cell cycle progression, and programmed cell death (Galan and Cossart, 2005). One of these translocated effectors, Cif, from pathogenic *E. coli* strains is associated with a cytopathic effect involving cell cycle arrest and cell distension (Marches et al., 2003). The molecular basis for this activity has remained elusive. Though not all pathogenic *E. coli* strains harbor the *cif* gene, Cif could be advantageous for bacteria in two ways. First, Cif's perturbation of host cellular processes could prolong bacterial colonization by blocking mitosis, which results in the disruption of intestinal epithelium integrity. Second, Cif could possibly impair the immune system by inhibiting clonal expansion of lymphocytes (Cortes-Bratti et al., 2001; Gelfanova et al., 1999; Shenker et al., 2001). Furthermore, with the capability of the *cif*-encoding lambdoid prophage from *E. coli* E22 to propagate to new strains by lysogenic conversion *in vitro* (Loukiadis et al., 2008), it is possible for strains lacking *cif*, such as the deadly *E. coli* O157:H7, to acquire such a virulent attribute.

The Cif crystal structure presented here shows that this T3SS substrate is a divergent member of the cysteine protease superfamily. Cif possesses a conserved core fold containing a putative catalytic triad composed of cysteine, histidine, and glutamine, which aligns well with the catalytic triads of other cysteine proteases and acetyltransferases. It is the first member from its superfamily to possess glutamine as its

third catalytic residue instead of the typical aspartate or asparagine. Mutation of any of these residues results in a Cif molecule that cannot induce cell cycle arrest during infection or when introduced exogenously into cells.

Given its structural resemblance to both cysteine proteases and acetyltransferases, Cif may possess either protease activity or acetyltransferase activity. In fact, as has been shown to be the case with the *Yersinia* effector YopJ (Mukherjee et al., 2006; Orth et al., 2000), it could very well possess both activities. To date, we have been unable to show protease activity of purified recombinant Cif *in vitro*. Cif might possess such a high specificity for its intracellular substrate that it does not cleave a general protease substrate in an *in vitro* assay. On the other hand, as with clostridial neurotoxins and anthrax lethal factor (Pellizzari et al., 1996; Vitale et al., 2000), Cif could interact with a specific structural motif to fully engage its substrate. It is also possible that full-length Cif could be inactive (as a zymogen) and thus requires a eukaryotic host factor for activation through processing or modification. Such activation could resemble that of the auto-proteolytic activity of the *Pseudomonas syringae* effector AvrRpt2, which requires the eukaryotic host factor cyclophilin to induce self-cleavage (Coaker et al., 2005). A unique structural feature to Cif within the cysteine protease superfamily is a loop that occludes part of the active site. It is conceivable that interaction with and/or modification by eukaryotic partner(s) could induce conformational changes allowing more efficient access of the substrate to the catalytic site. On the other hand, Cif may only possess acetyltransferase activity. Cif's activity is not affected by the cysteine protease inhibitor E64, suggesting that the Cif-induced phenotype does not require the activity of a cysteine protease. Cif could also utilize a common catalytic mechanism from cysteine proteases

to accommodate it own enzymatic mechanism. Our results show that it is able to auto-acetylate *in vitro*, and that this activity is tied directly to the active site triad.

The identification of the host cellular targets for Cif will provide the direct means to examine the enzymatic activities suggested by the structure. Towards these efforts, we identified the small, ubiquitin-like molecule Nedd8 as a potential Cif host-binding partner. Nedd8 is a key regulator of host cell cycle progression. It is covalently attached to Cullin-1 of the SCF complex. Given the central role of the SCF complex in cell cycle regulation and the large number of SCF substrates, Cif could act on any of the SCF substrates through its interaction with Nedd8. Our biochemical findings indicate that Nedd8 and Cif do interact to form a stable protein complex. In particular, the N-terminal domain seems to be the stronger contributor to Nedd8 binding. Intriguingly, the N-terminal domain is also absolutely required for the Cif-induced cytopathic effect, showing that both domains of Cif are critical to the function of this virulence factor in the host cell. We speculate that Cif subverts the cellular neddylation machinery to facilitate bacterial colonization.

Currently, our data has not demonstrated any proteolytic activity of Cif on Nedd8 or neddylated Cullin-1. The lack of any detectable enzymatic activity of Cif toward Nedd8 raises the question of what biological function this interaction possesses. Instead of being a direct substrate, Nedd8 could act as an allosteric activator of Cif. For example, Nedd8 could bind Cif and cause a conformational change in the occluding loop that opens the active site for substrate binding. In addition, Nedd8 could simply serve as a targeting protein, bringing Cif to other proteins in the SCF complex or other cullin related complexes. For instance, there is evidence for the accumulation of p21 in cells that are

infected with Cif (personal communication with E. Oswald). p21 is a Cdk inhibitor that prevents cells from progressing into S phase. p21 is ubiquitylated by the SCF complex and then targeted for 26S Proteasome for degradation. Cif could prevent the ubiquitylation of p21 by interacting with any components of the SCF complex via Nedd8.

In summary, Cif belongs to a superfamily of enzymes that includes cysteine proteases and acetyltransferases. Both types of enzymes have been shown to function in bacterial pathogenesis. Mutagenesis of the predicted catalytic triad residues in Cif results in the loss of host cell cycle arrest normally induced by the pathogen, linking the proposed protease or acetyltransferase function from the crystal structure to the cytopathic phenotype. Future studies will focus on finding the host cell substrate(s) for Cif (enzymatic activity) and elucidating its functional role in causing cytopathic effects. Understanding how Cif is able to modulate the host cell cycle for the benefit of the pathogen may also give a new perspective on the mechanisms of host cell cycle regulation.

REFERENCES

(1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr*, **50**, 760-763.

Andreeva, A., Howorth, D., Brenner, S.E., Hubbard, T.J., Chothia, C. and Murzin, A.G. (2004) SCOP database in 2004: refinements integrate structure and sequence family data. *Nucleic Acids Res*, **32**, D226-229.

Axtell, M.J., Chisholm, S.T., Dahlbeck, D. and Staskawicz, B.J. (2003) Genetic and molecular evidence that the Pseudomonas syringae type III effector protein AvrRpt2 is a cysteine protease. *Mol Microbiol*, **49**, 1537-1546.

Bai, C., Sen, P., Hofmann, K., Ma, L., Goebl, M., Harper, J.W. and Elledge, S.J. (1996) SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. *Cell*, **86**, 263-274.

Barrett, A.J. and Rawlings, N.D. (2001) Evolutionary lines of cysteine peptidases. *Biol Chem*, **382**, 727-733.

Bentley, R. and Meganathan, R. (1982) Biosynthesis of vitamin K (menaquinone) in bacteria. *Microbiol Rev*, **46**, 241-280.

Birtalan, S.C., Phillips, R.M. and Ghosh, P. (2002) Three-dimensional secretion signals in chaperone-effector complexes of bacterial pathogens. *Mol Cell*, **9**, 971-980.

Blattner, F.R., Plunkett, G., 3rd, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B. and Shao, Y. (1997) The complete genome sequence of Escherichia coli K-12. *Science*, **277**, 1453-1474.

Blocker, A., Jouihri, N., Larquet, E., Gounon, P., Ebel, F., Parsot, C., Sansonetti, P. and Allaoui, A. (2001) Structure and composition of the Shigella flexneri "needle complex", a part of its type III secreton. *Mol Microbiol*, **39**, 652-663.

Bornstein, G., Bloom, J., Sitry-Shevah, D., Nakayama, K., Pagano, M. and Hershko, A. (2003) Role of the SCFSkp2 ubiquitin ligase in the degradation of p21Cip1 in S phase. *J Biol Chem*, **278**, 25752-25757.

Bray, J. (1945) Isolation of antigenically homogeneous strains of Bact. coli neopolitanum from summer diarrhoea of infants. *J. Pathol. Bacteriol.*, **57**, 239-247.

Buttner, D. and Bonas, U. (2002) Getting across--bacterial type III effector proteins on their way to the plant cell. *Embo J*, **21**, 5313-5322.

Campellone, K.G., Robbins, D. and Leong, J.M. (2004) EspFU is a translocated EHEC effector that interacts with Tir and N-WASP and promotes Nck-independent actin assembly. *Dev Cell*, **7**, 217-228.

Cardozo, T. and Pagano, M. (2004) The SCF ubiquitin ligase: insights into a molecular machine. *Nat Rev Mol Cell Biol*, **5**, 739-751.

Carrano, A.C., Eytan, E., Hershko, A. and Pagano, M. (1999) SKP2 is required for ubiquitin-mediated degradation of the CDK inhibitor p27. *Nat Cell Biol*, **1**, 193-199.

CDC. Update on Multi-State Outbreak of E. coli O157:H7 Infections From Fresh Spinach. National Center for Infectious Diseases.

Charpentier, X. and Oswald, E. (2004) Identification of the secretion and translocation domain of the enteropathogenic and enterohemorrhagic Escherichia coli effector Cif, using TEM-1 beta-lactamase as a new fluorescence-based reporter. *J Bacteriol*, **186**, 5486-5495.

Chaudhuri, R.R., Khan, A.M. and Pallen, M.J. (2004) coliBASE: an online database for Escherichia coli, Shigella and Salmonella comparative genomics. *Nucleic Acids Res*, **32**, D296-299.

Coaker, G., Falick, A. and Staskawicz, B. (2005) Activation of a phytopathogenic bacterial effector protein by a eukaryotic cyclophilin. *Science*, **308**, 548-550.

Cohen-Fix, O., Peters, J.M., Kirschner, M.W. and Koshland, D. (1996) Anaphase initiation in Saccharomyces cerevisiae is controlled by the APC-dependent degradation of the anaphase inhibitor Pds1p. *Genes Dev*, **10**, 3081-3093.

Comayras, C., Tasca, C., Peres, S.Y., Ducommun, B., Oswald, E. and De Rycke, J. (1997) Escherichia coli cytolethal distending toxin blocks the HeLa cell cycle at the G2/M transition by preventing cdc2 protein kinase dephosphorylation and activation. *Infect Immun*, **65**, 5088-5095.

Cortes-Bratti, X., Chaves-Olarte, E., Lagergard, T. and Thelestam, M. (1999) The cytolethal distending toxin from the chancroid bacterium Haemophilus ducreyi induces cell-cycle arrest in the G2 phase. *J Clin Invest*, **103**, 107-115.

Cortes-Bratti, X., Karlsson, C., Lagergard, T., Thelestam, M. and Frisan, T. (2001) The Haemophilus ducreyi cytolethal distending toxin induces cell cycle arrest and apoptosis via the DNA damage checkpoint pathways. *J Biol Chem*, **276**, 5296-5302.

Crepin, V.F., Prasannan, S., Shaw, R.K., Wilson, R.K., Creasey, E., Abe, C.M., Knutton, S., Frankel, G. and Matthews, S. (2005) Structural and functional studies of the enteropathogenic Escherichia coli type III needle complex protein EscJ. *Mol Microbiol*, **55**, 1658-1670.

Dahan, S., Wiles, S., La Ragione, R.M., Best, A., Woodward, M.J., Stevens, M.P., Shaw, R.K., Chong, Y., Knutton, S., Phillips, A. and Frankel, G. (2005) EspJ is a prophage-carried type III effector protein of attaching and effacing pathogens that modulates infection dynamics. *Infect Immun*, **73**, 679-686.

Daniell, S.J., Delahay, R.M., Shaw, R.K., Hartland, E.L., Pallen, M.J., Booy, F., Ebel, F., Knutton, S. and Frankel, G. (2001a) Coiled-coil domain of enteropathogenic Escherichia coli type III secreted protein EspD is involved in EspA filament-mediated cell attachment and hemolysis. *Infect Immun*, **69**, 4055-4064.

Daniell, S.J., Kocsis, E., Morris, E., Knutton, S., Booy, F.P. and Frankel, G. (2003) 3D structure of EspA filaments from enteropathogenic Escherichia coli. *Mol Microbiol*, **49**, 301-308.

Daniell, S.J., Takahashi, N., Wilson, R., Friedberg, D., Rosenshine, I., Booy, F.P., Shaw, R.K., Knutton, S., Frankel, G. and Aizawa, S. (2001b) The filamentous type III secretion translocon of enteropathogenic Escherichia coli. *Cell Microbiol*, **3**, 865-871.

Dealy, M.J., Nguyen, K.V., Lo, J., Gstaiger, M., Krek, W., Elson, D., Arbeit, J., Kipreos, E.T. and Johnson, R.S. (1999) Loss of Cul1 results in early embryonic lethality and dysregulation of cyclin E. *Nat Genet*, **23**, 245-248.

Dean, P. and Kenny, B. (2004) Intestinal barrier dysfunction by enteropathogenic Escherichia coli is mediated by two effector molecules and a bacterial surface protein. *Mol Microbiol*, **54**, 665-675.

Delahay, R.M., Knutton, S., Shaw, R.K., Hartland, E.L., Pallen, M.J. and Frankel, G. (1999) The coiled-coil domain of EspA is essential for the assembly of the type III secretion translocon on the surface of enteropathogenic Escherichia coli. *J Biol Chem*, **274**, 35969-35974.

DeLano, W.L. (2002) The PyMOL User's Manual. *DeLano Scientific, Palo Alto, CA, USA*.

Deshaies, R.J. (1999) SCF and Cullin/Ring H2-based ubiquitin ligases. *Annu Rev Cell Dev Biol*, **15**, 435-467.

Dreyfus, G., Williams, A.W., Kawagishi, I. and Macnab, R.M. (1993) Genetic and biochemical analysis of Salmonella typhimurium FliI, a flagellar protein related to the catalytic subunit of the F0F1 ATPase and to virulence proteins of mammalian and plant pathogens. *J Bacteriol*, **175**, 3131-3138.

DuPont, H.L. (2007) The growing threat of foodborne bacterial enteropathogens of animal origin. *Clin Infect Dis*, **45**, 1353-1361.

Eichelberg, K., Ginocchio, C.C. and Galan, J.E. (1994) Molecular and functional characterization of the Salmonella typhimurium invasion genes invB and invC: homology of InvC to the F0F1 ATPase family of proteins. *J Bacteriol*, **176**, 4501-4510.

Elliott, S.J., Krejany, E.O., Mellies, J.L., Robins-Browne, R.M., Sasakawa, C. and Kaper, J.B. (2001) EspG, a novel type III system-secreted protein from enteropathogenic Escherichia coli with similarities to VirA of Shigella flexneri. *Infect Immun*, **69**, 4027-4033.

EMBL-EBI website: http://www.ebi.ac.uk

Escherich, T. (1886) Die darmbakterien des neugeborenen und säuglings. *Fortschr. Med*, **3**, 515–522; 547–554.

Fan, F. and Macnab, R.M. (1996) Enzymatic characterization of FliI. An ATPase involved in flagellar assembly in Salmonella typhimurium. *J Biol Chem*, **271**, 31981-31988.

Feldman, M.F. and Cornelis, G.R. (2003) The multitalented type III chaperones: all you can do with 15 kDa. *FEMS Microbiol Lett*, **219**, 151-158.

Feldman, R.M., Correll, C.C., Kaplan, K.B. and Deshaies, R.J. (1997) A complex of Cdc4p, Skp1p, and Cdc53p/cullin catalyzes ubiquitination of the phosphorylated CDK inhibitor Sic1p. *Cell*, **91**, 221-230.

Fivaz, M. and van der Goot, F.G. (1999) The tip of a molecular syringe. *Trends Microbiol*, **7**, 341-343.

Frankel, G., Candy, D.C., Everest, P. and Dougan, G. (1994) Characterization of the C-terminal domains of intimin-like proteins of enteropathogenic and enterohemorrhagic Escherichia coli, Citrobacter freundii, and Hafnia alvei. *Infect Immun*, **62**, 1835-1842.

Frithz-Lindsten, E., Rosqvist, R., Johansson, L. and Forsberg, A. (1995) The chaperone-like protein YerA of Yersinia pseudotuberculosis stabilizes YopE in the cytoplasm but is dispensible for targeting to the secretion loci. *Mol Microbiol*, **16**, 635-647.

Fromont-Racine, M., Rain, J.C. and Legrain, P. (1997) Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nat Genet*, **16**, 277-282.

Funabiki, H., Yamano, H., Kumada, K., Nagao, K., Hunt, T. and Yanagida, M. (1996) Cut2 proteolysis required for sister-chromatid seperation in fission yeast. *Nature*, **381**, 438-441.

Furukawa, M., Zhang, Y., McCarville, J., Ohta, T. and Xiong, Y. (2000) The CUL1 C-terminal sequence and ROC1 are required for efficient nuclear accumulation, NEDD8 modification, and ubiquitin ligase activity of CUL1. *Mol Cell Biol*, **20**, 8185-8197.

Galan, J.E. and Collmer, A. (1999) Type III secretion machines: bacterial devices for protein delivery into host cells. *Science*, **284**, 1322-1328.

Galan, J.E. and Cossart, P. (2005) Host-pathogen interactions: a diversity of themes, a variety of molecular machines. *Curr Opin Microbiol*, **8**, 1-3.

Galan, J.E. and Wolf-Watz, H. (2006) Protein delivery into eukaryotic cells by type III secretion machines. *Nature*, **444**, 567-573.

Gan-Erdene, T., Nagamalleswari, K., Yin, L., Wu, K., Pan, Z.Q. and Wilkinson, K.D. (2003) Identification and characterization of DEN1, a deneddylase of the ULP family. *J Biol Chem*, **278**, 28892-28900.

Garmendia, J., Frankel, G. and Crepin, V.F. (2005) Enteropathogenic and enterohemorrhagic Escherichia coli infections: translocation, translocation, translocation. *Infect Immun*, **73**, 2573-2585.

Garmendia, J., Phillips, A.D., Carlier, M.F., Chong, Y., Schuller, S., Marches, O., Dahan, S., Oswald, E., Shaw, R.K., Knutton, S. and Frankel, G. (2004) TccP is an enterohaemorrhagic Escherichia coli O157:H7 type III effector protein that couples Tir to the actin-cytoskeleton. *Cell Microbiol*, **6**, 1167-1183.

Gauthier, A., Puente, J.L. and Finlay, B.B. (2003) Secretin of the enteropathogenic Escherichia coli type III secretion system requires components of the type III apparatus for assembly and localization. *Infect Immun*, **71**, 3310-3319.

Gelfanova, V., Hansen, E.J. and Spinola, S.M. (1999) Cytolethal distending toxin of Haemophilus ducreyi induces apoptotic death of Jurkat T cells. *Infect Immun*, **67**, 6394-6402.

Gruenheid, S., Sekirov, I., Thomas, N.A., Deng, W., O'Donnell, P., Goode, D., Li, Y., Frey, E.A., Brown, N.F., Metalnikov, P., Pawson, T., Ashman, K. and Finlay, B.B. (2004) Identification and characterization of NleA, a non-LEE-encoded type III translocated virulence factor of enterohaemorrhagic Escherichia coli O157:H7. *Mol Microbiol*, **51**, 1233-1249.

Guex, N. and Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714-2723.

Hardwidge, P.R., Deng, W., Vallance, B.A., Rodriguez-Escudero, I., Cid, V.J., Molina, M. and Finlay, B.B. (2005) Modulation of host cytoskeleton function by the enteropathogenic Escherichia coli and Citrobacter rodentium effector protein EspG. *Infect Immun*, **73**, 2586-2594.

Harper, J.W., Burton, J.L. and Solomon, M.J. (2002) The anaphase-promoting complex: it's not just for mitosis any more. *Genes Dev*, **16**, 2179-2206.

Hartland, E.L., Daniell, S.J., Delahay, R.M., Neves, B.C., Wallis, T., Shaw, R.K., Hale, C., Knutton, S. and Frankel, G. (2000) The type III protein translocation system of enteropathogenic Escherichia coli involves EspA-EspB protein interactions. *Mol Microbiol*, **35**, 1483-1492.

Hochstrasser, M. (1998) There's the rub: a novel ubiquitin-like modification linked to cell cycle regulation. *Genes Dev*, **12**, 901-907.

Hofmann, B., Schomburg, D., Hecht, H.-J. . (1993) Crystal Structure of a Thiol Proteinase from Staphylococcus Aureus V-8 in the E-64 Inhibitor Complex. *Acta Crystallography*, **v49**, 1.

Holm, L. and Sander, C. (1995) Dali: a network tool for protein structure comparison. *Trends Biochem Sci*, **20**, 478-480.

Hoover, D.M. and Lubkowski, J. (2002) DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucleic Acids Res*, **30**, e43.

Hori, T., Osaka, F., Chiba, T., Miyamoto, C., Okabayashi, K., Shimbara, N., Kato, S. and Tanaka, K. (1999) Covalent modification of all members of human cullin family proteins by NEDD8. *Oncogene*, **18**, 6829-6834.

Huang, D.T. and Schulman, B.A. (2005) Expression, purification, and characterization of the E1 for human NEDD8, the heterodimeric APPBP1-UBA3 complex. *Methods Enzymol*, **398**, 9-20.

Hueck, C.J. (1998) Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol Mol Biol Rev*, **62**, 379-433.

Ide, T., Laarmann, S., Greune, L., Schillers, H., Oberleithner, H. and Schmidt, M.A. (2001) Characterization of translocation pores inserted into plasma membranes by type III-secreted Esp proteins of enteropathogenic Escherichia coli. *Cell Microbiol*, **3**, 669-679.

Jarvis, K.G., Giron, J.A., Jerse, A.E., McDaniel, T.K., Donnenberg, M.S. and Kaper, J.B. (1995) Enteropathogenic Escherichia coli contains a putative type III secretion system necessary for the export of proteins involved in attaching and effacing lesion formation. *Proc Natl Acad Sci U S A*, **92**, 7996-8000.

Jiang, J. and Struhl, G. (1998) Regulation of the Hedgehog and Wingless signalling pathways by the F-box/WD40-repeat protein Slimb. *Nature*, **391**, 493-496.

Jin, J., Cardozo, T., Lovering, R.C., Elledge, S.J., Pagano, M. and Harper, J.W. (2004) Systematic analysis and nomenclature of mammalian F-box proteins. *Genes Dev*, **18**, 2573-2580.

Jones, T.A., Zou, J.Y., Cowan, S.W. and Kjeldgaard. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A*, **47 (Pt 2)**, 110-119.

Kalman, D., Weiner, O.D., Goosney, D.L., Sedat, J.W., Finlay, B.B., Abo, A. and Bishop, J.M. (1999) Enteropathogenic E. coli acts through WASP and Arp2/3 complex to form actin pedestals. *Nat Cell Biol*, **1**, 389-391.

Kamphuis, I.G., Kalk, K.H., Swarte, M.B. and Drenth, J. (1984) Structure of papain refined at 1.65 A resolution. *J Mol Biol*, **179**, 233-256.

Kamura, T., Koepp, D.M., Conrad, M.N., Skowyra, D., Moreland, R.J., Iliopoulos, O., Lane, W.S., Kaelin, W.G., Jr., Elledge, S.J., Conaway, R.C., Harper, J.W. and Conaway, J.W. (1999) Rbx1, a component of the VHL tumor suppressor complex and SCF ubiquitin ligase. *Science*, **284**, 657-661.

Kanack, K.J., Crawford, J.A., Tatsuno, I., Karmali, M.A. and Kaper, J.B. (2005) SepZ/EspZ is secreted and translocated into HeLa cells by the enteropathogenic Escherichia coli type III secretion system. *Infect Immun*, **73**, 4327-4337.

Kaper, J.B. (1998) Enterohemorrhagic Escherichia coli. *Curr Opin Microbiol*, **1**, 103-108.

Kaper, J.B., Nataro, J.P. and Mobley, H.L. (2004) Pathogenic Escherichia coli. *Nat Rev Microbiol*, **2**, 123-140.

Kee, Y., Lyon, N. and Huibregtse, J.M. (2005) The Rsp5 ubiquitin ligase is coupled to and antagonized by the Ubp2 deubiquitinating enzyme. *Embo J*, **24**, 2414-2424.

Kellogg, D.R. (2003) Wee1-dependent mechanisms required for coordination of cell growth and cell division. *J Cell Sci*, **116**, 4883-4890.

Kenny, B., Ellis, S., Leard, A.D., Warawa, J., Mellor, H. and Jepson, M.A. (2002) Co-ordinate regulation of distinct host cell signalling pathways by multifunctional enteropathogenic Escherichia coli effector molecules. *Mol Microbiol*, **44**, 1095-1107.

Kenny, B. and Jepson, M. (2000) Targeting of an enteropathogenic Escherichia coli (EPEC) effector protein to host mitochondria. *Cell Microbiol*, **2**, 579-590.

King, R.W., Peters, J.M., Tugendreich, S., Rolfe, M., Hieter, P. and Kirschner, M.W. (1995) A 20S complex containing CDC27 and CDC16 catalyzes the mitosis-specific conjugation of ubiquitin to cyclin B. *Cell*, **81**, 279-288.

Knutton, S., Rosenshine, I., Pallen, M.J., Nisan, I., Neves, B.C., Bain, C., Wolff, C., Dougan, G. and Frankel, G. (1998) A novel EspA-associated surface organelle of enteropathogenic Escherichia coli involved in protein translocation into epithelial cells. *Embo J*, **17**, 2166-2176.

Kodama, T., Akeda, Y., Kono, G., Takahashi, A., Imura, K., Iida, T. and Honda, T. (2002) The EspB protein of enterohaemorrhagic Escherichia coli interacts directly with alpha-catenin. *Cell Microbiol*, **4**, 213-222.

Kresse, A.U., Rohde, M. and Guzman, C.A. (1999) The EspD protein of enterohemorrhagic Escherichia coli is required for the formation of bacterial surface appendages and is incorporated in the cytoplasmic membranes of target cells. *Infect Immun*, **67**, 4834-4842.

Kubori, T., Matsushima, Y., Nakamura, D., Uralil, J., Lara-Tejero, M., Sukhan, A., Galan, J.E. and Aizawa, S.I. (1998) Supramolecular structure of the Salmonella typhimurium type III protein secretion system. *Science*, **280**, 602-605.

Kumar, S., Yoshida, Y. and Noda, M. (1993) Cloning of a cDNA which encodes a novel ubiquitin-like protein. *Biochem Biophys Res Commun*, **195**, 393-399.

Lai, L.C., Wainwright, L.A., Stone, K.D. and Donnenberg, M.S. (1997) A third secreted protein that is encoded by the enteropathogenic Escherichia coli pathogenicity island is required for transduction of signals and for attaching and effacing activities in host cells. *Infect Immun*, **65**, 2211-2217.

Lammer, D., Mathias, N., Laplaza, J.M., Jiang, W., Liu, Y., Callis, J., Goebl, M. and Estelle, M. (1998) Modification of yeast Cdc53p by the ubiquitin-related protein rub1p affects function of the SCFCdc4 complex. *Genes Dev*, **12**, 914-926.

Lara-Tejero, M. and Galan, J.E. (2000) A bacterial toxin that controls cell cycle progression as a deoxyribonuclease I-like protein. *Science*, **290**, 354-357.

Laskowski, R.A., Moss, D.S. and Thornton, J.M. (1993) Main-chain bond lengths and bond angles in protein structures. *J Mol Biol*, **231**, 1049-1067.

Lee, B. and Richards, F.M. (1971) The interpretation of protein structures: estimation of static accessibility. *J Mol Biol*, **55**, 379-400.

Lee, E. (2006) E. coli Bedevils Produce Industry. *The Atlanta Journal-Constitution*, Atlanta.

Lee, S.H. and Galan, J.E. (2004) Salmonella type III secretion-associated chaperones confer secretion-pathway specificity. *Mol Microbiol*, **51**, 483-495.

Li, T., Pavletich, N.P., Schulman, B.A. and Zheng, N. (2005) High-level expression and purification of recombinant SCF ubiquitin ligases. *Methods Enzymol*, **398**, 125-142.

Liakopoulos, D., Doenges, G., Matuschewski, K. and Jentsch, S. (1998) A novel protein modification pathway related to the ubiquitin system. *Embo J*, **17**, 2208-2214.

Lilic, M., Vujanac, M. and Stebbins, C.E. (2006) A common structural motif in the binding of virulence factors to bacterial secretion chaperones. *Mol Cell*, **21**, 653-664.

Liu, J., Furukawa, M., Matsumoto, T. and Xiong, Y. (2002) NEDD8 modification of CUL1 dissociates p120(CAND1), an inhibitor of CUL1-SKP1 binding and SCF ligases. *Mol Cell*, **10**, 1511-1518.

Loukiadis, E., Nobe, R., Herold, S., Tramuta, C., Ogura, Y., Ooka, T., Morabito, S., Kerouredan, M., Brugere, H., Schmidt, H., Hayashi, T. and Oswald, E. (2008) Distribution, functional expression, and genetic organization of Cif, a phage-encoded type III-secreted effector from enteropathogenic and enterohemorrhagic Escherichia coli. *J Bacteriol*, **190**, 275-285.

Mainil, J.G., Duchesnes, C.J., Whipp, S.C., Marques, L.R., O'Brien, A.D., Casey, T.A. and Moon, H.W. (1987) Shiga-like toxin production and attaching effacing activity of Escherichia coli associated with calf diarrhea. *Am J Vet Res*, **48**, 743-748.

Mansfield, J., Jenner, C., Hockenhull, R., Bennett, M.A. and Stewart, R. (1994) Characterization of avrPphE, a gene for cultivar-specific avirulence from Pseudomonas syringae pv. phaseolicola which is physically linked to hrpY, a new hrp gene identified in the halo-blight bacterium. *Mol Plant Microbe Interact*, **7**, 726-739.

Marches, O., Covarelli, V., Dahan, S., Cougoule, C., Bhatta, P., Frankel, G. and Caron, E. (2008) EspJ of enteropathogenic and enterohaemorrhagic Escherichia coli inhibits opsono-phagocytosis. *Cell Microbiol*.

Marches, O., Ledger, T.N., Boury, M., Ohara, M., Tu, X., Goffaux, F., Mainil, J., Rosenshine, I., Sugai, M., De Rycke, J. and Oswald, E. (2003) Enteropathogenic and enterohaemorrhagic Escherichia coli deliver a novel effector called Cif, which blocks cell cycle G2/M transition. *Mol Microbiol*, **50**, 1553-1567.

McDaniel, T.K., Jarvis, K.G., Donnenberg, M.S. and Kaper, J.B. (1995) A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *Proc Natl Acad Sci U S A*, **92**, 1664-1668.

McGowan, C.H. and Russell, P. (2004) The DNA damage response: sensing and signaling. *Curr Opin Cell Biol*, **16**, 629-633.

McNamara, B.P. and Donnenberg, M.S. (1998) A novel proline-rich protein, EspF, is secreted from enteropathogenic Escherichia coli via the type III export pathway. *FEMS Microbiol Lett*, **166**, 71-78.

McNamara, B.P., Koutsouris, A., O'Connell, C.B., Nougayrede, J.P., Donnenberg, M.S. and Hecht, G. (2001) Translocated EspF protein from enteropathogenic Escherichia coli disrupts host intestinal barrier function. *J Clin Invest*, **107**, 621-629.

Mendoza, H.M., Shen, L.N., Botting, C., Lewis, A., Chen, J., Ink, B. and Hay, R.T. (2003) NEDP1, a highly conserved cysteine protease that deNEDDylates Cullins. *J Biol Chem*, **278**, 25637-25643.

Michael, W.M. and Newport, J. (1998) Coupling of mitosis to the completion of S phase through Cdc34-mediated degradation of Wee1. *Science*, **282**, 1886-1889.

Moon, H.W., Whipp, S.C., Argenzio, R.A., Levine, M.M. and Giannella, R.A. (1983) Attaching and effacing activities of rabbit and human enteropathogenic Escherichia coli in pig and rabbit intestines. *Infect Immun*, **41**, 1340-1351.

Morimoto, M., Nishida, T., Honda, R. and Yasuda, H. (2000) Modification of cullin-1 by ubiquitin-like protein Nedd8 enhances the activity of SCF(skp2) toward p27(kip1). *Biochem Biophys Res Commun*, **270**, 1093-1096.

Morris, R.J., Perrakis, A. and Lamzin, V.S. (2002) ARP/wARP's model-building algorithms. I. The main chain. *Acta Crystallogr D Biol Crystallogr*, **58**, 968-975.

Moshe, Y., Boulaire, J., Pagano, M. and Hershko, A. (2004) Role of Polo-like kinase in the degradation of early mitotic inhibitor 1, a regulator of the anaphase promoting complex/cyclosome. *Proc Natl Acad Sci U S A*, **101**, 7937-7942.

Mukherjee, S., Keitany, G., Li, Y., Wang, Y., Ball, H.L., Goldsmith, E.J. and Orth, K. (2006) Yersinia YopJ acetylates and inhibits kinase activation by blocking phosphorylation. *Science*, **312**, 1211-1214.

Mundy, R., Jenkins, C., Yu, J., Smith, H. and Frankel, G. (2004a) Distribution of espI among clinical enterohaemorrhagic and enteropathogenic Escherichia coli isolates. *J Med Microbiol*, **53**, 1145-1149.

Mundy, R., Petrovska, L., Smollett, K., Simpson, N., Wilson, R.K., Yu, J., Tu, X., Rosenshine, I., Clare, S., Dougan, G. and Frankel, G. (2004b) Identification of a novel Citrobacter rodentium type III secreted protein, EspI, and roles of this and other secreted proteins in infection. *Infect Immun*, **72**, 2288-2302.

Murray, A.W. (2004) Recycling the cell cycle: cyclins revisited. *Cell*, **116**, 221-234.

Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr*, **53**, 240-255.

Murshudov, G.N., Vagin, A.A., Lebedev, A., Wilson, K.S. and Dodson, E.J. (1999) Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr D Biol Crystallogr*, **55 (Pt 1)**, 247-255.

Nagai, T., Abe, A. and Sasakawa, C. (2005) Targeting of enteropathogenic Escherichia coli EspF to host mitochondria is essential for bacterial pathogenesis: critical role of the 16th leucine residue in EspF. *J Biol Chem*, **280**, 2998-3011.

Nakayama, K.I., Hatakeyama, S. and Nakayama, K. (2001) Regulation of the cell cycle at the G1-S transition by proteolysis of cyclin E and p27Kip1. *Biochem Biophys Res Commun*, **282**, 853-860.

Nesic, D., Hsu, Y. and Stebbins, C.E. (2004) Assembly and function of a bacterial genotoxin. *Nature*, **429**, 429-433.

Nougayrede, J.P., Boury, M., Tasca, C., Marches, O., Milon, A., Oswald, E. and De Rycke, J. (2001) Type III secretion-dependent cell cycle block caused in HeLa cells by enteropathogenic Escherichia coli O103. *Infect Immun*, **69**, 6785-6795.

Nougayrede, J.P. and Donnenberg, M.S. (2004) Enteropathogenic Escherichia coli EspF is targeted to mitochondria and is required to initiate the mitochondrial death pathway. *Cell Microbiol*, **6**, 1097-1111.

Nougayrede, J.P., Taieb, F., De Rycke, J. and Oswald, E. (2005) Cyclomodulins: bacterial effectors that modulate the eukaryotic cell cycle. *Trends Microbiol*, **13**, 103-110.

O'Conner, P.M. and Jackman, J. (1995) Synchronisation of mammalian cells. In Pagano, M. (ed.), *Cell Cycle - Materials and Methods*. Springer-Verlag, New York, pp. 61-74.

Obrig, T.G. (1997) Shiga toxin mode of action in E. coli O157:H7 disease. *Front Biosci*, **2**, d635-642.

Olsen, S.J., MacKinnon, L.C., Goulding, J.S., Bean, N.H. and Slutsker, L. (2000) Surveillance for foodborne-disease outbreaks--United States, 1993-1997. *MMWR CDC Surveill Summ*, **49**, 1-62.

Orth, K., Xu, Z., Mudgett, M.B., Bao, Z.Q., Palmer, L.E., Bliska, J.B., Mangel, W.F., Staskawicz, B. and Dixon, J.E. (2000) Disruption of signaling by Yersinia effector YopJ, a ubiquitin-like protein protease. *Science*, **290**, 1594-1597.

Osaka, F., Kawasaki, H., Aida, N., Saeki, M., Chiba, T., Kawashima, S., Tanaka, K. and Kato, S. (1998) A new NEDD8-ligating system for cullin-4A. *Genes Dev*, **12**, 2263-2268.

Oswald, E., Nougayrede, J.P., Taieb, F. and Sugai, M. (2005) Bacterial toxins that modulate host cell-cycle progression. *Curr Opin Microbiol*, **8**, 83-91.

Passmore, L.A. and Barford, D. (2004) Getting into position: the catalytic mechanisms of protein ubiquitylation. *Biochem J*, **379**, 513-525.

Pellizzari, R., Rossetto, O., Lozzi, L., Giovedi, S., Johnson, E., Shone, C.C. and Montecucco, C. (1996) Structural determinants of the specificity for synaptic vesicle-associated membrane protein/synaptobrevin of tetanus and botulinum type B and G neurotoxins. *J Biol Chem*, **271**, 20353-20358.

Perrakis, A., Morris, R. and Lamzin, V.S. (1999) Automated protein model building combined with iterative structure refinement. *Nat Struct Biol*, **6**, 458-463.

Perrakis, A., Sixma, T.K., Wilson, K.S. and Lamzin, V.S. (1997) wARP: improvement and extension of crystallographic phases by weighted averaging of multiple-refined dummy atomic models. *Acta Crystallogr D Biol Crystallogr*, **53**, 448-455.

Phan, J., Tropea, J.E. and Waugh, D.S. (2004) Structure of the Yersinia pestis type III secretion chaperone SycH in complex with a stable fragment of YscM2. *Acta Crystallogr D Biol Crystallogr*, **60**, 1591-1599.

Pickart, C.M. and Cohen, R.E. (2004) Proteasomes and their kin: proteases in the machine age. *Nat Rev Mol Cell Biol*, **5**, 177-187.

Potterton, E., Briggs, P., Turkenburg, M. and Dodson, E. (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr D Biol Crystallogr*, **59**, 1131-1137.

Rawlings, N.D. and Barrett, A.J. (1994) Families of cysteine peptidases. *Methods Enzymol*, **244**, 461-486.

Rawlings, N.D., Morton, F.R. and Barrett, A.J. (2006) MEROPS: the peptidase database. *Nucleic Acids Res*, **34**, D270-272.

Read, M.A., Brownell, J.E., Gladysheva, T.B., Hottelet, M., Parent, L.A., Coggins, M.B., Pierce, J.W., Podust, V.N., Luo, R.S., Chau, V. and Palombella, V.J. (2000) Nedd8 modification of cul-1 activates SCF(beta(TrCP))-dependent ubiquitination of IkappaBalpha. *Mol Cell Biol*, **20**, 2326-2333.

Rodrigues-Lima, F., Delomenie, C., Goodfellow, G.H., Grant, D.M. and Dupret, J.M. (2001) Homology modelling and structural analysis of human arylamine N-acetyltransferase NAT1: evidence for the conservation of a cysteine protease catalytic domain and an active-site loop. *Biochem J*, **356**, 327-334.

Rost, B., Sander, C. and Schneider, R. (1994) PHD--an automatic mail server for protein secondary structure prediction. *Comput Appl Biosci*, **10**, 53-60.

Saff, E.a.K., ABJ. (1997) *The Mathematical Intelligencer*, **19**, 5-11.

Scheffner, M., Nuber, U. and Huibregtse, J.M. (1995) Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade. *Nature*, **373**, 81-83.

Schubot, F.D., Jackson, M.W., Penrose, K.J., Cherry, S., Tropea, J.E., Plano, G.V. and Waugh, D.S. (2005) Three-dimensional structure of a macromolecular assembly that regulates type III secretion in Yersinia pestis. *J Mol Biol*, **346**, 1147-1161.

Sekiya, K., Ohishi, M., Ogino, T., Tamano, K., Sasakawa, C. and Abe, A. (2001) Supermolecular structure of the enteropathogenic Escherichia coli type III secretion system and its direct interaction with the EspA-sheath-like structure. *Proc Natl Acad Sci U S A*, **98**, 11638-11643.

Seol, J.H., Feldman, R.M., Zachariae, W., Shevchenko, A., Correll, C.C., Lyapina, S., Chi, Y., Galova, M., Claypool, J., Sandmeyer, S., Nasmyth, K., Deshaies, R.J., Shevchenko, A. and Deshaies, R.J. (1999) Cdc53/cullin and the essential Hrt1 RING-H2 subunit of SCF define a ubiquitin ligase module that activates the E2 enzyme Cdc34. *Genes Dev*, **13**, 1614-1626.

Sert, V., Cans, C., Tasca, C., Bret-Bennis, L., Oswald, E., Ducommun, B. and De Rycke, J. (1999) The bacterial cytolethal distending toxin (CDT) triggers a G2 cell cycle checkpoint in mammalian cells without preliminary induction of DNA strand breaks. *Oncogene*, **18**, 6296-6304.

Shao, F., Merritt, P.M., Bao, Z., Innes, R.W. and Dixon, J.E. (2002) A Yersinia effector and a Pseudomonas avirulence protein define a family of cysteine proteases functioning in bacterial pathogenesis. *Cell*, **109**, 575-588.

Shenker, B.J., Hoffmaster, R.H., Zekavat, A., Yamaguchi, N., Lally, E.T. and Demuth, D.R. (2001) Induction of apoptosis in human T cells by Actinobacillus actinomycetemcomitans cytolethal distending toxin is a consequence of G2 arrest of the cell cycle. *J Immunol*, **167**, 435-441.

Sherr, C.J. (1996) Cancer cell cycles. *Science*, **274**, 1672-1677.

Sherr, C.J. and Roberts, J.M. (1999) CDK inhibitors: positive and negative regulators of G1-phase progression. *Genes Dev*, **13**, 1501-1512.

Sinclair, J.C., Sandy, J., Delgoda, R., Sim, E. and Noble, M.E. (2000) Structure of arylamine N-acetyltransferase reveals a catalytic triad. *Nat Struct Biol*, **7**, 560-564.

Skowyra, D., Craig, K.L., Tyers, M., Elledge, S.J. and Harper, J.W. (1997) F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex. *Cell*, **91**, 209-219.

Sorg, I., Goehring, U.M., Aktories, K. and Schmidt, G. (2001) Recombinant Yersinia YopT leads to uncoupling of RhoA-effector interaction. *Infect Immun*, **69**, 7535-7543.

Stebbins, C.E. (2004) Structural insights into bacterial modulation of the host cytoskeleton. *Curr Opin Struct Biol*, **14**, 731-740.

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

Stebbins, C.E. and Galan, J.E. (2003) Priming virulence factors for delivery into the host. *Nat Rev Mol Cell Biol*, **4**, 738-743.

Stemmer, W.P., Crameri, A., Ha, K.D., Brennan, T.M. and Heyneker, H.L. (1995) Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene*, **164**, 49-53.

Sudakin, V., Ganoth, D., Dahan, A., Heller, H., Hershko, J., Luca, F.C., Ruderman, J.V. and Hershko, A. (1995) The cyclosome, a large complex containing cyclin-selective ubiquitin ligase activity, targets cyclins for destruction at the end of mitosis. *Mol Biol Cell*, **6**, 185-197.

Sutterluty, H., Chatelain, E., Marti, A., Wirbelauer, C., Senften, M., Muller, U. and Krek, W. (1999) p45SKP2 promotes p27Kip1 degradation and induces S phase in quiescent cells. *Nat Cell Biol*, **1**, 207-214.

Taieb, F., Nougayrede, J.P., Watrin, C., Samba-Louaka, A. and Oswald, E. (2006) Escherichia coli cyclomodulin Cif induces G2 arrest of the host cell cycle without activation of the DNA-damage checkpoint-signalling pathway. *Cell Microbiol*, **8**, 1910-1921.

Tamano, K., Aizawa, S., Katayama, E., Nonaka, T., Imajoh-Ohmi, S., Kuwae, A., Nagai, S. and Sasakawa, C. (2000) Supramolecular structure of the Shigella type III secretion machinery: the needle part is changeable in length and essential for delivery of effectors. *Embo J*, **19**, 3876-3887.

Terwilliger, T.C. and Berendzen, J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr D Biol Crystallogr*, **55 (Pt 4)**, 849-861.

Tsvetkov, L.M., Yeh, K.H., Lee, S.J., Sun, H. and Zhang, H. (1999) p27(Kip1) ubiquitination and degradation is regulated by the SCF(Skp2) complex through phosphorylated Thr187 in p27. *Curr Biol*, **9**, 661-664.

Tu, X., Nisan, I., Yona, C., Hanski, E. and Rosenshine, I. (2003) EspH, a new cytoskeleton-modulating effector of enterohaemorrhagic and enteropathogenic Escherichia coli. *Mol Microbiol*, **47**, 595-606.

Viswanathan, V.K., Koutsouris, A., Lukic, S., Pilkinton, M., Simonovic, I., Simonovic, M. and Hecht, G. (2004) Comparative analysis of EspF from enteropathogenic and enterohemorrhagic Escherichia coli in alteration of epithelial barrier function. *Infect Immun*, **72**, 3218-3227.

Vitale, G., Bernardi, L., Napolitani, G., Mock, M. and Montecucco, C. (2000) Susceptibility of mitogen-activated protein kinase kinase family members to proteolysis by anthrax lethal factor. *Biochem J*, **352 Pt 3**, 739-745.

Weber, W.W. and Hein, D.W. (1985) N-acetylation pharmacogenetics. *Pharmacol Rev*, **37**, 25-79.

Weissman, A.M. (2001) Themes and variations on ubiquitylation. *Nat Rev Mol Cell Biol*, **2**, 169-178.

Whitehouse, C.A., Balbo, P.B., Pesci, E.C., Cottle, D.L., Mirabito, P.M. and Pickett, C.L. (1998) Campylobacter jejuni cytolethal distending toxin causes a G2-phase cell cycle block. *Infect Immun*, **66**, 1934-1940.

Wilson, R.K., Shaw, R.K., Daniell, S., Knutton, S. and Frankel, G. (2001) Role of EscF, a putative needle complex protein, in the type III protein translocation system of enteropathogenic Escherichia coli. *Cell Microbiol*, **3**, 753-762.

Winn, M.D., Isupov, M.N. and Murshudov, G.N. (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr*, **57**, 122-133.

Winn, M.D., Murshudov, G.N. and Papiz, M.Z. (2003) Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol*, **374**, 300-321.

Woestyn, S., Allaoui, A., Wattiau, P. and Cornelis, G.R. (1994) YscN, the putative energizer of the Yersinia Yop secretion machinery. *J Bacteriol*, **176**, 1561-1569.

Woestyn, S., Sory, M.P., Boland, A., Lequenne, O. and Cornelis, G.R. (1996) The cytosolic SycE and SycH chaperones of Yersinia protect the region of YopE and YopH involved in translocation across eukaryotic cell membranes. *Mol Microbiol*, **20**, 1261-1271.

Wolff, C., Nisan, I., Hanski, E., Frankel, G. and Rosenshine, I. (1998) Protein translocation into host epithelial cells by infecting enteropathogenic Escherichia coli. *Mol Microbiol*, **28**, 143-155.

Wu, G., Xu, G., Schulman, B.A., Jeffrey, P.D., Harper, J.W. and Pavletich, N.P. (2003a) Structure of a beta-TrCP1-Skp1-beta-catenin complex: destruction motif binding and lysine specificity of the SCF(beta-TrCP1) ubiquitin ligase. *Mol Cell*, **11**, 1445-1456.

Wu, K., Chen, A. and Pan, Z.Q. (2000) Conjugation of Nedd8 to CUL1 enhances the ability of the ROC1-CUL1 complex to promote ubiquitin polymerization. *J Biol Chem*, **275**, 32317-32324.

Wu, K., Yamoah, K., Dolios, G., Gan-Erdene, T., Tan, P., Chen, A., Lee, C.G., Wei, N., Wilkinson, K.D., Wang, R. and Pan, Z.Q. (2003b) DEN1 is a dual function protease capable of processing the C terminus of Nedd8 and deconjugating hyper-neddylated CUL1. *J Biol Chem*, **278**, 28882-28891.

Yaron, A., Hatzubai, A., Davis, M., Lavon, I., Amit, S., Manning, A.M., Andersen, J.S., Mann, M., Mercurio, F. and Ben-Neriah, Y. (1998) Identification of the receptor component of the IkappaBalpha-ubiquitin ligase. *Nature*, **396**, 590-594.

Yoshida, S., Katayama, E., Kuwae, A., Mimuro, H., Suzuki, T. and Sasakawa, C. (2002) Shigella deliver an effector protein to trigger host microtubule destabilization, which promotes Rac1 activity and efficient bacterial internalization. *Embo J*, **21**, 2923-2935.

Yu, Z.K., Gervais, J.L. and Zhang, H. (1998) Human CUL-1 associates with the SKP1/SKP2 complex and regulates p21(CIP1/WAF1) and cyclin D proteins. *Proc Natl Acad Sci U S A*, **95**, 11324-11329.

Zheng, J., Yang, X., Harrell, J.M., Ryzhikov, S., Shim, E.H., Lykke-Andersen, K., Wei, N., Sun, H., Kobayashi, R. and Zhang, H. (2002a) CAND1 binds to unneddylated CUL1 and regulates the formation of SCF ubiquitin E3 ligase complex. *Mol Cell*, **10**, 1519-1526.

Zheng, N., Schulman, B.A., Song, L., Miller, J.J., Jeffrey, P.D., Wang, P., Chu, C., Koepp, D.M., Elledge, S.J., Pagano, M., Conaway, R.C., Conaway, J.W., Harper, J.W. and Pavletich, N.P. (2002b) Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. *Nature*, **416**, 703-709.

Zhu, M., Shao, F., Innes, R.W., Dixon, J.E. and Xu, Z. (2004) The crystal structure of Pseudomonas avirulence protein AvrPphB: a papain-like fold with a distinct substrate-binding site. *Proc Natl Acad Sci U S A*, **101**, 302-307.

Zumbihl, R., Aepfelbacher, M., Andor, A., Jacobi, C.A., Ruckdeschel, K., Rouot, B. and Heesemann, J. (1999) The cytotoxin YopT of Yersinia enterocolitica induces modification and cellular redistribution of the small GTP-binding protein RhoA. *J Biol Chem*, **274**, 29289-29293.