12-2019

The manipulation of host transcription by the ANKH effector of legionella.

Juanita E. Von Dwingelo

THE MANIPULATION OF HOST TRANSCRIPTION BY THE ANKH EFFECTOR OF *LEGIONELLA*

By

Juanita E Von Dwingelo
B.Sc., Murray State University, 2010
M.S., Murray State University, 2012
M.S., University of Louisville, 2014

A Dissertation
Submitted to the Faculty of the
School of Medicine of the University of Louisville
In Partial Fulfillment of the Requirements
For the Degree of

Doctor of Philosophy
In Microbiology and Immunology

Department of Microbiology and Immunology
University of Louisville
Louisville, Kentucky

December 2019

THE MANIPULATION OF HOST TRANSCRIPTION BY THE ANKH EFFECTOR

OF *LEGIONELLA*

By

Juanita E. Von Dwingelo
B.Sc., Murray State Univesity, 2010
M.S., Murray State University, 2012
M.S., University of Louisville, 2014

A Dissertation Approved on

August 15, 2019

By the following Dissertation Committee:

_____
Dissertation Director, Dr. Yousef Abu Kwaik

_____
Dr. Richard Miller

_____
Dr. Matthew Lawrenz

_____
Dr. Richard Lamont

_____
Dr. Thomas Mitchell

DEDICATION

I would like to dedicate my dissertation to my parents who have continued to be supportive no matter what, to my husband Adam who has been by my side for almost the entirety of my PhD work and has been my rock and to my son Kai, who will hopefully grow up with a love of all things science.

ACKNOWLEDGMENTS

I would first like to thank my mentor, Dr. Yousef Abu Kwaik – if you had not agreed to take me on as a student I would not be here. I would also like to thank my committee members: Dr. Richard Miller, Dr. Matthew Lawrenz, Dr. Thomas Mitchell and Dr. Richard Lamont – you insight and guidance through my time here has been beyond valuable.

I would like to thank the two amazing friends I started this journey with but was unable to finish with – Dr. Ashley Best and Dr. Nikole Warner. I can't thank you each enough for the help you've provided be it just talking through ideas or being there for a shoulder to cry on when things went wrong. I love you both.

I would like to thank my fellow Abu Kwaik lab members, particularly Dr. Chris Price and Snake Jones – you have both put up with so many stupid questions and annoyances from me for the past 7 years but have always been willing to lend a helping hand. I'm honored that I've had the opportunity to work alongside you both; Bethany Vaughn, you've been an amazing partner in crime; and Hannah Hanford, you're dad jokes are ridiculous but you're a brilliant young scientist.

I would lastly like to thank the friends I've made along the way who have had to deal with my antics; Amanda, Sarah and Amanda (Tom) you have all been great people to work next to and have always been extremely helpful in every way; Claire and Trey for being the best Pokémon Go buddies; Autumn for being the wonderful plant guru that you are; and Drew, where would we all be without your amazing baked goods! You're all wonderful people and I know you all have wonderfully exciting things ahead of you.

ABSTRACT

THE MANIPULATION OF HOST TRANSCRIPTION BY THE ANKH EFFECTOR OF *LEGIONELLA*

Juanita Von Dwingelo

August 15, 2019

Legionella pneumophila is a Gram-negative facultative intracellular bacterium that can be found dispersed throughout freshwater environments, where it primarily parasitizes amoebae and other protozoan species. Humans are an accidental host for *L. pneumophila*, and infection occurs upon inhalation of aerosolized water droplets that contain the bacteria. *L. pneumophila* is the causative agent of Legionnaires' Disease, which is the result of intracellular proliferation within alveolar macrophages. Pathogenesis of *L. pneumophila* is dependent on the Dot/Icm type 4 secretion system (T4SS) apparatus, which is comprised of 27 proteins and is responsible for translocating over 330 effector proteins into the host cell. Many of these effector proteins contain eukaryotic-like domains and motifs, which have been acquired through interkingdom horizontal gene transfer from various aquatic eukaryotic hosts. While *L. pneumophila* contains such a large repertoire of effector proteins, most of them are not required for survival and proliferation in mammalian macrophages, since single deletion of most effectors does not result in a defect in intracellular replication. Although this could be explained by effector redundancy, it is more likely that these effector proteins constitute a tool box utilized by *L. pneumophila* to survive and replicate within numerous species of protozoa. One

effector identified, that when deleted results in a defect in intracellular replication, is the AnkH effector. It has been shown that AnkH is required for robust intracellular replication of *L. pneumophila* within amoebae, human macrophages and the A/J mouse model of infection. It has previously been shown that AnkH is an effector that contains ankyrin repeats, which are eukaryotic-like domains, and function as a scaffold for protein-protein interactions. Other than requirement of AnkH during intracellular replication, its function and host targets remain unknown and are the focus of this work. We further characterized AnkH to elucidate its host target and function during infection of macrophages. Using a yeast 2 hybrid system, seven potential host interacting partners have been identified and one interacting partner, human La related protein 7 (LARP7), has been confirmed via co-immunoprecipitation. LARP7 is a component of a transcriptional regulatory complex, 7SK snRNP complex that negatively regulates transcriptional elongation. The AnkH -LARP7 interaction blocks LARP7 binding to components of the 7SK snRNP complex, resulting in the disruption of the complex. Knockdown of LARP7 using LARP7 specific RNAi results in a significant growth defect of the WT strain during infection of macrophages, and the growth defect of the Δ*ankH* null mutant becomes more severe. RNAseq has been performed on macrophages infected with either WT or Δ*ankH* strains of *L. pneumophila* to determine modulation of transcription during infection. The data show that there are a total of 405 genes that are differentially regulated in cells infected with WT versus the Δ*ankH* mutant. The crystal structure of AnkH has been resolved, and it revealed that AnkH contains 4 ankyrin repeats, 2 asparagine hydroxylation motifs, a cysteine-like protease domain and a cap domain. When residues are substituted within the ankyrin repeats, asparagine

hydroxylation sites and cysteine-like protease domain, a decrease in intracellular replication is observed, indicating these domains are critical for the function of AnkH. A substitution within the β-hairpin loop of the third ankyrin repeat results in diminished LARP7-AnkH interactions, and phenocopies the Δ*ankH* null mutant defect in intracellular growth. Taken together, these data suggest that the β-hairpin loop of the third ankyrin repeat of AnkH interacts with the host LARP7, which disrupts host cell transcription elongation by inhibiting assembly of the 7SK snRNP complex resulting in global modulation of transcription. This interaction is important for the intracellular replication of *L. pneumophila* in human macrophages. The ARDs, asparagine hydroxylation motifs and cysteine-like protease pocket are all required for the function of AnkH in intracellular replication of WT *L. pneumophila*. AnkH is an important effector protein that aids in the survival and replication of *L. pneumophila* in all hosts, the study of which would result in a better understanding of how *L. pneumophila* creates an environment within host cells that supports robust intracellular replication.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1:

INTRODUCTION

INTRODUCTION

Discovery of Legionnaires' Disease

In July of 1976, Philadelphia, PA was host to the bicentennial celebration of the formation of the United States as well as the 56th annual American Legion Convention. The 4-day gathering was attended by more than 2,000 American Legion delegates and was hosted at the Bellevue-Stratford Hotel. After the convention, numerous attendees suffered from pneumonia-like symptoms. In total, 182 of the convention attendees reported symptoms and a total of 34 individuals succumbed to the mysterious disease, nicknamed the "Philly Killer" [1].

At first, it was feared that this disease was caused by a new strain of Influenza. The outbreak prompted a high-profile investigation by the Centers for Disease Control and Prevention (CDC). It was determined early in the investigation that the outbreak was caused by a previously unidentified agent. The investigation lasted close to six months and was worked on by multiple teams of parasitologists, virologists, epidemiologists, bacteriologists, and toxicologists. In January 1977, the causative agent was identified but was called the Legionnaires' Disease bacterium until April 1979. The bacteria was given a name representing the disease caused as well as those who were affected by the first documented outbreak, *Legionella pneumophila* [1-3].

What was unknown when the investigation began was the unique nutritional requirements of *L. pneumophila* that makes it difficult to culture and isolate. Unless a special kind of agar plates are used. Moreover, the bacteria replicate within alveolar macrophages making identification from lung tissue secretion difficult. The use of guinea

pigs by the CDC led to the identification of *L. pneumophila* since guinea pigs are susceptible to infection with *Legionella* [2].

The *Legionella* genus consists of approximately 65 species with half (30 species) known to cause disease [4]. *L. pneumophila* remains the primary causative agent of Legionnaires' disease globally, causing 95% of reported cases. However, in Australia *L. longbeachae* causes most reported cases of disease [5, 6]. *L. pneumophila* can be subdivided into 16 serogroups with most of confirmed cases caused by serogroup 1 (Lp1) [4, 7, 8].

Legionnaires' Disease

Infection with *L. pneumophila* results in two distinct clinical manifestations – Pontiac Fever or Legionnaires' Disease [4, 9]. Pontiac Fever is a mild, self-limiting flu like illness which usually resolves in 2 to 5 days and does not benefit from any treatment with antibiotics. Legionnaires' disease can be a multisystem disease and is the pneumonic form of legionellosis that has a case fatality rate of 10%. In immunocompromised and immunosuppressed patients mortality is increased as much as 25% [9]. Immune compromised individuals and smokers are more susceptible to Legionnaires' disease but healthy individuals are also at risk for contracting the disease [10, 11]. The incubation period for both forms of disease varies with symptoms surfacing anywhere from 2-14 days after inoculation. The symptoms of Legionnaires' disease include cough, fever, headache, shortness of breath and muscle pains. Patients with a more severe form of the disease may show symptoms including diarrhea, bloody sputum, ataxia, vomiting, and/or loss of appetite [12]. Death is usually the result of multi organ failure or respiratory shock

[13]. No vaccines exist to protect from Legionnaires' Disease but the disease can be successfully treated with antibiotics including macrolides, tetracyclines, and quinolones, which to date are the most effective [14-16]. Legionnaires' disease cannot be successfully treated with penicillin and β-lactams as a result of the resistance of *L. pneumophila* to the antibiotics. Treatment with these antibiotics also leads to an increased mortality.

Legionnaires' disease is likely underreported in many countries because of a lack of diagnostics and surveillance systems [9, 17]. In 2016 alone, the CDC reported 6,100 confirmed cases of Legionnaires' disease in the United States and acknowledged that this number may be higher because of undiagnosed disease [18]. In many cases, when patients present with pneumonia, they are treated with antibiotics and no lab diagnostic tests are performed to determine the causative agent of the pneumonia. Urine-ELISA assays are needed to confirm *Legionella* infections but are becoming more commonly performed [19-21]. Without consistent patient testing, it is difficult to confirm the number of Legionnaires' disease cases annually. Importantly, roughly 50% of pneumonia cases are of unknown etiology [21], suggesting that *L. pneumophila* may be responsible for more cases then is currently appreciated.

Epidemiology of *L. pneumophila*

L. pneumophila are aquatic organisms globally distributed and natural bodies of water serve as the natural reservoir. As a result, Legionnaires' disease remains an important public health problem worldwide. Outbreaks of *L. pneumophila* is thought to have emerged in the 20th century because of alterations to the environment by humans that generate water aerosols that act as a vehicle to transmit *L. pneumophila* from

different water sources [4]. Some of these sources include air conditioning systems, cooling towers, grocery store misters, humidifiers, and hot tubs[4]. Generally, infection starts with the inhalation of contaminated aerosolized water droplets [22-25]. Until recently, it was widely accepted that *L. pneumophila* was exclusively transmitted in this manner [27]. A single recent case in Europe is the only report of person-to-person transmission [27]. The conditions surrounding this case aided in transmission from one person to another. In this case, an individual was taking care of a seriously ill close relative where frequent and lengthy exposure occurred leading to the transmission of *L. pneumophila* [26-28]. Thus, water serves as the natural reservoir for *L. pneumophila* and serves as the only source of transmission.

L. pneumophila can be controlled in water handling systems with proper maintenance. While this is simple enough, many water handling systems and water holding units are not properly cared for. This is a wide-spread problem which was illustrated by a study through the CDC that identified *Legionella* DNA in 84% of cooling towers tested in the United States [29]. For eradication, continual water treatment is required. Treatments include keeping hot tank water temperatures above 55 ºC and treatment with either monochloramine, chlorine dioxides, or copper-silver ions [30-32]. Short term interventions are the common method for treatment of contaminated water sources including biocides, overheating of water or single treatments using UV irradiation, but these methods are not successful for eradicating the bacteria from water sources [30, 31, 33].

Ecology of *L. pneumophila* and its Adaptation to Protoza

The natural reservoir for *L. pneumophila* is water; and this bacterium has been found in many freshwater environments, and in many man-made water systems, in close association with freshwater protozoa. Legionnaires' disease has only recently emerged because of human alterations to the environment which result in optimal conditions to support replication of the organism [4, 34]. When *L. pneumophila* are left in their natural aquatic environment it is unlikely that they would cause disease; and natural water environments have never been implicated in Legionnaires' disease outbreaks [4].

Protozoa are an important reservoir for *L. pneumophila,* and in aquatic environments, these bacteria parasitize and replicate within amoebae. There are 17 known species of amoebae and 7 species of non-amoebal protozoa that are capable of supporting *L. pneumophila* growth [35-49]. *L. pneumophila* infects the trophozoite form of amoebae and serves to protect the bacteria [50]. Amoebae do not only play an important role in enhancing the pathogenicity of *L. pneumophila,* enable the bacteria to persist in the environment thereby contributing to the pathogenesis of *Legionella.*

The ability to infect human macrophages is hypothesized to be a consequence of the prior adaptation of *L. pneumophila* to the intracellular life within the various protozoan hosts [51]. Respirable sized vesicles are released from protozoa, which contain bacteria that are highly resistant to biocides while the vesicles themselves are resistant to sonication and freezing [52]. When released from a protozoan host, *L. pneumophila* exhibits an enhanced ability to infect mammalian cells as well as being more invasive to cells [53]. *L. pneumophila* grown in protozoa show changes in biochemistry, physiology, and virulence potential relative to those grown *in vitro* [54].These changes include an

increased resistance to antibiotics, biocides, disinfectants and harsh conditions as well as altered fatty acid and protein profiles, decrease in size and motility, an increased ability to infect amoeba and mammalian cells, an increase in environmental fitness, and an increase in uptake via coiling phagocytosis [55-62]. During outbreaks, *L. pneumophila* and amoebae have been isolated from the same source of infection and these amoebae have been shown to support the growth and replication of *L. pneumophila* [63]. Some *L. pneumophila* that cannot be isolated using classical culturing methods have been culturable if in the presence of protozoa [51].

When conditions become unfavorable, protozoa can differentiate from their trophozoite form into a cyst form that protects the organisms and ensures their survival. *L. pneumophila* has also been shown to survive within amoebic cysts [64].This differentiation is a highly resistant developmental stage for the amoebae and contributes to the resistance of *L. pneumophila* to different chemical and physical agents [65].

Environmental stress plays an important role in the transition of *L. pneumophila* from environmental bacteria to an intracellular pathogen [66]. The relationship between *L. pneumophila* and amoebae plays an important role in the pathogenicity of the bacterium [67]. Contributing to the pathogenesis of *Legionella*, a great deal of evidence shows that growth of bacteria in amoebae also plays a role in transmission. First, there is no transmission of *L. pneumophila* between individuals under normal circumstances. Second, the number of free bacteria isolated from the environmental sources of Legionnaires' disease infections is usually low or undetectable [68]. Third, protozoa release respirable sized vesicles that contain *L. pneumophila* [52]. Fourth, the bacteria exhibit an enhanced ability to infect mammalian cells after being released from a

protozoa host [53]. Perhaps because *L. pneumophila* that has been grown in amoebae is more motile and invasive [52]. Fifth, the bacteria grown in a protozoan host show increased resistance to chemical disinfectants, biocides and antibiotics, which makes the bacteria better at establishing disease then free-living amoebae [60-62]. Sixth, bacteria also show an increased resistance to harsh conditions compared to those grown *in vitro* [59]. Seventh, bacteria and amoebae have been isolated from the same source of infection during outbreaks [63]. Lastly, *L. pneumophila* that cannot be cultured using classical methods can be cultured if they are co-cultured with protozoa [51, 69].

There are *Legionella*-like species that cannot be grown on bacteriologic media but must be co-cultured with protozoa and are referred to as *Legionella*-like amoebal pathogens (LLAP). LLAPs are closely related to *Legionella* phylogenetically and acquired their name because of their ability to infect and multiply within amoebae [70]. The genes that code rRNA in bacteria are highly conserved and are often used to compare the relatedness of different organisms. When comparing LLAPs rRNAs to those of *L. pneumophila*, LLAP rRNA shows 91.6-95.8% similarity to *L. pneumophila* rRNA indicating there is a phylogenetic relationship between the two organisms [70]. The LLAPs play a role in community-acquired pneumonia, usually as a co-pathogen and rarely as the sole pathogen [71]. LLAPs are remain a mystery and future studies sre needed to gain a better understanding of the significance of these organisms to human health.

Numerous methods have been employed to attempt to eradicate *L. pneumophila* from aquatic environments. These attempts, which include chemical biocides, overheating water and UV irradiation, have been successful for short periods after which

the bacteria can be again detected. It has been suggested that to eradicate *L. pneumophila* from aquatic environments treatments should be continuous and effective against both the bacteria and the protozoa host [51, 54, 65]. These findings support the hypothesis that amoeba play a key role in the ecology and pathogenesis of *L. pneumophila* and demonstrate the close and unique relationship between the two organisms.

Intracellular life cycle of *L. pneumophila* within Amoebae and Macrophages

The infection of human phagocytic cells occurs when an individual inhales contaminated aerosolized water [4]. Once *L. pneumophila* infects a human host it enters alveolar macrophages where the intracellular life cycle is strikingly similar to the life cycle observed when amoebae engulf *L. pneumophila* (Figure 1-1) [4]. The mode of uptake for both macrophages and amoeba has been described as coiling phagocytosis [72, 73]. Once inside the host cell the bacteria can be found inside a unique replicative vacuole whose biogenesis does not follow the endosomal-lysosomal degradation pathway [74]. This vacuole is termed the *L. pneumophila* containing vacuole (LCV). The LCV is associated with ribosome-studded membranes of the host cell endoplasmic reticulum [4]. Within the LCV, *L. pneumophila* replicates in high numbers which causes the LCV to rupture releasing the bacteria into the host cell cytosol where another 1-2 rounds of replication occur. During these final stages of replication, the bacteria become flagellated and virulent [51, 75, 76]. The final stage in the *L.pneumophila* intracellular lifecycle is lytic of the host cell and release of bacteria [77-79]. This cycle is repeated once the bacteria infects new host cells in the lungs.

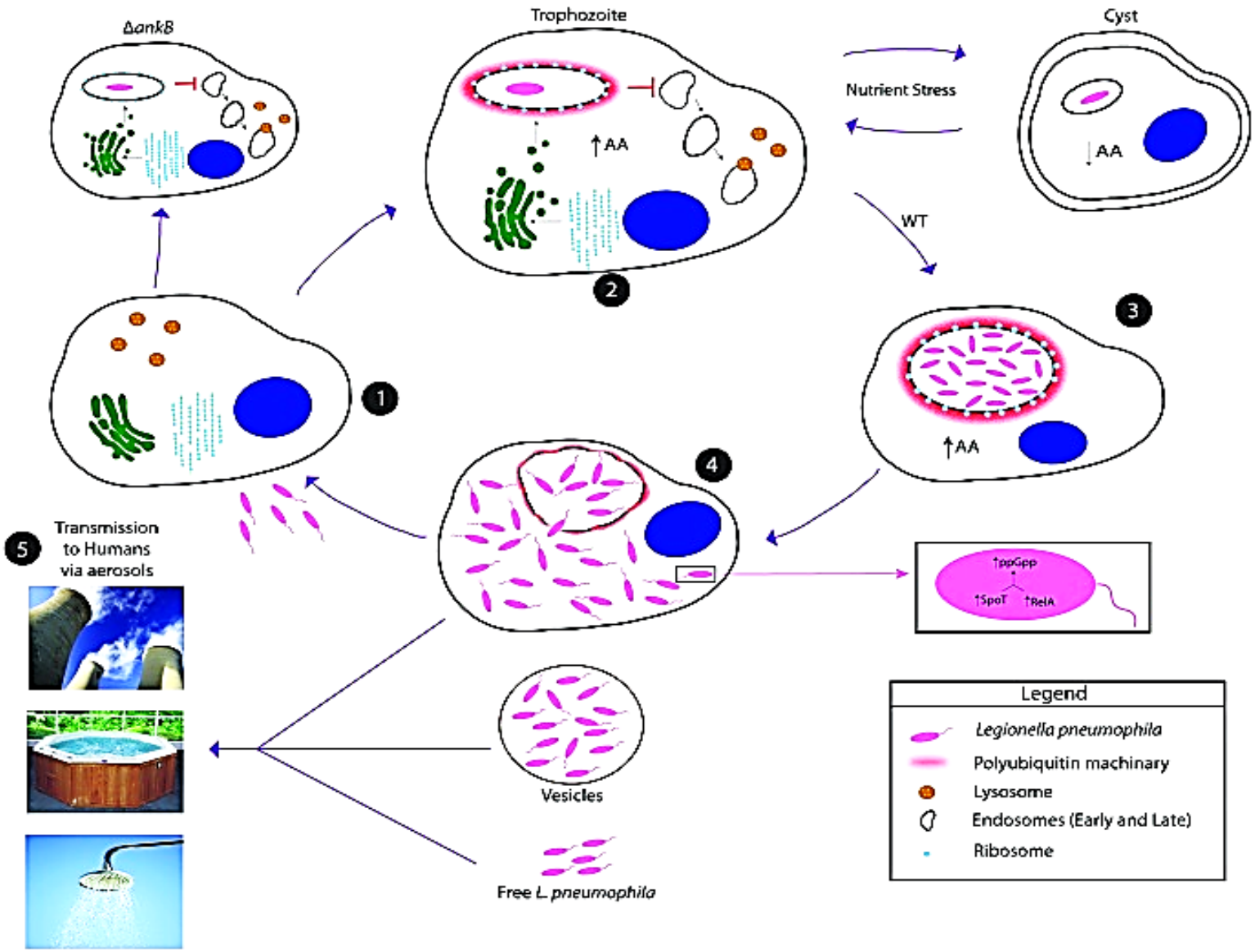
ΔankB
Trophozoite
Cyst
Nutrient Stress
↑AA
↓AA
WT
1
2
3
4
↑AA
5
Transmission
to Humans
via aerosols
↑ppGpp
↑SpoT
↑RelA
Vesicles
Free L. pneumophila
Legend
Legionella pneumophila
Polyubiquitin machinary
Lysosome
Endosomes (Early and Late)
Ribosome

Figure 1-1: The environmental life cycle of *L. pneumophila*. (1) Flagellated *L. pneumophila* infect protozoa in the aquatic environment. (2) The LCV evades the default endosomal–lysosomal degradation pathway and becomes rapidly remodeled by the ER through intercepting ER-to-Golgi vesicle traffic and becomes rapidly decorated with polyubiquitinated proteins in an AnkB-dependent manner. (3) Under unfavorable stress conditions, such as nutrient deprivation, amoebae encyst and bacterial proliferation will not occur due to nutrient limitation. Under growth-permissive conditions for the amoebae, the LCV is decorated with polyubiquitinated proteins, which are targeted for proteasomal degradation leading to elevated cellular levels of amino acids (AA) that power bacterial proliferation of the wild-type strain, while the ankB mutant is defective in this process and is unable to grow despite formation of ER-remodeled replicative LCV. (4) During late stages of infection, the LCV becomes disrupted leading to bacterial egress into the cytosol where the last 1–2 rounds of proliferations are completed. Upon nutrient depletion (see magnified box), RelA and SpoT are triggered leading to increased levels of ppGpp, which triggers phenotypic transition into a flagellated virulent phenotype followed by lysis of the amoeba and bacterial escape from the host cell. Excreted vesicles filled with bacteria are also released. The infectious particle is not known but may include excreted *Legionella*-filled vesicles, intact *Legionella*-filled amoebae, or free *Legionella* that have been released from host cell. (5) Transmission to humans occurs via aerosols generated from man-made devices and installations, such as cooling towers, whirlpools, and showerheads [80]. Adapted from Richards *et al* 2013 [79].

Biphasic Life Cycle of *L. pneumophila*

The intracellular lifecycle of *L. pneumophila* consists of a replicative phase, within the LCV, and a transmissive phase, exhibited upon escape into the cytosol [75, 81-83]. This biphasic lifestyle is characterized by dramatic changes in gene expression and phenotypes [84, 85]. During the replicative phase, the bacterium is undergoing exponential (E) growth, it is non-motile, avirulent, sodium resistant and represses its transmissive traits [86, 87]. A ‘stringent-like’ response pathway is triggered upon transition of *L. pneumophila* into post-exponential (PE) growth. The bacteria become virulent, cytotoxic, motile, and capable of lysosomal evasion. These changes are necessary to invade a new host and start a second cycle of proliferation [58, 88].

Transcriptional analysis of L. pneumophila during infection of *Acanthamoeba castellanii* showed that these two phases exist both *in vivo* and *in vitro* [85].

Replicative to transmissive phase transition is triggered by nutrient limitation and is a highly orchestrated event involving many factors [75, 76, 89]. Upon amino acid depletion, uncharged tRNAs activate RelA to synthesize the bacterial alarmone 3',5'-bispyrophosphate (ppGpp), a master regulator of numerous genes of *L. pneumophila* pathogenesis, which triggers differentiation into the post exponential (PE) phase [90, 91]. RpoS and several global response two-component regulators, such as LetA/S [76, 92-94], are required for phenotypic transition at the PE phase while the RNA-binding protein CsrA acts as a global repressor of the transition and needed later for replication [95]. Two small, non-coding RNAs, RsmY and RsmZ, are induced by LetA at the stationary phase to relieve the repression of CsrA from target genes, required to avoid lysosomal degradation [76, 87]. The ppGpp synthetase RelA monitors amino acid availability through its association with the ribosome [96] and works in conjunction with SpoT, a bifunctional synthetase/hydrolase that responds to fatty acid starvation, to control levels of ppGpp [96]. DskA, a RNA polymerase (RNAP) secondary channel interacting protein, mediates the physiological effects of ppGpp through interactions with RNAP [97]. Without DskA, *L. pneumophia* is defective in stationary phase survival, flagellar gene activation, lysosomal avoidance, and macrophage cytotoxicity [96].

The Type II Secretion System (T2SS)

There are 8 secretion systems that Gram-negative bacteria possess that permits the export of bacterial proteins from within the bacteria to the target host cell or into the

extracellular milieu [98]. *L. pneumophila* codes for two distinct secretion systems, type 2 and type 4, both of which contribute to the pathogenesis of the bacterium [99-101]. The Type 2 secretion system (T2SS) exists in many Gram-negative bacteria including both plant and animal pathogens.

The T2SS is composed of 12 core proteins with 4 subcomplexes. The first subcomplex is an outer membrane "secretin" which provides a pore through the membrane, second is an inner membrane platform which provides a connection to the secretin, third is a cytoplasmic ATPase which gets recruited to the inner membrane platform, and lastly is a periplasm-spanning pseudopilus [99, 102-104]. The T2S consists of a two-step process where proteins that are destined to be secreted are first trafficked into the periplasm, across the inner membrane of the bacteria by the Sec pathway or the Tat pathway [102]. The second step is responsible for secreting proteins that are recognized by the secretion apparatus to the extracellular milieu via an outer membrane pore [102]. The *L. pneumophila* T2SS is important for intracellular infection in host cells and amoebae, as well as growth in mouse models of disease [105-108]. Nearly all pathogens that express the T2SS system exist within aquatic and soil environments in addition to their higher organism hosts [102]. The Legionella type 2 secretion (Lsp) system, plays a major role in the infection of amoeba and is involved in promoting bacterial replication in at least four genera of amoebae [46, 100, 104, 109, 110]. The T2SS system functions at a temperature range of 22-37°C, temperatures commonly associated with aquatic niches thus implicating the T2SS system as being necessary for *L. pneumophila* survival in the environment [107].

To date, 25 proteins have been identified as substrates of the T2SS, many of which are enzymes responsible for degrading proteins and lipids as well as some proteins with novel functions [102]. The effector substrates of the T2SS system increase the likelihood of infection of amoebae with *L. pneumophila*. These effectors include the acyltransferase PlaC, ribonuclease SrnA, metalloprotease ProA, and two novel proteins – NttA and NttC [46, 104, 105, 111]. Each effector is important for infection and their importance varies depending on the species of amoebae. This suggests that the repertoire of *L. pneumophila* effectors secreted by the T2SS system has evolved to enhance the broad host range of this bacterium.

Various studies have considered the importance of the T2SS system in relation to its ability to survive in aquatic environments either as part of a multi-organismal biofilms or planktonically [106, 111-115]. One study from Söderberg *et al* [106] has shown that T2SS mutants show a decreased ability to survive extracellularly in tap water samples that have been incubated at a temperature range of 4-25°C [107, 112]. The secretome of *L. pneumophila* changes in relation to temperature changes, which suggests there are effectors secreted by the T2SS that facilitate survival in low temperatures [107, 112]. It has also been shown that mutants lacking the T2SS Lcl protein are not able to form biofilms as efficiently as bacteria containing a functional T2SS system [113]. Lastly, it was shown that T2SS systems mutants have impaired gliding motility, which is likely the result of an inability to secret a novel surfactant [114-116].

While evidence shows the T2SS is important for infection of amoebae and persistence in the aquatic environment, it also aids in *L. pneumophila* growth within the lung. For example, T2SS mutants show impaired growth within the mouse and Guinea

pig disease models [106, 109, 117]. Studies have shown that the T2SS is not required for entry of *L. pneumophila* into alveolar macrophages, nor is it required for the evasion of the host phagosome-lysosome degradation pathway but is required for bacterial replication at 4-8 hours post infection and for the ability to replicate to large numbers within the LCV at and beyond 12 hours post infection [108]. The T2SS is responsible for dampening the cytokine output of infected macrophages as well as the secondary host during infection, epithelial cells [117]. The T2SS suppression of the innate immune response is hypothesized to limit inflammatory cell infiltrates into the lung initially, which aids in prolonged bacterial growth [118]. The intracellular localization of T2SS substrates is unknown for many of the substrates secreted via this system [119]. It has been suggested that T2SS are not restricted to the LCV lumen due to the observation that T2SS mutants show an impaired ability to retain Rab1B on the LCV, which suggests that a T2SS substrate may exit the LCV where it can engage host GTPases in the cytoplasm [119]. It has also been shown that T2SS mutants trigger immune response pathways, like MyD88 and Toll Like receptor 2, resulting in an increase in inflammatory cytokine levels and suggesting T2SS effector proteins dampen cytosolic innate immunity sensors [118, 120]. A recent study by Truchan *et al* elucidated the cellular localization of 2 T2SS effectors, ProA (a metalloproteinase) and ChiA (chitinase), which escape the LCV and then form a ring-like pattern around the LCV membrane in the host cell cytosol [121]. These studies help to shift views from the predominant paradigm in the *L. pneumophila* field, which is that only type 4 secretion system (T4SS) effectors are able gain access to the host cell cytoplasm during infection [121].

The Dot/Icm Type 4 Secretion System

While there is still much to learn about the T2SS, more is known about the T4SS. The Dot/Icm T4SS system is an important virulence system that is involved in almost all aspects of the intracellular biology of *L. pneumophila* [122]. This set of roughly 30 genes were named *dot* (defective in organelle trafficking) or *icm* (intracellular multiplication) [123-129]. The Dot/Icm system, which is classified as a T4SS system because of its similarity with conjugation systems, is composed of 27 proteins that form a syringe-like apparatus responsible for translocating effector proteins from the inside the bacteria across the LVC membrane into the host cell cytosol [130, 131]. This system is responsible for translocating ~330 effector proteins into the host cell cytoplasm with the help of a secretion system coupling complex comprised of DotL, DotM, and DotN [132]. DotL, DotM, and DotN form a coupling complex that is responsible for recruiting protein substrates in the bacterial cytoplasm and delivering them to the translocation channel in the inner membrane [132]. This complex works in association with three chaperone proteins, IcmW, IcmS and LvgA, which recruit some of the substrates to DotL via a 20-amino acid region translocation sequence located within the C terminal portion of the protein [130, 132-136]. This need for a chaperone is not the case for all effectors as some can still bind to the coupling complex without the help of the chaperone proteins [132].

The ~330 substrates of the Dot/Icm system are known as effector proteins, which accounts for approximately 10% of the *L. pneumophila* genome coding capacity [137-140]. *L. pneumophila* contains the largest repertoire of effector proteins, followed by *Coxiella burnetii* which contains ~100 effector proteins [141]. Some of the *L. pneumophila* effector proteins are translocated upon attachment of the bacterium to a host

cell, but the translocation of these proteins occurs throughout the intracellular growth phase of the bacteria [142, 143]. RalF was the first effector shown to be translocated from the bacteria to the host cell in a Dot/Icm dependent manner [144]. Since the discovery of RalF, numerous other effectors have been identified that have been shown to play roles in modifying host cell processes in order to establish a replicative niche that supports robust replication of *L. pneumophila*. Some of the better characterized examples of effectors whose functions have been identified play roles in vesicular trafficking pathways (LidA & AnkX), host protein synthesis (Lgt1 & Sidl), cell apoptosis (SdhA & SidF), and host ubiquitination pathways (AnkB & LubX) [145-151].

The accumulation of the alarmone ppGpp increases mRNA for T4SS components, secreted host regulators, and effectors [96]. Many of the substrates for the T4SS are strongly upregulated during the transmissive phase [152]. Many of the effectors upregulated during this phase of growth are involved in inhibition of phagosome maturation, altering trafficking and proteins involved in egress from the amoeba [152]. An important effector molecule for intracellular survival and replication of *L. pneumophila* is the eukaryotic-like protein AnkB which is temporally and differentially regulated at the PE phase by RelA [56]. AnkB is injected into the host cytoplasm by the Dot/Icm system immediately upon bacterial attachment to the plasma membrane, and anchors into the LCV [54, 150]. Where it plays a role in creating nutrients required for *L. pneumophila* survival and replication.

The Dot/Icm system is located at the poles of the bacteria, which is an important for the bacteria since studies have shown that non-polar localization results in the failure of lysosomal evasion by *L. pneumophila* [153]. While this system is responsible for

translocating many substrates, only 2-4 Dot/Icm apparatus complexes have been identified at the poles of the LCV [153]. Not all effectors are translocated at the same concentration or point during infection, which could be a result of the number of apparatuses located on the LCV [154, 155].

Functional Redundancy of Effector Proteins

Deletion of very few effector proteins result in intracellular growth defects in macrophages of *L.pneumophila*, which is likely the result of functional redundancy among the effector repertoire [156-158]. One explanation for functional redundancy is that over time, *L. pneumophila* has acquired a toolbox of effector proteins as a result of inter-kingdom horizontal gene transfer. Different effectors are likely specific for infection of specific hosts, which explains why deletion of effectors does not result in a growth defect in macrophages. Eliminating up to 31% of the effectors has been shown to barely cause any replication defects in mouse macrophages [158].

Effector redundancy occurs via different mechanisms including pathway redundancy, cellular process redundancy, target redundancy, molecular redundancy and system redundancy [156]. Redundant effector proteins have been shown to perform the same function in host cells and interact with the same host cell targets. One well characterized example of this is the SidE family of effectors [157]. The SidE family consists of four effector proteins (SidE, SdeA, B, and C) which function to catalyze the addition of ubiquitin moieties to the host proteins Reticulon 4 and Rab33b [159, 160]. When each of these effectors are individually deleted, there is no replication defect detected, but when all four effector proteins are deleted there is a significant decrease in

intracellular replication of *L. pneumophila*, which can be restored with complementation of SdeA alone [160, 161]. An example of pathway redundancy is that of VipD and SidK, both of which target different components of the host endocytic pathway [154, 162]. Effectors that show redundancy in targets include SidM and AnkX, both of which modulate Rab1 activity but through different mechanisms [163-165]. SidF and SidP are two examples of effectors that show cellular process redundancy as both effectors modulate host lipid metabolism and phosphoinositide abundance at the LCV by targeting redundant or complementary host pathways governed by a single process [156, 166, 167]. Lastly, system redundant effectors are effectors that are responsible for modulating more than one cellular processes in host cells in order to accomplish the same task. Examples of this include LegK1, that activates NF-κB by degrading IκB, Lgt1/2/3 blocks host protein synthesis by restoring IκB, and SidF, which inhibits the host pro-apoptotic proteins [165, 167]. These effectors work to satisfy the end goal of inducing mechanisms of host cell survival [146, 166, 168].

Genome Plasticity

A hallmark of the *L. pnumophila* genome is its low GC content. The GC content of the *L. pneumophila* genome is roughly 38.3% and the GC content of protozoan genomes is also similarly low [169]. The long-term co-evolution of *L.pneumophila* with different protozoan hosts has likely affected genome structure of this bacteria primarily through inter-kingdom horizontal gene transfer (HGT) [89, 170-173]. A high degree of plasticity is observed even between strains of the same species of *L. pneumophila* [169]. When comparing the *L. pneumophila* strains Paris and Lens, there are 2,664 genes

conserved but 428 and 280 are strain-specific, respectively [169]. Different potential "hot spots" for genomic rearrangement that contribute to the plasticity of *Legionella* have been identified [169, 174]. Some *L. pneumophila* strains contain plasmids that have been inserted into the chromosome, which has also contributed to the plasticity of the genome [169].

Co-evolution of *L. pneumophila* with amoebae and other protozoan hosts has likely contributed to the plasticity of the genome. *L. pneumophila* is naturally competent and is capable of natural transformation of DNA uptake through conjugation machinery [67, 129, 175]. Amoebae have possibly played the role of a melting pot for *L. pneumophila* which has resulted in long term convergent evolution and gene modification via HGT, which probably explains both the genome plasticity, large repertoire of effector proteins, effectors containing eukaryotic-like domains and motifs, as well as effector redundancy of the organism [169, 171, 176].

Eukaryotic-Like Proteins of *L. pneumophila* and their Origin

L. pneumophila harbors a plethora of eukaryotic-like effectors that interfere with host processes by mimicking eukaryotic functions. One of the best described examples of effector proteins with conserved eukaryotic domains that are necessary for intracellular proliferation of *L. pneumophila* is AnkB [54, 150, 177]. The AnkB protein consists of Ankyrin domains (ANK), an asparagine hydroxylation motif, a eukaryotic CaaX motif ("C' cysteine, "a" aliphatic amino acid, "X" any amino acid), and an F-box domain [177-180]. Ankyrin repeat domains are 33-amino acid domains that primarily function as scaffolds to mediate protein-protein interactions [181, 182]. These domains are responsible for targeting effector proteins in the host and are one of the most versatile

domains present in *Legionella* effectors [137, 138]. The CaaX motif allows the protein to be farnesylated, which is a highly conserved posttranslational modification that confers hydrophobicity allowing the lipidated protein to be anchored in membranes [183]. Another example of an effector of *L. pneumophila* that hijacks conserved host eukaryotic systems is the SidE family of effectors. This family of effectors exploits the ubiquitination machinery and specifically ubiquitinates Reticulon4 and Rab33 [158, 159]. Reticulon4 and Rab33 do not have homologs in all species of protozoa that can be hosts for *L. pnueumophila* and only a few have Rtn4 and Rab33 homologs indicating the SideE family has host-specific functions [208]. Many of these translocated effectors are functionally and structurally similar to eukaryotic proteins and interact with various eukaryotic processes such as signaling, protein synthesis, apoptosis, posttranslational modification, vesicular trafficking, ubiquitination, and proteasomal degradation [57, 184].

The long-term coevolution of *L. pneumophila* with various protists and metazoa has influenced the genomic structure of the bacteria through inter-kingdom HGT. Translocated effectors contain many motifs and domains normally found only in eukaryotic proteins. Long term modification of the acquired host genes through cquisition of prokaryotic promoters and regulators, as well as translocation motifs is essential to evolve the proteins to become functionally active effectors in the host cell [159]. A good example is AnkB. The F-box domain of AnkB contains the ANK domain which is common for amoeba F-box proteins but not for mammalian F-box proteins indicating AnkB was likely acquired from a primitive eukaryotic host [89, 180, 185]. It is to be

expected that many of the eukaryotic-like proteins are still undergoing modifications that might allow for them to be translocated and/or to act as effector proteins [89].

Domain shuffling has also played a major role in the evolution of *Legionella* effector proteins. Two recent studies, one conducted by Burstein *et al.* and one conducted by Gomez-Valero *et al*., analyzed the genomes of multiple species of *Legionella* [137, 138]. The Burstein study identified and analyzed effector proteins in 41 *Legionella* genomes. Analysis was performed using two criteria – first was the similarity to known domains in a domain database, and the second was the conservation of effector regions across orthologous groups of effectors [137]. The group identified 99 distinct domains including 53 well characterized domains and 46 new conserved domains. Next, they analyzed protein architecture, or domain combinations, and found that the same domains were commonly present in different architectures or combinations [137]. The Gomez-Valero study expanded on the data found in the Burstein study [137]. This group sequenced 58 *Legionella* species and analyzed them with publicly available genomes (80 genomes total) [138]. This group identified a total of 137 different eukaryotic motifs/domains present in the strains studied. Both studies found that the Ankyrin repeat domain (ANK) was the most common domain [137, 138]. This domain appeared in combination with a variety of other domains and architectures across the *Legionella* genome. Over 300 *Legionella* effector proteins contain an ANK domain [137, 138]. Ankyrin repeats commonly appear in combination with other protein domains in numerous effector proteins [137, 138]. Some Ankyrin repeats are found in combinations that are species-specific effectors in *Legionella*, while others were conserved across the genus.

L. pneumophila has an extraordinary number of effectors that are in its toolbox and are the probable source of ability of *L. pneumophila* to exploit many host processes within many different amoebae hosts, thus effectively increasing its fitness as a generalist pathogen [89]. Long term coevolution of *L. pneumophila* with its protozoan hosts and inter-kingdom horizontal gene transfer has likely resulted in an accidental ability to cause disease in humans, perpetuated by changes in human lifestyle. Understanding its association with amoeba will give us a better understanding of how *L. pneumophila* is able to cause disease though the exploitation of evolutionary conserved eukaryotic processes.

Core Effectors of *L. pneumophila*

In total, Burstein *et al*. identified 5,885 putative effectors present within the *Legionella* genus [137]. *Legionella* genes that consisted of $\geq$ 80% predicted effectors were split into orthologous groups which were designated *Legionella* effector ortholog groups (LEOGs) [137]. A total of 608 LEOGs were identified and it was observed that most of the LEOGs were shared by a small subset of species. Roughly 63% of the effector repertoire (3,715 effectors in 269 LEGOs) consisted of orthologs of validated effectors from *L. longbeachae* and *L. pneumophila*. The remainder (2,170 effectors in 339 LEGOs) represent new putative effectors which may show novel functionality [137]. The study conducted by Gomez-Valero *et. al.* identified roughly 18,000 effector proteins representing more than 1,600 orthologous groups [138].

Interestingly, both studies only identified seven core effectors. The study conducted by Gomez-Valero *et al.* also identified one other effector that was not included in the search by Burstein *et al.* which brings the total number of core effectors to eight.

[137, 138]. Six of the 8 core effectors are conserved among all species of *Legionella*, sequenced. One effector (MavN) had orthologs in all the sequenced *Legionella* as well as one other bacterium encoding the DOT/ICM T4SS, *Rickettsiella gyrlli* [137]. Remarkably, only one of the core effectors (AnkH/LegA3/Lpg2300) is not only conserved across the *Legionella* genus but also contains orthologs in other organisms containing the Dot/Icm T4SS, including *Coxiella burnetii* and *Rickettsiella grylli* [137, 138]. The conservation of the AnkH effector among bacterial species encoding the Dot/Icm T4SS indicated that it is involved in modulating host cell processes that are evolutionarily conserved and required by various intracellular pathogens.

Structure of AnkH and its role in the intracellular survival and replication of *L. pneumophila*

The AnkH effector is a 467-amino acid protein that contains four eukaryotic-like ANK domains [186]. Previous studies in the Abu Kwaik lab have established that AnkH is successfully translocated into the host cell cytoplasm [187]. This effector is required for intracellular replication within multiple host cells, including human monocyte-derived macrophages (hMDM), *Acanthamoeba polyphaga*, *Hartmanella vermiformis*, and for intrapulmonary proliferation in the mouse model of infection [186, 187]. This replication defect as a result of deletion of *ankH* has been shown using both colony forming units (CFUs) and confocal microscopy. With the later, we have shown that when cells are infected with a *L. pneumophila* strain lacking the AnkH effector (Δ*ankH)*, the LCV contains fewer bacteria than the LCV of cells infected with the wild type strain [187]. The Δ*ankH* mutant is rescued by complementation and can also be trans-rescued in

eukaryotic cells transfected with the *ankH* gene [186]. This was also the case when the Δ*ankH* mutant shared communal or distinct LCVs within the same cell as the WT strain during co-infection studies [187].

When cells are infected with the Δ*ankH* strain, the LCV biogenesis and its fusion to ER-derived vesicles is indistinguishable from the LCV harboring the WT strain. When AnkH is ectopically expressed, there was no significant difference observed between the association of the tagged AnkH protein with trafficking markers including Lamp2, cathepsin D, GM130, KDEL actin, tubulin, or mitochondrial protein [186] (Figure 1-2).

In addition to ANK repeats, there have also been two asparagine hydroxylation motifs identified within AnkH [179]. The crystal structure of AnkH has revealed two Asn hydroxylation motifs, four ANK domains, as well as a cysteine-like protease domain and a CAP domain [188]. We have previously shown that two of the ANK domains are required for proper function of the protein, since deletion of either domain results in an intracellular replication defect [185]. It has also been shown that one of the asparagine hydroxylation motifs is hydroxylated (at N59) and the motifs are required for proper function of AnkH [179]. Overall, previous data has shown that AnkH is an important effector for *L. pneumophila* and defining the function of AnkH during infection would result in a better understanding of the pathogenesis of *L. pneumophila*.

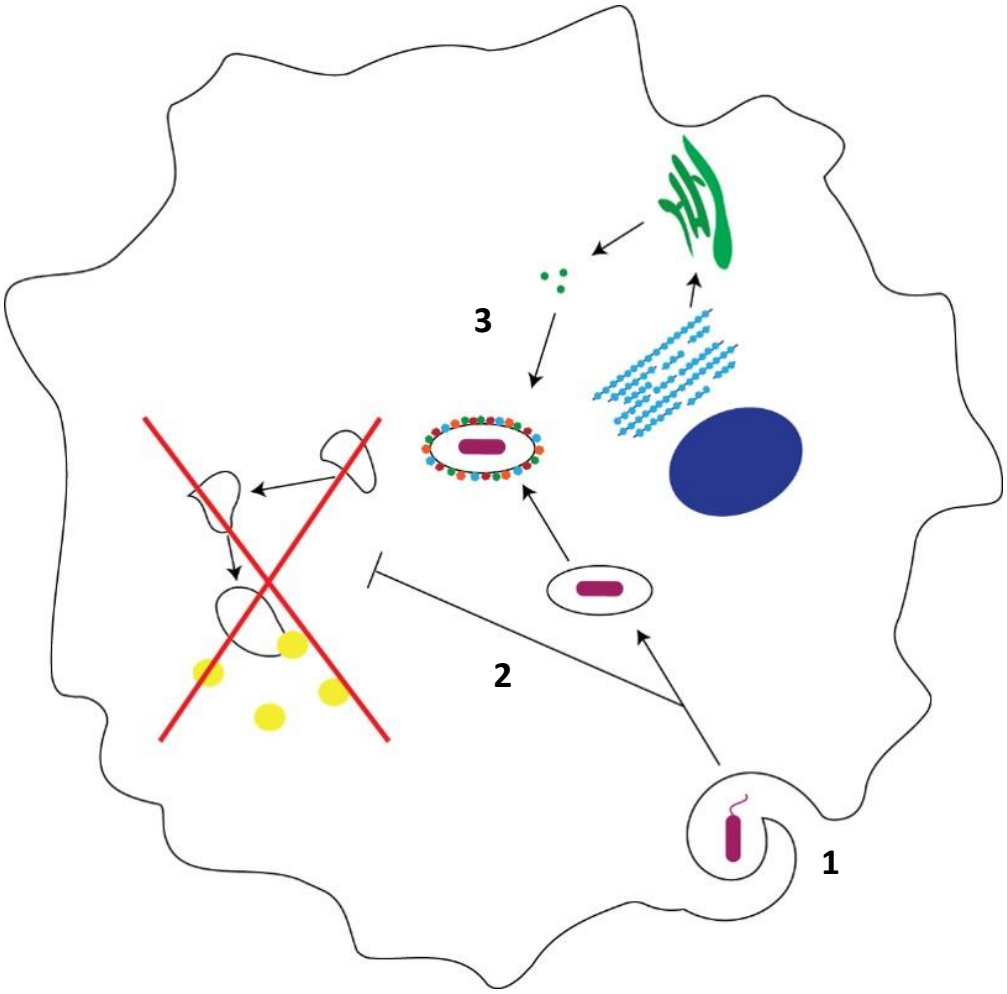

Figure 1-2: AnkH is not involved in LCV formation or evasion of host degradation pathways. Once engulfed by the host cell (1), LCVs harboring the WT or ankH mutant strain inhabit similar LCVs which evade the host endosome-lysosome degradation pathway (2) and go on to intercept ER derived vesicles that help to create a replicative niche for the bacterium (3).

SPECIFIC AIMS

L. pneumophila has spent a considerable amount of time co-evolving with primitive eukaryotic organisms resulting in a plethora of effectors with eukaryotic like protein domains. These protein domains are found dispersed among *L. pneumophila* effector proteins resulting in a toolbox that aids in the infection of specific *L. pneumophila* host organisms. Many of the ~330 effector proteins translocated by *L. pneumophila* contain eukaryotic-like protein domains demonstrating the importance of these domains in the function of many of the effectors, which contributes to intracellular survival of *L. pneumophila*. Ankyrin repeat domains were identified as the most commonly occurring eukaryotic-like protein domain among effectors in the *Legionella* genus [136, 137]. Ankyrin repeat domains are involved in protein-protein interactions and act as a scaffold for these interactions [180, 181]. *L. pneumophila* contains 11 effector proteins containing ANK domains [186]. One of the ANK domain containing effector proteins, AnkH, has been shown to be required for intracellular replication of *L. pneumophila* in amoebae, human macrophages, and for intrapulmonary proliferation in the mouse model of infection. The crystal structure revealed that AnkH consists of four ANK domains, a cysteine-like protease domain, two asparagine hydroxylation motifs, and a cap domain [188]. The *Legionella* genus codes for ~18,000 effector proteins. Of those effector proteins, AnkH is the only effector that is conserved among all sequenced

Species of *Legionella* as well as among organisms that contain the Dot/Icm T4SS [137, 138].

I hypothesize that the AnkH effector interacts with a specific highly conserved host target and modulates an evolutionarily conserved process in eukaryotic cells.

This hypothesis will be tested through three specific aims:

Specific Aim 1: Identify the host-cell interacting partners for AnkH.

Specific Aim 2: Identify downstream cellular processes altered by the interaction between AnkH and its host cell target proteins.

Specific Aim 3: Determine the role of various domains and motifs of AnkH in its function.

CHAPTER 2:

INTERACTION OF THE ANKYRIN H CORE EFFECTOR OF *LEGIONELLA* WITH THE HOST LARP7 COMPONENT OF THE 7SK SNRNP COMPLEX*

---

* Von Dwingelo, J., Chung, I., Price, C.T., Li, L., Jones, S., Cygler, M., Abu Kwaik, Y. Interaction of the Ankyrin H Core Effector of Legionella with the Host LARP7 Component of the 7SK snRNP Complex. mBio, 2019, 10(4) e01942-19; DOI: 10.1128/mBio.01942-19

Summary

The *Legionella* genus encode at least 18,000 effector proteins that are translocated through the Dot/Icm type IVB translocation system into macrophages and protist hosts to enable intracellular growth. Eight effectors, including Ankyrin H (AnkH), are common to all *Legionella* species. The AnkH effector is also present in *Coxiella* and *Rickettsiella*. To date, no pathogenic effectors have ever been described that directly interfere with host cell transcription. We identified the host nuclear protein LARP7, which is a component of the 7SK snRNP complex, to interact with AnkH in the host cell nucleus. The AnkH-LARP7 interaction partially impedes interaction of the 7SK snRNP components with LARP7, interfering with transcriptional elongation by Pol II. Consistent with that, our data show AnkH-dependent global reprogramming of transcription of macrophages infected by *L. pneumophila*. The crystal structure of AnkH shows that it contains N-terminal four ankyrin repeats, followed by a cysteine protease-like domain and an α-helical C-terminal domain. A substitution within the β-hairpin loop of the third ankyrin repeat results in diminished LARP7-AnkH interactions and phenocopies the *ankH* null mutant defect in intracellular growth. LARP7 knockdown partially suppresses intracellular proliferation of WT bacteria and increases severity of the defect of the Δ*ankH* mutant indicating a role for LARP7 in permissiveness of host cells to intracellular bacterial infection. We conclude that AnkH-LARP7 interaction impedes interaction of LARP7 with 7SK snRNP, which would block transcriptional elongation by Pol II leading to host global transcriptional reprogramming and permissiveness to *L. pneumophila*.

Importance

In order for intracellular pathogens to thrive in host cells, an environment that supports the survival and replication needs to be established. *L. pneumophila* accomplishes this through the ~330 effector proteins that are injected into host cells during infection. Effector functions range from hijacking host cells trafficking pathways to altering host cell machinery resulting in altered cell biology and innate immunity. One such pathway is the host protein synthesis pathway. Currently, 5 *L. pneumophila* effectors have been identified that alter host cell translation while only 2 effectors have been identified that indirectly affect host cell transcription. To date, no pathogenic effectors have ever been described that directly interfere with host cell transcription. Here we show direct interaction of the AnkH effector with a host cell transcription complex involved in transcriptional elongation. We identify a novel process by which AnkH interferes with host transcriptional elongation through interference with formation of a functional complex and this interference is required for pathogen proliferation.

Introduction

Legionella pneumophila is a Gram-negative intracellular pathogen that is ubiquitous in freshwater environments [189] where it primarily parasitizes a wide range of protozoan hosts, which serve as the bacterial natural hosts [40, 190, 191] and contribute to pathogenesis and ecology of the pathogen [4, 50, 65, 192]. When humans encounter contaminated water sources, aerosolized water droplets can be inhaled and reach the lung where bacteria can invade and proliferate within alveolar macrophages, causing pneumonia [193]. To date, approximately 65 species of *Legionella* have been identified with almost half of the species capable of causing disease in humans [137, 138, 194, 195]. *L. pneumophila* in particular is responsible for 90% of Legionnaires disease cases globally [8].

The life cycle of *L. pneumophila* within amoebae and alveolar macrophages, is strikingly similar [196-201]. After the bacteria are engulfed by the cell, ER-derived vesicles fuse to the phagosome to generate the *Legionella*-containing vacuole (LCV) [198, 200, 202, 203], which evades the host endosomal-lysosomal degradation pathway but communicates with early secretory vesicle trafficking pathways [156, 204]. Biogenesis of the LCV is dependent on the Dot/Icm type IV secretion system that is responsible for translocation of at least 330 effector proteins into the host cell cytoplasm [205, 206]. The injected effectors interact with specific host targets to modulate a plethora of host cell processes that remodel the macrophage and amoeba host into a proliferative niche [205, 207-209]. In most cases, the deletion of a single *L. pneumophila* effector gene does not result in a growth defect in mammalian macrophages or amoeba [158]. Although this is thought to be due to redundancy, it is more likely that many of

this arsenal of effectors are host-specific and constitute a “toolbox” from which specific tools are utilized in specific environmental eukaryotic hosts [156, 210]. Genomic analysis of 58 *Legionella* species have shown that the legionella genus has ~18,000 effectors but only 8 of these effectors (MavN, VipF, RavC, CetLp1, lpg2832, lpg3000, lpg1356/lpp1310 and AnkH/LegA3/Lpg2300) are conserved among all *Legionella* species and are designated as core effectors [137, 138]. Of the 8 core effectors, AnkH is the only effector conserved among all bacterial pathogens harboring the Dot/Icm T4SS, including *Coxiella burnetii* and *Rickettsiella grylli* [137, 138]. It is therefore likely that AnkH is involved in altering an evolutionarily conserved eukaryotic process required for the infection by many obligate and facultative intracellular pathogens.

A large number of the Dot/Icm-translocated effector proteins contain eukaryotic-like motifs and domains, which is likely the result of long-term co-evolution of *L. pneumophila* with its various protozoan hosts, leading to inter-kingdom horizontal gene transfer [89, 169-172, 192, 210]. Examples of these eukaryotic domains include F box and prenylation motifs, U box domain, leucine-rich repeats, and ankyrin repeat domains (ARDs), which are protein-protein interactions domains [54, 150, 211-213].

The ankyrin repeat (AR) is a structural fold composed of two α-helices forming a helix-turn-helix motif. It is one of the most commonly structural motifs found in eukaryotic proteins [54, 169]. AR-containing domains (ARD) usually contain multiple ARs [181, 214-218] and function predominantly as protein-protein interactions scaffolds [219, 220]. Many bacterial pathogens that inject protein effectors into host cells harbor eukaryotic-like ARD-containing protein effectors that interact with specific host targets [89, 165, 221]. Among 58 sequenced species of *Legionella*, 1134 ARD-containing

effectors have been identified in various combinations with other eukaryotic domains [137, 138, 187].

While many *L. pneumophila* effectors are dispensable for intracellular growth of the pathogen in macrophages, we have previously shown that the AnkH ARD-containing effector is one of very few effectors required for intracellular replication in macrophages, amoebae, and for intrapulmonary proliferation in the A/J mouse model [186, 187]. We have also shown that AnkH is one of the effector proteins that contains an asparagine hydroxylation motif (Lxxxxx(D/E)(ILVA)N(ILVA)), which is hydroxylated in human macrophages [179, 186].

While no *L. pneumophila* effectors have been shown to interfere directly with host transcription machinery, few *L. pneumophila* effectors have been identified that modulate host translation machinery. Five effectors (Lgt1, Lgt2, Lgt3, SidI, and SidL) act on host translation machinery primarily by interfering with the host elongation factors eELF1A and eELF1Bγ [145, 146, 222, 223]. In contrast, the RomA (or LegAS4) effectors are SET-domain containing proteins that directly modify host chromatin through histone modification but the effect on host transcription is not known [224, 225]. The LegK7 effector interferes with the host Hippo signaling pathway, which results in the degradation of TAZ and YAP1 transcriptional regulators to alter the transcriptional profile of mammalian macrophages [226].

No bacterial effector has been shown to modulate the function of 7SK small nuclear ribonucleoprotein (7SK snRNP). The La related protein 7 (LARP7) is a component of the 7SK snRNP complex which controls the pausing time of Pol II at the initiation of transcriptional elongation at almost all metazoan genes [227-229]. Binding of

LARP7 to the 7SK 3'-terminal U-rich sequence protects 7SK from nucleolytic degradation [229-233]. The canonical 7SK snRNP core complex consists of 7SK, LARP7, and γ-methylphosphate capping enzyme (MePCE) [227-229]. Formation of the 7SK snRNP core complex enables recruitment of transcription elongation factor b (P-TEFb; Cdk9-cyclin T1 heterodimer) and HEXIM1/2 dimer to the complex [228, 231, 234-237]. Binding and sequestration of P-TEFb within the 7SK snRNP complex results in inhibition of its kinase activity and continued pause in Pol II transcription elongation [230, 233, 238, 239]. P-TEFb is the critical factor that controls the release of paused Pol II into productive elongation at almost all metazoan genes. Various stimuli trigger the release of P-TEFb from the 7SK snRNP complex, leading to activation of its kinase activity and transition of Pol II into productive transcriptional elongation [240, 241]. Our data indicate that the β-hairpin loop of the third ankyrin repeat of AnkH interacts with LARP7. The AnkH-LARP7 interaction impedes interaction of LARP7 with the 7SK snRNP complex components, which would trigger transcriptional elongation by Pol II leading to host global transcriptional reprogramming.

Results

Interaction of AnkH with the LARP7 host protein

We utilized the yeast two-hybrid system to identify potential host cell interacting partners of AnkH. The full-length coding sequence of AnkH served as the bait construct and the normalized universal human library was used for the prey. After mating of the two yeast strains, a total of 1004 potentially positive clones were identified, and their growth on a selective media narrowed the number of positive clones to 37. After multiple

rounds of co-transformations of AnkH and the 37 positive clones, seven potential interacting partners of AnkH were identified (Table 2-1). Of the seven host proteins candidates, LARP7 was the only positive in all co-transformations and we pursued verification of its interaction with AnkH.

Table 2-1. Potential interacting partners identified in Y2H screen.

Proteins identified by Yeast 2 Hybrid Assay	
LA related protein 7 (LARP7)	Involved in global transcription regulation
Intersectin 2 (INST2)	Adaptor protein involved in trafficking of endocytic vesicles
Ubiquitin specific peptidase like 1 (USPL1)	SUMO specific isopeptidase involved in protein desumoylation
ANK repeat domain 18A (ANKRD18A)	Possible role in global regulation of platelet function and number
TOX4	Involved in regulating chromatin structure and cell cycle progression
Sodium channel modifier 1 (SNCM1)	Zinc finger protein and putative splicing factor
HLA-DQA1	Involved in process of presenting antigens on cell surface

The LARP7 protein is a component of the 7SK snRNP complex, which enables continued pause of Pol II elongation through sequestering and inhibiting the kinase activity of P-TEFb [242]. To confirm the AnkH-LARP7 interaction, tagged-AnkH and LARP7 were co-transfected into human embryonic kidney (HEK293T) cells and subjected to reciprocal co-immunoprecipitation (co-IP) by IP of AnkH or LARP7 (Fig. 2-1A). The data showed that LARP7 was pulled down with AnkH in the reciprocal co-IPs (Fig. 2-1A, third lane from left). To determine if AnkH-LARP7 interaction impacted recruitment of critical components essential for sequestration of P-TEFb in the 7SK snRNP complex, we determined whether the LARP7-AnkH complex interacted with the

7SK snRNP components. The AnkH co-IP was probed in immunoblots for components of the 7SK snRNP complex (CDK9, CyclinT1, MePCE, HEXIM 1/2). The data showed that none of the other complex components were immunoprecipitated with the LARP7-AnkH complex, similar to the vector control (Fig. 2-1B). However, MePCE was immunoprecipitated with the LARP7-AnkH complex 60% of the time (3 out of 5 replicates). This could be the result of expression of MePCE and the transient formation of the 7SK snRNP complex or that these MePCE positive samples were immunoprecipitated in instances where LARP7 is part of the complex and has not yet been removed from the complex via the LARP7-AnkH interaction. Importantly, in the absence of AnkH, all the 7SK snRNP components immunoprecipitated in a complex with LARP7 (Fig. 2-1C). Our data show that AnkH specifically interacts *in vivo* with the LARP7 protein and this impedes interaction of LARP7 with critical components of the 7SK snRNP complex required for the sequestration of P-TEFb in the 7SK snRNP complex.

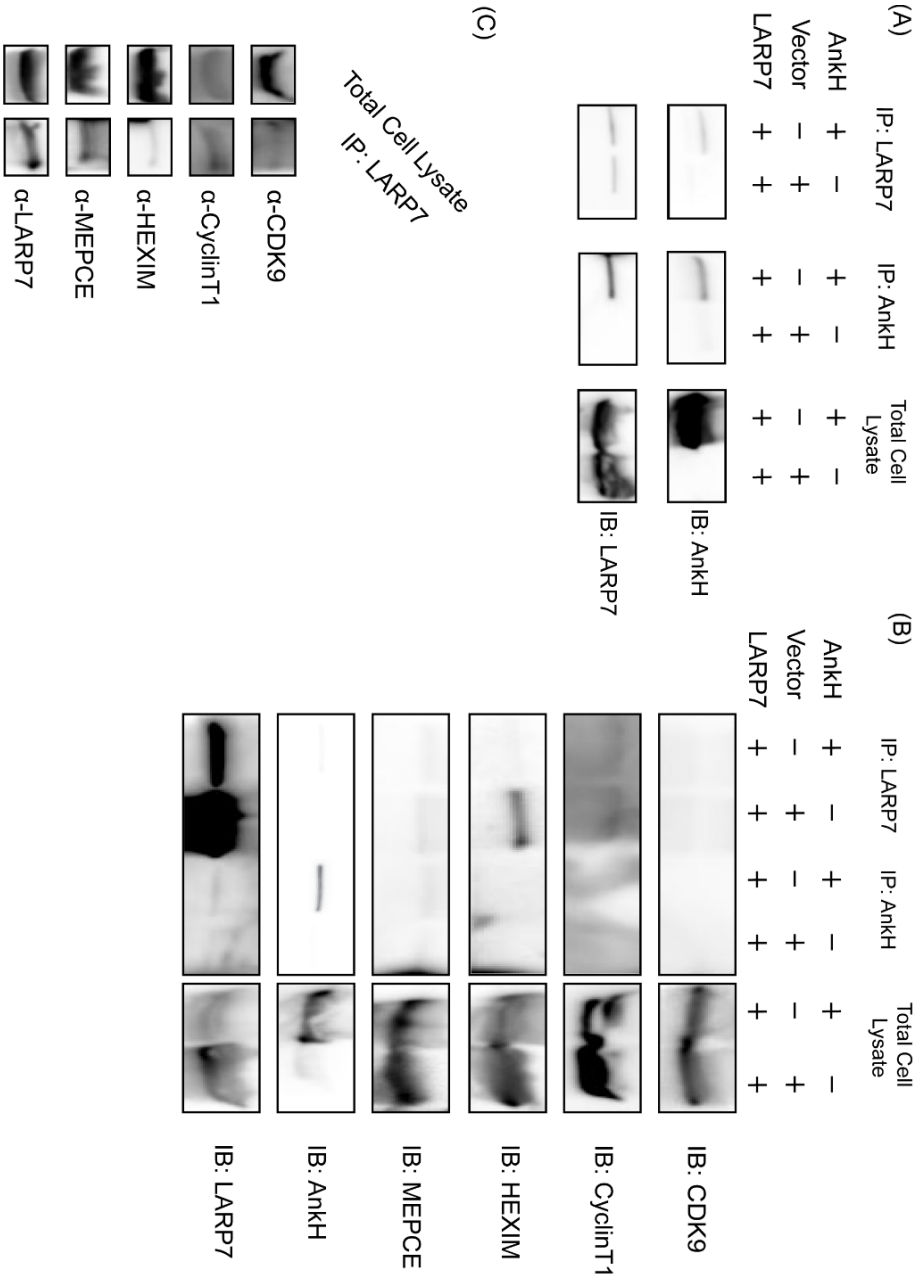

Figure 2-1. Interaction of LARP7 with the AnkH effector. (A) HEK293T cells were transiently transfected with 3xFLAG-AnkH or 3xFLAG-BAP and c-myc-LARP7, and immuneprecipitated with anti-FLAG or anti-myc antibody, and the co-IP was immunoblotted to detect the presence of AnkH and LARP7. (B) The AnkH co-IP was immunoblotted against 7SK snRNP complex components. (C) HEK293T cells were transiently transfected with c-myc-LARP7 and immunoprecipitated with anti-myc antibody and the IP was immunoblotted to detect the presence of 7SK snRNP complex

components. Lanes for total cell lysates of the immunoblot were imaged for less time due to high intensity signal. Results are representative of five independent experiments.

Localization of AnkH with LARP7 to the host cell nucleus

Consistent with its role in transcription, LARP7 is localized primarily in the nucleus [242]. Since AnkH interacts with LARP7, we determined whether the AnkH effector was targeted to the nucleus. HEK293T cells were transfected with a plasmid containing tagged AnkH and subcellular localization of AnkH was examined using confocal microscopy (Fig. 2-2A). In 85% of transfected cells, the AnkH effector was predominantly localized to the nucleus in addition to some cytosolic localization (Fig. 2-2A). In contrast, the AnkB effector control was primarily localized to the plasma membrane (92%) (Fig. 2-2A) [177].

To confirm sub-cellular localization of AnkH, nuclear and cytoplasmic fractions were analyzed by immunoblotting. In cells transfected with tagged AnkH, the majority of AnkH was present in both the nuclear and cytoplasmic fractions (Fig. 2-2B), while the AnkB effector control was mainly localized to the cytoplasmic fraction (Fig. 2-2B). Cellular fractionation was confirmed using the nuclear protein Lamin as a control (Fig. 2-2B).

To determine if AnkH and LARP7 were simultaneously localized to the nucleus, HEK293T cells were transfected with tagged AnkH and LARP7 and confocal microscopy was performed. The tagged bacterial alkaline phosphatase (BAP) was used as the control. Our data confirmed that ~70% of the cells showed simultaneous localization of AnkH and LARP7 in the nucleus (Student *t*-test $p < 0.01$) (Fig 2-2C), compared to the

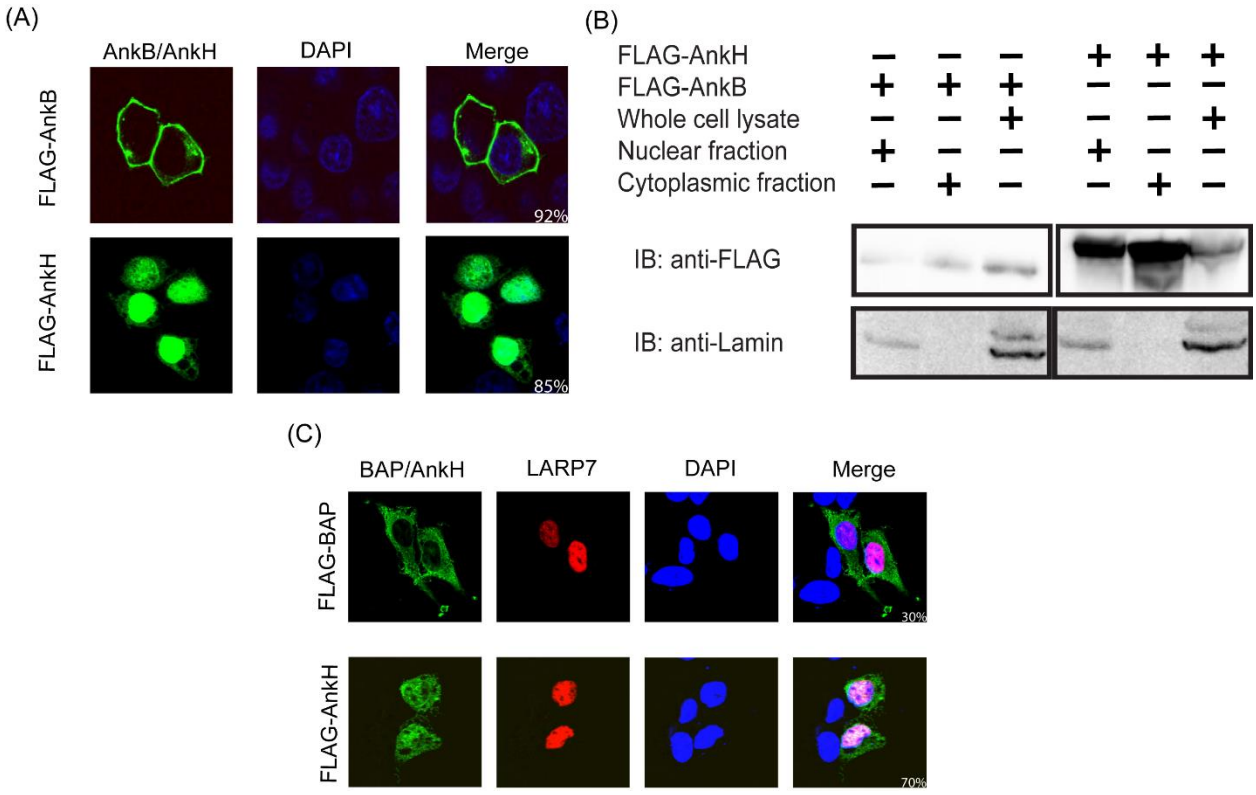

Figure 2-2. Localization of AnkH with LARP7 in the nucleus. (A) Representative confocal microscopy images of HEK293T cells transiently transfected with 3xFLAG-AnkH or 3xFLAG-AnkB control. The cells were labeled with anti-FLAG antibody (green), and the nucleus is stained with DAPI (blue). (B) Representative confocal microscopy images of HEK293T cells transiently co-transfected with 3xFLAG-AnkH and c-myc-LARP7 or 3xFLAG-BAP and c-myc-LARP7. The cells were labeled with anti-FLAG (green), anti-myc (red), and the nucleus is stained with DAPI (blue). Numbers in the merged images in (A) and (B) are quantification of % of nuclear localizations of AnkH and LARP7 proteins in HEK293T cells. For (A) and (B), 100 transfected cells were analyzed from multiple coverslips. Results are representative of three independent experiments performed in triplicate. (C) HEK293T cells transiently transfected with 3xFLAG-AnkH or 3xFLAG-AnkB control were subjected to nuclear fractionation. Cell fractions were separated by SDS-PAGE and analyzed by immunoblotting. AnkH and AnkB were detected using anti-FLAG monoclonal antibody. Fractionation was confirmed by detection of the nuclear protein Lamin.

control BAP (~30%), which is a highly expressed protein (Fig. 2-2C). Our data showed that AnkH and LARP7 are localized to the nucleus, consistent with their interaction.

Role of LARP7 in intracellular replication of *L. pneumophila* in hMDMs

We have previously shown that AnkH is required for intracellular replication of *L. pneumophila* in macrophage and amoeba [186, 187]. Depletion of either LARP7 or MePCE via RNAi triggers 7SK degradation in cells [231, 235, 243]. Since AnkH interacts with the LARP7 component of the 7SK snRNP complex, we investigated if LARP7 was also required for replication of *L. pneumophila*. We utilized a lentiviral RNAi system to knockdown expression of LARP7 in human monocyte derived macrophages (hMDMs) that were infected with the WT strain of *L. pneumophila* or the Δ*ankH* null mutant. Knockdown of LARP7 was confirmed by immunoblot (Fig. 2-3A). Interestingly, when LARP7 was knocked down and cells were infected with the Δ*ankH*, the defective phenotype was exacerbated. Surprisingly, the knockdown of LARP7 resulted in a partial but significant decrease in intracellular replication of the WT strain (Student *t*-test $p < 0.05$), which was not observed in non-treated or control RNAi-treated cells (Fig. 2-3B). These data support our findings for the role of AnkH-LARP7 interaction in intracellular replication of *L. pneumophila* in hMDMs and indicates that LARP7 is involved in transcription of genes involved in permissiveness to *L. pneumophila*.

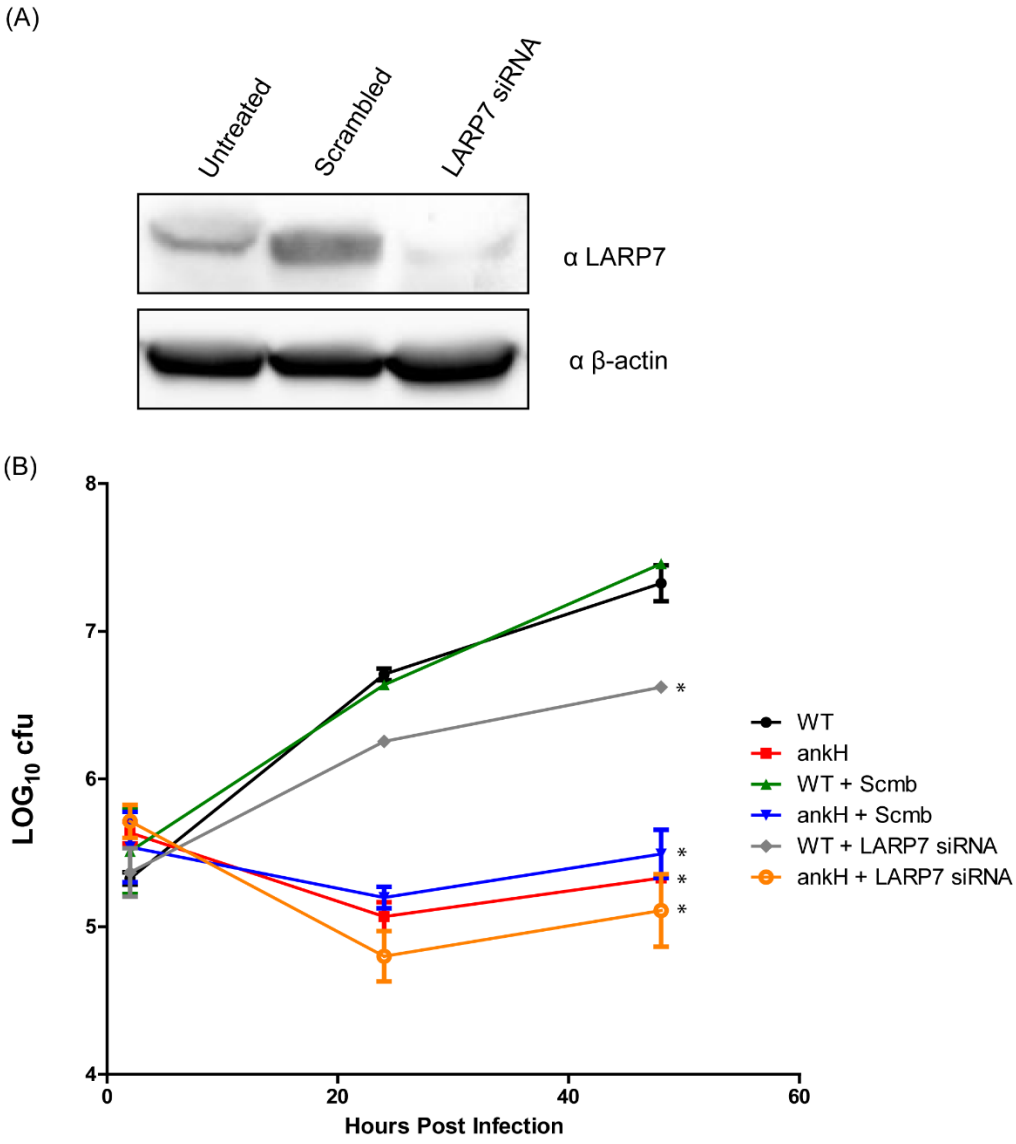

Figure 2-3. Requirement of LARP7 for intracellular replication of *L. pneumophila*. (A) Cells were treated with LARP7 RNAi for 24 hours then infected. Knockdown of LARP7 was determined by immunoblotting with anti-LARP7 polyclonal antibody. (B) Intracellular growth kinetics of *L. pneumophila* in hMDMs treated with LARP7-specific or scrambled RNAi. The results are representative of three independent experiments performed in triplicate. Statistical analysis was performed comparing all conditions to WT untreated using Student's T test where *, $P < 0.05$.

Alteration of host global transcription by AnkH

Our data showed that the LARP7-AnkH complex impedes interaction of LARP7 with critical components of the 7SK snRNP complex required for the sequestration of P-TEFb in the 7SK snRNP complex, which indicates an active P-TEFb kinase and release of Pol II from pause sites and transitions into productive transcriptional elongation [231, 242]. We utilized RNAseq to examine modulation of global gene expression in hMDMs infected with either the WT strain or the Δ*ankH* null mutant. The data showed deletion of AnkH had a dramatic effect on global transcription of *L. pneumophila*-infected hMDMs, with a total of 405 genes that were differentially regulated in cells infected with the WT strain compared to the Δ*ankH* mutant, the top 10 of each based on log-fold change are listed in Table 2-2 (Full list in Table 2-3 & 2-4). MetaCore was used to determine which pathways were differentially regulated based on p values. Certain cellular pathways were downregulated in AnkH-dependent manner, including apoptosis, autophagy and certain signaling pathways including STK3/4 pathway and JNK pathway, indicating negative regulation of these pathways by AnkH during infection (Table 2-5). When cells were infected with the Δ*ankH* null mutant, transcription and immune response pathways were downregulated compared to cells infected with the WT strain, indicating their upregulation by AnkH (Table 2-5). An array of 10 cytokines were tested based on RNAseq data to determine which were altered by the presence of AnkH (Figure 2-4A-H). Human monocyte derived macrophages were infected with WT or Δ*ankH L. pneumophila* for 6 hours. Supernatants were collected and centrifuged to remove debris then cytokine levels were tested using a 10-panel cytokine multiplex. Of the 10 cytokines tested, only eight produced levels that were in a detectable range, and only IL-1α, showed

a significant difference in cells infected with WT versus Δ*ankH* (Figure 2-4G). These data show that AnkH triggers a dramatic reprogramming of cellular transcription and that is most likely mediated by interaction with several host substrates, one of which is LARP7.

Table 2-2: Top 10 up regulated and top 10 down regulated genes in cells infected with *ankH* mutant compared to cells infected with WT.

Gene Symbol | Description	Log_2FC	Gene Symbol | Description	Log_2FC
HSPA1B| Heat shock 70kDa protein 1B	+2.9979	TMC8|Transmembrane channel-like 8	-2.75204
EGR1| Early growth response 1	+2.25815	HMHA1| Histocompatibility (minor) HA-1	-1.99385
DNAJB1| DnaJ (Hsp40) homolog, subunit B, member 1	+1.51215	DAPK3|Death associated protein kinase 3	-1.91505
DUSP1| Dual specificity phosphatase 1	+1.24951	PDLIM2|PDZ and LIM domain 2	-1.79007
FOS| FBJ murine osteosarcoma viral oncogene homolog	+1.15258	SDF2L1|Stromal cell-derived factor 2-like 1	-1.65596
TDO2| Tryptophan metabolism	+1.01074	TOR2A|Torsin family 2, member A	-1.62058
MS4A4E| Membrane-spanning 4-domains, subfamily A, member 4E	+2.36464	LMF2|Lipase maturation factor 2	-1.59493
PKIB| Protein kinase inhibitor beta	+1.33129	NOTCH3| Notch 3	-1.57732
PEG3| Paternally espressed 3	+2.64637	IL27| Interleukin 27	-1.55472
GRIK2| Glutamate receptor, ionotropic, kainate 2	+1.78851	CPSF1|Cleavage and polyadenylation specific factor 1	-1.89935

Table 2-3: Complete list of genes upregulated in hMDMs infected with Δ*ankH* null mutant compared to WT strain of *L. pneumophila*.

ENSEMBL GENE	ENTREZ ID	GENE SYMBOL|DESCRIPTION	log2FC (ankh/wild_type)	p_value
ENSG00000120129	1843	DUSP1|dual specificity phosphatase 1	1.24951	5.00E-05
ENSG00000120738	1958	EGR1|early growth response 1	2.25815	5.00E-05
ENSG00000132002	3337	DNAJB1|DnaJ (Hsp40) homolog, subfamily B, member 1	1.51215	5.00E-05
ENSG00000135549	5570	PKIB|protein kinase (cAMP-dependent, catalytic) inhibitor beta	1.33129	0.00105
ENSG00000151790	6999	TDO2|tryptophan 2,3-dioxygenase	1.01074	0.0006
ENSG00000152380	167555	FAM151B|family with sequence similarity 151, member B	1.08992	0.00015
ENSG00000164418	2898	GRIK2|glutamate receptor, ionotropic, kainate 2	1.78851	0.00245
ENSG00000165694	90167	FRMD7|FERM domain containing 7	1.3049	5.00E-05
ENSG00000170345	2353	FOS|FBJ murine osteosarcoma viral oncogene homolog	1.15258	5.00E-05
ENSG00000185842	127602	DNAH14|dynein, axonemal, heavy chain 14	1.00125	0.0014
ENSG00000198300	5178	PEG3|paternally expressed 3	2.64637	0.0013
ENSG00000204388	3304	HSPA1B|heat shock 70kDa protein 1B	2.9979	5.00E-05
ENSG00000214787	643680	MS4A4E|membrane-spanning 4-domains, subfamily A, member 4E	2.36464	0.001
ENSG00000225465		RFPL1S|	2.61193	0.00195
ENSG00000226047			1.04367	0.0009
ENSG00000227028	100128590	SLC8A1-AS1|SLC8A1 antisense RNA 1	2.66729	5.00E-05
ENSG00000229956	100852410	ZRANB2-AS2|ZRANB2 antisense RNA 2 (head to head)	1.22918	0.00125
ENSG00000234506	101927015	LINC01506|long intergenic non-protein coding RNA 1506	1.10281	5.00E-05
ENSG00000245573	497258	BDNF-AS|BDNF antisense RNA	1.01595	0.00595
ENSG00000262097	101927311	|uncharacterized LOC101927311	1.32999	5.00E-05
ENSG00000279348			1.02381	5.00E-05

Table 2-4: Complete list of genes downregulated in hMDMs infected with Δ*ankH* null mutant compared to WT strain of *L. pneumophila*.

ENSEMBL GENE	ENTREZ ID	GENE SYMBOL|DESCRIPTION	log2FC (ankh/wild_type)	p_value
ENSG00000005075	5439	POLR2J|polymerase (RNA) II (DNA directed) polypeptide J, 13.3kDa	-1.00719	0.0004
ENSG00000008382	84954	MPND|MPN domain containing	-1.13737	0.00305
ENSG00000008441	4784	NFIX|nuclear factor I/X (CCAAT-binding transcription factor)	-1.56187	0.00215
ENSG00000010295	25900	IFFO1|intermediate filament family orphan 1	-1.02159	5.00E-05
ENSG00000011028	9902	MRC2|mannose receptor, C type 2	-1.58459	0.0001
ENSG00000014164	23144	ZC3H3|zinc finger CCCH-type containing 3	-1.16549	5.00E-05
ENSG00000015285	7454	WAS|Wiskott-Aldrich syndrome	-1.02431	5.00E-05
ENSG00000021762	114879	OSBPL5|oxysterol binding protein-like 5	-1.0211	0.00325
ENSG00000023191	6050	RNH1|ribonuclease/angiogenin inhibitor 1	-1.21807	5.00E-05
ENSG00000025770	29781	NCAPH2|non-SMC condensin II complex, subunit H2	-1.01792	0.00085
ENSG00000029534	286	ANK1|ankyrin 1, erythrocytic	-1.56839	0.00025
ENSG00000030582	2896	GRN|granulin	-1.10623	5.00E-05
ENSG00000037042	27175	TUBG2|tubulin, gamma 2	-1.06627	0.001
ENSG00000050820	9564	BCAR1|breast cancer anti-estrogen resistance 1	-1.06863	0.00475
ENSG00000051128	9454	HOMER3|homer scaffolding protein 3	-1.92266	0.0026
ENSG00000051523	1535	CYBA|cytochrome b-245, alpha polypeptide	-1.24454	5.00E-05
ENSG00000061938	10188	TNK2|tyrosine kinase, non-receptor, 2	-1.22665	0.00035
ENSG00000063245	29924	EPN1|epsin 1	-1.40878	5.00E-05
ENSG00000063854	3029	HAGH|hydroxyacylglutathione hydrolase	-1.02578	0.003
ENSG00000064490	8625	RFXANK|regulatory factor X-associated ankyrin-containing protein	-1.10454	0.00115
ENSG00000064687	10347	ABCA7|ATP-binding cassette, sub-family A (ABC1), member 7	-1.16013	0.00305
ENSG00000065268	57418	WDR18|WD repeat domain 18	-1.24648	0.0052
ENSG00000065	9423	NTN1|netrin 1	-1.27328	0.0001

20				5
ENSG00000068001	8692	HYAL2|hyaluronoglucosaminidase 2	-1.02605	0.0037
ENSG00000070404	10272	FSTL3|follistatin-like 3 (secreted glycoprotein)	-1.47082	0.00035
ENSG00000070413	9993	DGCR2|DiGeorge syndrome critical region gene 2	-1.08748	0.00105
ENSG00000071859	9130	FAM50A|family with sequence similarity 50, member A	-1.03637	0.0007
ENSG00000071889	60343	FAM3A|family with sequence similarity 3, member A	-1.10308	5.00E-05
ENSG00000071894	29894	CPSF1|cleavage and polyadenylation specific factor 1, 160kDa	-1.89935	5.00E-05
ENSG00000072110	87	ACTN1|actinin, alpha 1	-1.02917	5.00E-05
ENSG00000072786	6793	STK10|serine/threonine kinase 10	-1.09231	5.00E-05
ENSG00000074181	4854	NOTCH3|notch 3	-1.57732	5.00E-05
ENSG00000074964	55160	ARHGEF10L|Rho guanine nucleotide exchange factor (GEF) 10-like	-1.11483	0.00095
ENSG00000075618	6624	FSCN1|fascin actin-bundling protein 1	-1.54859	5.00E-05
ENSG00000076924	56949	XAB2|XPA binding protein 2	-1.45048	5.00E-05
ENSG00000077454	4034	LRCH4|leucine-rich repeats and calponin homology (CH) domain containing 4	-1.18425	5.00E-05
ENSG00000078269	8871	SYNJ2|synaptojanin 2	-1.03271	0.00515
ENSG00000078808	51150	SDF4|stromal cell derived factor 4	-1.34504	5.00E-05
ENSG00000079432	23152	CIC|capicua transcriptional repressor	-1.26161	0.0007
ENSG00000080573	50509	COL5A3|collagen, type V, alpha 3	-1.26729	5.00E-05
ENSG00000083838	55663	ZNF446|zinc finger protein 446	-1.70362	0.00435
ENSG00000085117	3732	CD82|CD82 molecule	-1.09771	5.00E-05
ENSG00000088256	2767	GNA11|guanine nucleotide binding protein (G protein), alpha 11 (Gq class)	-1.47458	5.00E-05
ENSG00000090013	645	BLVRB|biliverdin reductase B	-1.05725	5.00E-05
ENSG00000095397	25861	DFNB31|deafness, autosomal recessive 31	-1.94109	0.00035
ENSG00000099817	5434	POLR2E|polymerase (RNA) II (DNA directed) polypeptide E, 25kDa	-1.17485	0.00135

ENSG00000099821	5442	POLRMT|polymerase (RNA) mitochondrial (DNA directed)	-1.29291	0.00335
ENSG00000099995	10291	SF3A1|splicing factor 3a, subunit 1, 120kDa	-1.15539	0.0006
ENSG00000100056	8220	DGCR14|DiGeorge syndrome critical region gene 14	-1.05868	0.00045
ENSG00000100075	6576	SLC25A1|solute carrier family 25 (mitochondrial carrier; citrate transporter), member 1	-1.07913	5.00E-05
ENSG00000100097	3956	LGALS1|lectin, galactoside-binding, soluble, 1	-1.05694	5.00E-05
ENSG00000100147	79879	CCDC134|coiled-coil domain containing 134	-1.1867	0.00015
ENSG00000100241	6305	SBF1|SET binding factor 1	-1.6413	0.00015
ENSG00000100258	91289	LMF2|lipase maturation factor 2	-1.59493	5.00E-05
ENSG00000100292	3162	HMOX1|heme oxygenase 1	-1.08059	5.00E-05
ENSG00000100299	410	ARSA|arylsulfatase A	-1.05355	0.0006
ENSG00000100300	706	TSPO|translocator protein (18kDa)	-1.12915	5.00E-05
ENSG00000100319	55954	ZMAT5|zinc finger, matrin-type 5	-1.14104	5.00E-05
ENSG00000100417	5372	PMM1|phosphomannomutase 1	-1.44872	5.00E-05
ENSG00000100425	23774	BRD1|bromodomain containing 1	-1.0918	0.0017
ENSG00000100429	83933	HDAC10|histone deacetylase 10	-1.30876	5.00E-05
ENSG00000100949	5875	RABGGTA|Rab geranylgeranyltransferase, alpha subunit	-1.06608	5.00E-05
ENSG00000100985	4318	MMP9|matrix metallopeptidase 9	-1.06531	5.00E-05
ENSG00000101194	63910	SLC17A9|solute carrier family 17 (vesicular nucleotide transporter), member 9	-1.14327	5.00E-05
ENSG00000101335	10398	MYL9|myosin, light chain 9, regulatory	-1.0549	0.00225
ENSG00000101443	10406	WFDC2|WAP four-disulfide core domain 2	-1.6624	0.0038
ENSG00000101986	215	ABCD1|ATP-binding cassette, sub-family D (ALD), member 1	-1.45046	5.00E-05
ENSG00000101997	28952	CCDC22|coiled-coil domain containing 22	-1.08562	5.00E-05
ENSG00000102265	7076	TIMP1|TIMP metallopeptidase inhibitor 1	-1.0606	5.00E-05
ENSG00000102879	11151	CORO1A|coronin, actin binding protein, 1A	-1.06787	5.00E-05

ENSG00000103145	54985	HCFC1R1|host cell factor C1 regulator 1 (XPO1 dependent)	-1.00296	0.00025
ENSG00000103257	8140	SLC7A5|solute carrier family 7 (amino acid transporter light chain, L system), member 5	-1.07431	5.00E-05
ENSG00000103335	9780	PIEZO1|piezo-type mechanosensitive ion channel component 1	-1.04014	0.003
ENSG00000103653	1445	CSK|c-src tyrosine kinase	-1.20218	5.00E-05
ENSG00000104368	5327	PLAT|plasminogen activator, tissue	-1.51427	5.00E-05
ENSG00000104907	55621	TRMT1|tRNA methyltransferase 1	-1.04471	0.00115
ENSG00000104973	81857	MED25|mediator complex subunit 25	-1.51656	0.0024
ENSG00000104976	6618	SNAPC2|small nuclear RNA activating complex, polypeptide 2, 45kDa	-1.68231	0.00015
ENSG00000105204	9149	DYRK1B|dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1B	-1.28731	0.0018
ENSG00000105223	23646	PLD3|phospholipase D family, member 3	-1.04403	5.00E-05
ENSG00000105229	51588	PIAS4|protein inhibitor of activated STAT, 4	-1.88599	0.00125
ENSG00000105374	4818	NKG7|natural killer cell granule protein 7	-1.11691	5.00E-05
ENSG00000105669	11316	COPE|coatomer protein complex, subunit epsilon	-1.00067	0.00035
ENSG00000105701	23770	FKBP8|FK506 binding protein 8, 38kDa	-1.50269	5.00E-05
ENSG00000105717	80714	PBX4|pre-B-cell leukemia homeobox 4	-1.63027	0.0019
ENSG00000105723	2931	GSK3A|glycogen synthase kinase 3 alpha	-1.21067	5.00E-05
ENSG00000105732	64763	ZNF574|zinc finger protein 574	-1.53883	0.00015
ENSG00000106009	221927	BRAT1|BRCA1-associated ATM activator 1	-1.35966	0.00445
ENSG00000106348	3614	IMPDH1|IMP (inosine 5'-monophosphate) dehydrogenase 1	-1.2246	5.00E-05
ENSG00000106683	3984	LIMK1|LIM domain kinase 1	-1.28683	5.00E-05
ENSG00000107816	84445	LZTS2|leucine zipper, putative tumor suppressor 2	-1.12868	0.00105
ENSG00000108639	9144	SYNGR2|synaptogyrin 2	-1.0749	5.00E-05
ENSG00000108840	10014	HDAC5|histone deacetylase 5	-1.07427	0.00025
ENSG00000109736	10227	MFSD10|major facilitator superfamily domain containing 10	-1.23104	5.00E-05

ENSG00000110025	29907	SNX15|sorting nexin 15	-5.02403	0.0016
ENSG00000110046	23130	ATG2A|autophagy related 2A	-1.21033	0.0002
ENSG00000110446	51296	SLC15A3|solute carrier family 15 (oligopeptide transporter), member 3	-1.04048	5.00E-05
ENSG00000110717	4728	NDUFS8|NADH dehydrogenase (ubiquinone) Fe-S protein 8, 23kDa (NADH-coenzyme Q reductase)	-1.04348	5.00E-05
ENSG00000110944	51561	IL23A|interleukin 23, alpha subunit p19	-1.04283	5.00E-05
ENSG00000111321	4055	LTBR|lymphotoxin beta receptor (TNFR superfamily, member 3)	-1.05988	5.00E-05
ENSG00000111678	113246	C12orf57|chromosome 12 open reading frame 57	-1.18795	0.0036
ENSG00000113494	5618	PRLR|prolactin receptor	-1.28143	0.00215
ENSG00000113657	1809	DPYSL3|dihydropyrimidinase-like 3	-1.12821	5.00E-05
ENSG00000114554	5361	PLXNA1|plexin A1	-1.40661	0.00265
ENSG00000114626	80325	ABTB1|ankyrin repeat and BTB (POZ) domain containing 1	-1.08265	0.0018
ENSG00000115085	7535	ZAP70|zeta-chain (TCR) associated protein kinase 70kDa	-2.2362	0.00025
ENSG00000115286	374291	NDUFS7|NADH dehydrogenase (ubiquinone) Fe-S protein 7, 20kDa (NADH-coenzyme Q reductase)	-1.2702	0.00285
ENSG00000115718	5624	PROC|protein C (inactivator of coagulation factors Va and VIIIa)	-1.5442	0.0027
ENSG00000116691	60672	MIIP|migration and invasion inhibitory protein	-1.06346	5.00E-05
ENSG00000116809	7709	ZBTB17|zinc finger and BTB domain containing 17	-1.2056	5.00E-05
ENSG00000117984	1509	CTSD|cathepsin D	-1.18663	5.00E-05
ENSG00000120899	2185	PTK2B|protein tyrosine kinase 2 beta	-1.15075	5.00E-05
ENSG00000120913	64236	PDLIM2|PDZ and LIM domain 2 (mystique)	-1.7752	5.00E-05
ENSG00000120949	943	TNFRSF8|tumor necrosis factor receptor superfamily, member 8	-1.03051	5.00E-05
ENSG00000121057	8165	AKAP1|A kinase (PRKA) anchor protein 1	-1.03472	0.00055
ENSG00000123143	5585	PKN1|protein kinase N1	-1.00546	5.00E-05
ENSG00000123453	1757	SARDH|sarcosine dehydrogenase	-1.14408	0.0002
ENSG00000124216	6615	SNAI1|snail family zinc finger 1	-1.78349	0.00195

ENSG00000125089	54436	SH3TC1|SH3 domain and tetratricopeptide repeats 1	-1.0519	0.00015
ENSG00000125148	4502	MT2A|metallothionein 2A	-1.06199	5.00E-05
ENSG00000125503	54776	PPP1R12C|protein phosphatase 1, regulatory subunit 12C	-1.29062	5.00E-05
ENSG00000125534	79144	PPDPF|pancreatic progenitor cell differentiation and proliferation factor	-1.42636	5.00E-05
ENSG00000125656	8192	CLPP|caseinolytic mitochondrial matrix peptidase proteolytic subunit	-1.14732	5.00E-05
ENSG00000125726	970	CD70|CD70 molecule	-1.07511	0.00035
ENSG00000125817	1059	CENPB|centromere protein B, 80kDa	-1.06327	5.00E-05
ENSG00000125912	56926	NCLN|nicalin	-1.60475	5.00E-05
ENSG00000126062	11070	TMEM115|transmembrane protein 115	-1.19645	5.00E-05
ENSG00000126254	79171	RBM42|RNA binding motif protein 42	-1.67206	5.00E-05
ENSG00000126353	1236	CCR7|chemokine (C-C motif) receptor 7	-1.03117	5.00E-05
ENSG00000126461	58506	SCAF1|SR-related CTD-associated factor 1	-1.67049	5.00E-05
ENSG00000126561	6776	STAT5A|signal transducer and activator of transcription 5A	-1.08881	5.00E-05
ENSG00000126903	8273	SLC10A3|solute carrier family 10, member 3	-1.01445	5.00E-05
ENSG00000126934	5605	MAP2K2|mitogen-activated protein kinase kinase 2	-1.03897	5.00E-05
ENSG00000127663	23030	KDM4B|lysine (K)-specific demethylase 4B	-1.5128	5.00E-05
ENSG00000127666	148022	TICAM1|toll-like receptor adaptor molecule 1	-1.0486	5.00E-05
ENSG00000128228	23753	SDF2L1|stromal cell-derived factor 2-like 1	-1.65596	5.00E-05
ENSG00000128271	135	ADORA2A|adenosine A2a receptor	-1.51909	0.00415
ENSG00000128342	3976	LIF|leukemia inhibitory factor	-1.17852	5.00E-05
ENSG00000129911	83855	KLF16|Kruppel-like factor 16	-1.04128	5.00E-05
ENSG00000129925	58986	TMEM8A|transmembrane protein 8A	-1.35995	5.00E-05
ENSG00000130165	84337	ELOF1|ELF1 homolog, elongation factor 1	-1.04415	5.00E-05
ENSG00000130203	348	APOE|apolipoprotein E	-1.35627	5.00E-05
ENSG00000130222	10912	GADD45G|growth arrest and DNA-damage-inducible, gamma	-1.11982	0.0008

ENSG00000130255	25873	RPL36|ribosomal protein L36	-1.11055	0.0001
ENSG00000130313	25796	PGLS|6-phosphogluconolactonase	-1.22736	5.00E-05
ENSG00000130479	55201	MAP1S|microtubule-associated protein 1S	-1.1206	5.00E-05
ENSG00000130706	11047	ADRM1|adhesion regulating molecule 1	-1.1761	5.00E-05
ENSG00000130726	10155	TRIM28|tripartite motif containing 28	-1.0238	5.00E-05
ENSG00000131165	5119	CHMP1A|charged multivesicular body protein 1A	-1.09835	5.00E-05
ENSG00000131459	9945	GFPT2|glutamine-fructose-6-phosphate transaminase 2	-1.10419	5.00E-05
ENSG00000131653	84231	TRAF7|TNF receptor-associated factor 7, E3 ubiquitin protein ligase	-1.41604	0.00015
ENSG00000131669	4814	NINJ1|ninjurin 1	-1.35514	5.00E-05
ENSG00000131759	5914	RARA|retinoic acid receptor, alpha	-1.14887	5.00E-05
ENSG00000132017	90379	DCAF15|DDB1 and CUL4 associated factor 15	-1.20722	0.00035
ENSG00000132382	10514	MYBBP1A|MYB binding protein (P160) 1a	-1.00679	0.0008
ENSG00000133027	10400	PEMT|phosphatidylethanolamine N-methyltransferase	-1.50047	5.00E-05
ENSG00000133069	9911	TMCC2|transmembrane and coiled-coil domain family 2	-1.31481	5.00E-05
ENSG00000133275	1455	CSNK1G2|casein kinase 1, gamma 2	-1.29917	0.00475
ENSG00000135094	10993	SDS|serine dehydratase	-1.05989	5.00E-05
ENSG00000135723	29109	FHOD1|formin homology 2 domain containing 1	-1.22406	0.00095
ENSG00000136286	64005	MYO1G|myosin IG	-1.06235	5.00E-05
ENSG00000136717	274	BIN1|bridging integrator 1	-1.00875	5.00E-05
ENSG00000136877	2356	FPGS|folylpolyglutamate synthase	-1.20062	5.00E-05
ENSG00000137166	116113	FOXP4|forkhead box P4	-1.45906	0.0004
ENSG00000137221	93643	TJAP1|tight junction associated protein 1 (peripheral)	-1.18844	0.00065
ENSG00000137266	63027	SLC22A23|solute carrier family 22, member 23	-1.00819	0.00505
ENSG00000137818	6176	RPLP1|ribosomal protein, large, P1	-1.18368	5.00E-05
ENSG00000138080	11117	EMILIN1|elastin microfibril interfacer 1	-1.52419	5.00E-05
ENSG00000139594	94	ACVRL1|activin A receptor type II-	-1.09297	0.0025

67		like 1		
ENSG00000140548	374655	ZNF710|zinc finger protein 710	-1.03225	0.0063
ENSG00000140854	10300	KATNB1|katanin p80 (WD repeat containing) subunit B 1	-1.00905	0.0025
ENSG00000140939	8996	NOL3|nucleolar protein 3 (apoptosis repressor with CARD domain)	-1.33415	0.0001
ENSG00000141526	9123	SLC16A3|solute carrier family 16 (monocarboxylate transporter), member 3	-1.11006	0.00055
ENSG00000141985	6455	SH3GL1|SH3-domain GRB2-like 1	-1.01881	5.00E-05
ENSG00000141994	56931	DUS3L|dihydrouridine synthase 3-like	-1.48655	0.00085
ENSG00000142186	57410	SCYL1|SCY1-like, kinase-like 1	-1.14476	5.00E-05
ENSG00000142546	51070	NOSIP|nitric oxide synthase interacting protein	-1.00975	5.00E-05
ENSG00000143373	57592	ZNF687|zinc finger protein 687	-1.07338	0.0059
ENSG00000143774	2987	GUK1|guanylate kinase 1	-1.12173	5.00E-05
ENSG00000144476	57007	ACKR3|atypical chemokine receptor 3	-1.37842	0.0023
ENSG00000144579	58190	CTDSP1|CTD (carboxy-terminal domain, RNA polymerase II, polypeptide A) small phosphatase 1	-1.14305	5.00E-05
ENSG00000145901	10318	TNIP1|TNFAIP3 interacting protein 1	-1.22933	5.00E-05
ENSG00000145936	3779	KCNMB1|potassium channel subfamily M regulatory beta subunit 1	-1.08752	5.00E-05
ENSG00000147443	9046	DOK2|docking protein 2, 56kDa	-1.19973	5.00E-05
ENSG00000148343	84895	FAM73B|family with sequence similarity 73, member B	-1.53125	0.0006
ENSG00000149781	83706	FERMT3|fermitin family member 3	-1.03321	5.00E-05
ENSG00000149782	5331	PLCB3|phospholipase C, beta 3 (phosphatidylinositol-specific)	-1.09659	5.00E-05
ENSG00000149925	226	ALDOA|aldolase A, fructose-bisphosphate	-1.32281	5.00E-05
ENSG00000150672	1740	DLG2|discs, large homolog 2 (Drosophila)	-1.14705	0.0054
ENSG00000151651	101	ADAM8|ADAM metallopeptidase domain 8	-1.08517	5.00E-05
ENSG00000153443	124402	UBALD1|UBA-like domain containing 1	-2.02675	0.00385
ENSG00000154040	123872	DNAAF1|dynein, axonemal,	-1.18584	0.0001

99		assembly factor 1		
ENSG00000156966	93010	B3GNT7|UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 7	-1.07511	0.00015
ENSG00000157353	197258	FUK|fucokinase	-1.20746	5.00E-05
ENSG00000158517	653361	NCF1|neutrophil cytosolic factor 1	-1.18042	5.00E-05
ENSG00000158941	57805	CCAR2|cell cycle and apoptosis regulator 2	-1.03328	0.0002
ENSG00000159069	54461	FBXW5|F-box and WD repeat domain containing 5	-1.37742	0.0044
ENSG00000159166	3898	LAD1|ladinin 1	-1.50991	0.00365
ENSG00000159189	714	C1QC|complement component 1, q subcomponent, C chain	-1.17726	5.00E-05
ENSG00000159314	201176	ARHGAP27|Rho GTPase activating protein 27	-1.75118	0.0017
ENSG00000159363	23400	ATP13A2|ATPase type 13A2	-1.22485	0.00035
ENSG00000159496	266747	RGL4|ral guanine nucleotide dissociation stimulator-like 4	-2.44684	0.00025
ENSG00000160211	2539	G6PD|glucose-6-phosphate dehydrogenase	-1.21866	5.00E-05
ENSG00000160285	4047	LSS|lanosterol synthase (2,3-oxidosqualene-lanosterol cyclase)	-1.12374	5.00E-05
ENSG00000160326	11182	SLC2A6|solute carrier family 2 (facilitated glucose transporter), member 6	-1.12955	5.00E-05
ENSG00000160404	27433	TOR2A|torsin family 2, member A	-1.62058	5.00E-05
ENSG00000160446	84885	ZDHHC12|zinc finger, DHHC-type containing 12	-1.41004	5.00E-05
ENSG00000160703	79671	NLRX1|NLR family member X1	-1.04198	0.0013
ENSG00000160789	4000	LMNA|lamin A/C	-1.37469	5.00E-05
ENSG00000160877	112939	NACC1|nucleus accumbens associated 1, BEN and BTB (POZ) domain containing	-1.09164	5.00E-05
ENSG00000161011	8878	SQSTM1|sequestosome 1	-1.15104	5.00E-05
ENSG00000161016	6132	RPL8|ribosomal protein L8	-1.01671	5.00E-05
ENSG00000161653	162417	NAGS|N-acetylglutamate synthase	-1.02547	0.00445
ENSG00000162104	115	ADCY9|adenylate cyclase 9	-1.14224	0.0039
ENSG00000162302	8986	RPS6KA4|ribosomal protein S6 kinase, 90kDa, polypeptide 4	-1.01178	0.0002
ENSG00000162767	93185	IGSF8|immunoglobulin	-1.28496	5.00E-

29		superfamily, member 8		05
ENSG00000162897	83953	FCAMR|Fc receptor, IgA, IgM, high affinity	-1.0373	5.00E-05
ENSG00000163430	11167	FSTL1|follistatin-like 1	-1.53001	0.00275
ENSG00000163702	84818	IL17RC|interleukin 17 receptor C	-1.2268	5.00E-05
ENSG00000163870	131601	TPRA1|transmembrane protein, adipocyte asscociated 1	-1.11974	5.00E-05
ENSG00000163931	7086	TKT|transketolase	-1.00824	5.00E-05
ENSG00000164896	10922	FASTK|Fas-activated serine/threonine kinase	-1.07824	5.00E-05
ENSG00000164897	83590	TMUB1|transmembrane and ubiquitin-like domain containing 1	-1.22881	0.0008
ENSG00000165178	654817	NCF1C|neutrophil cytosolic factor 1C pseudogene	-1.27805	5.00E-05
ENSG00000165233	84270	CARD19|chromosome 9 open reading frame 89	-1.44697	5.00E-05
ENSG00000165886	80019	UBTD1|ubiquitin domain containing 1	-1.08225	5.00E-05
ENSG00000166133	27079	RPUSD2|RNA pseudouridylate synthase domain containing 2	-1.01157	5.00E-05
ENSG00000166140	84936	ZFYVE19|zinc finger, FYVE domain containing 19	-1.02447	0.00435
ENSG00000166165	1152	CKB|creatine kinase, brain	-2.00337	5.00E-05
ENSG00000166189	79803	HPS6|Hermansky-Pudlak syndrome 6	-1.07737	5.00E-05
ENSG00000166484	5598	MAPK7|mitogen-activated protein kinase 7	-1.02829	5.00E-05
ENSG00000166816	197257	LDHD|lactate dehydrogenase D	-1.10143	0.00285
ENSG00000166831	348093	RBPMS2|RNA binding protein with multiple splicing 2	-1.04779	0.00085
ENSG00000166925	81628	TSC22D4|TSC22 domain family, member 4	-1.25566	5.00E-05
ENSG00000167173	56905	C15orf39|chromosome 15 open reading frame 39	-1.15298	0.0007
ENSG00000167302	146705	ENTHD2|ENTH domain containing 2	-1.79226	0.0001
ENSG00000167508	4597	MVD|mevalonate (diphospho) decarboxylase	-1.50744	5.00E-05
ENSG00000167578	53916	RAB4B|RAB4B, member RAS oncogene family	-1.0446	5.00E-05
ENSG00000167657	1613	DAPK3|death-associated protein kinase 3	-1.91505	5.00E-05
ENSG00000167703	124935	SLC43A2|solute carrier family 43 (amino acid system L transporter), member 2	-1.24864	5.00E-05
ENSG00000167	84798	C19orf48|chromosome 19 open	-1.0195	0.0056

47		reading frame 48		5
ENSG00000167779	3489	IGFBP6|insulin-like growth factor binding protein 6	-1.10574	5.00E-05
ENSG00000167797	10263	CDK2AP2|cyclin-dependent kinase 2 associated protein 2	-1.05986	5.00E-05
ENSG00000167895	147138	TMC8|transmembrane channel-like 8	-2.75204	5.00E-05
ENSG00000167962	90850	ZNF598|zinc finger protein 598	-1.11773	0.0013
ENSG00000168056	4054	LTBP3|latent transforming growth factor beta binding protein 3	-1.12471	0.0055
ENSG00000168071	283234	CCDC88B|coiled-coil domain containing 88B	-1.58332	0.00395
ENSG00000168096	124401	ANKS3|ankyrin repeat and sterile alpha motif domain containing 3	-1.20393	0.00395
ENSG00000168487	649	BMP1|bone morphogenetic protein 1	-1.11063	0.0003
ENSG00000168528	347735	SERINC2|serine incorporator 2	-1.73007	5.00E-05
ENSG00000169026	84179	MFSD7|major facilitator superfamily domain containing 7	-1.08125	0.0011
ENSG00000169188	27301	APEX2|APEX nuclease (apurinic/apyrimidinic endonuclease) 2	-1.21942	5.00E-05
ENSG00000169692	10555	AGPAT2|1-acylglycerol-3-phosphate O-acyltransferase 2	-1.47395	5.00E-05
ENSG00000169710	2194	FASN|fatty acid synthase	-1.67319	0.00155
ENSG00000169738	51181	DCXR|dicarbonyl/L-xylulose reductase	-1.2648	5.00E-05
ENSG00000169976	83443	SF3B5|splicing factor 3b, subunit 5, 10kDa	-1.05208	5.00E-05
ENSG00000170458	929	CD14|CD14 molecule	-1.1104	0.0002
ENSG00000170638	80305	TRABD|TraB domain containing	-1.48765	0.0003
ENSG00000170906	4696	NDUFA3|NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 3, 9kDa	-1.03021	0.00455
ENSG00000171105	3643	INSR|insulin receptor	-1.09733	0.002
ENSG00000171298	2548	GAA|glucosidase, alpha; acid	-1.31969	5.00E-05
ENSG00000172183	3669	ISG20|interferon stimulated exonuclease gene 20kDa	-1.3279	0.0035
ENSG00000172354	2783	GNB2|guanine nucleotide binding protein (G protein), beta polypeptide 2	-1.24667	5.00E-05
ENSG00000172375	9854	C2CD2L|C2CD2-like	-1.04602	0.0033
ENSG00000172572	1521	CTSW|cathepsin W	-1.00814	0.0001

43				5
ENSG00000172663	80194	TMEM134|transmembrane protein 134	-1.30309	5.00E-05
ENSG00000172724	6363	CCL19|chemokine (C-C motif) ligand 19	-1.52581	5.00E-05
ENSG00000173264	56834	GPR137|G protein-coupled receptor 137	-1.33788	5.00E-05
ENSG00000173369	713	C1QB|complement component 1, q subcomponent, B chain	-1.0098	5.00E-05
ENSG00000173372	712	C1QA|complement component 1, q subcomponent, A chain	-1.07418	5.00E-05
ENSG00000173457	26472	PPP1R14B|protein phosphatase 1, regulatory (inhibitor) subunit 14B	-1.07777	5.00E-05
ENSG00000173540	29925	GMPPB|GDP-mannose pyrophosphorylase B	-1.1117	0.00015
ENSG00000173546	1464	CSPG4|chondroitin sulfate proteoglycan 4	-1.1675	5.00E-05
ENSG00000174775	3265	HRAS|Harvey rat sarcoma viral oncogene homolog	-1.59219	5.00E-05
ENSG00000174886	126328	NDUFA11|NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 11, 14.7kDa	-1.18046	0.0001
ENSG00000174938	26470	SEZ6L2|seizure related 6 homolog (mouse)-like 2	-1.28382	5.00E-05
ENSG00000174939	253982	ASPHD1|aspartate beta-hydroxylase domain containing 1	-1.0931	0.00105
ENSG00000175573	83638	C11orf68|chromosome 11 open reading frame 68	-1.06967	5.00E-05
ENSG00000175756	54998	AURKAIP1|aurora kinase A interacting protein 1	-1.36108	5.00E-05
ENSG00000176101	8636	SSNA1|Sjogren syndrome nuclear autoantigen 1	-1.53953	5.00E-05
ENSG00000176170	8877	SPHK1|sphingosine kinase 1	-1.0981	0.0046
ENSG00000176454	254531	LPCAT4|lysophosphatidylcholine acyltransferase 4	-1.31541	0.0005
ENSG00000176973	23625	FAM89B|family with sequence similarity 89, member B	-1.24525	5.00E-05
ENSG00000176978	29952	DPP7|dipeptidyl-peptidase 7	-1.53086	5.00E-05
ENSG00000177030	10522	DEAF1|DEAF1 transcription factor	-1.01725	0.00495
ENSG00000177106	64787	EPS8L2|EPS8-like 2	-1.39058	0.0013
ENSG00000177542	79751	SLC25A22|solute carrier family 25 (mitochondrial carrier: glutamate), member 22	-1.57688	0.00015
ENSG00000177600	6181	RPLP2|ribosomal protein, large, P2	-1.01189	5.00E-05
ENSG00000178209	5339	PLEC|plectin	-1.41348	5.00E-05

ENSG00000178605	8225	GTPBP6|GTP binding protein 6 (putative)	-1.08428	5.00E-05
ENSG00000179409	50628	GEMIN4|gem (nuclear organelle) associated protein 4	-1.17819	5.00E-05
ENSG00000179593	247	ALOX15B|arachidonate 15-lipoxygenase, type B	-1.01864	5.00E-05
ENSG00000179922	147808	ZNF784|zinc finger protein 784	-1.28387	0.00235
ENSG00000180448	23526	HMHA1|histocompatibility (minor) HA-1	-1.99385	5.00E-05
ENSG00000181577	221416	C6orf223|chromosome 6 open reading frame 223	-1.01552	0.00015
ENSG00000182087	91304	TMEM259|transmembrane protein 259	-1.03373	5.00E-05
ENSG00000182487	654816	NCF1B|neutrophil cytosolic factor 1B pseudogene	-1.20503	5.00E-05
ENSG00000182504	79598	CEP97|centrosomal protein 97kDa	-2.16027	5.00E-05
ENSG00000183092	57596	BEGAIN|brain-enriched guanylate kinase-associated	-2.03118	0.00305
ENSG00000183570	54039	PCBP3|poly(rC) binding protein 3	-1.32106	0.0008
ENSG00000183684	10189	ALYREF|Aly/REF export factor	-1.21359	0.0002
ENSG00000183751	10607	TBL3|transducin (beta)-like 3	-1.49669	5.00E-05
ENSG00000184281	10078	TSSC4|tumor suppressing subtransferable candidate 4	-1.44619	0.00075
ENSG00000184489	11156	PTP4A3|protein tyrosine phosphatase type IVA, member 3	-1.09171	0.00025
ENSG00000184730	55911	APOBR|apolipoprotein B receptor	-1.1535	5.00E-05
ENSG00000184897	8971	H1FX|H1 histone family, member X	-1.03817	0.0001
ENSG00000185033	10509	SEMA4B|sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4B	-1.18097	5.00E-05
ENSG00000185187	59307	SIGIRR|single immunoglobulin and toll-interleukin 1 receptor (TIR) domain	-1.21446	0.0011
ENSG00000185201	10581	IFITM2|interferon induced transmembrane protein 2	-1.04698	5.00E-05
ENSG00000185504	80233	FAAP100|Fanconi anemia core complex associated protein 100	-1.02877	0.00285
ENSG00000185507	3665	IRF7|interferon regulatory factor 7	-1.52581	0.00025
ENSG00000185669	333929	SNAI3|snail family zinc finger 3	-1.06225	0.0014
ENSG00000185803	79581	SLC52A2|solute carrier family 52 (riboflavin transporter), member 2	-1.02906	0.00315

ENSG00000185885	8519	IFITM1|interferon induced transmembrane protein 1	-1.3133	0.00055
ENSG00000186010	51079	NDUFA13|NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 13	-1.0379	5.00E-05
ENSG00000186174	283149	BCL9L|B-cell CLL/lymphoma 9-like	-1.37879	0.0001
ENSG00000186501	84065	TMEM222|transmembrane protein 222	-1.30675	5.00E-05
ENSG00000186635	116985	ARAP1|ArfGAP with RhoGAP domain, ankyrin repeat and PH domain 1	-1.10892	5.00E-05
ENSG00000186891	8784	TNFRSF18|tumor necrosis factor receptor superfamily, member 18	-1.226	5.00E-05
ENSG00000186907	349667	RTN4RL2|reticulon 4 receptor-like 2	-1.7492	0.00185
ENSG00000187531	51547	SIRT7|sirtuin 7	-1.1814	0.00115
ENSG00000187608	9636	ISG15|ISG15 ubiquitin-like modifier	-1.68484	0.00015
ENSG00000187688	51393	TRPV2|transient receptor potential cation channel, subfamily V, member 2	-1.06984	5.00E-05
ENSG00000187838	57048	TMEM256-PLSCR3|phospholipid scramblase 3	-1.15105	0.00045
ENSG00000188372	7784	ZP3|zona pellucida glycoprotein 3 (sperm receptor)	-1.13303	5.00E-05
ENSG00000188566	27158	NDOR1|NADPH dependent diflavin oxidoreductase 1	-1.12096	0.0023
ENSG00000188976	26155	NOC2L|NOC2-like nucleolar associated transcriptional repressor	-1.05362	5.00E-05
ENSG00000189077	83862	TMEM120A|transmembrane protein 120A	-1.08651	5.00E-05
ENSG00000196365	9361	LONP1|lon peptidase 1, mitochondrial	-1.0442	5.00E-05
ENSG00000196453	27153	ZNF777|zinc finger protein 777	-1.06024	5.00E-05
ENSG00000196498	9612	NCOR2|nuclear receptor corepressor 2	-1.03301	0.00045
ENSG00000196576	23654	PLXNB2|plexin B2	-1.37984	0.00055
ENSG00000196843	10865	ARID5A|AT rich interactive domain 5A (MRF1-like)	-1.11183	5.00E-05
ENSG00000196878	3914	LAMB3|laminin, beta 3	-1.08183	0.0001
ENSG00000196924	2316	FLNA|filamin A, alpha	-1.22521	5.00E-05
ENSG00000197114	84619	ZGPAT|zinc finger, CCCH-type with G patch domain	-2.04932	0.00035
ENSG00000197150	11194	ABCB8|ATP-binding cassette, sub-family B (MDR/TAP), member 8	-1.24739	5.00E-05

ENSG00000197272	246778	IL27|interleukin 27	-1.55472	5.00E-05
ENSG00000197903	85236	HIST1H2BK|histone cluster 1, H2bk	-1.34856	5.00E-05
ENSG00000198026	63925	ZNF335|zinc finger protein 335	-1.1129	5.00E-05
ENSG00000198055	2870	GRK6|G protein-coupled receptor kinase 6	-1.15645	5.00E-05
ENSG00000198517	7975	MAFK|v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog K	-1.40139	5.00E-05
ENSG00000198917	51490	C9orf114|chromosome 9 open reading frame 114	-1.48285	0.001
ENSG00000205414			-1.02484	0.00085
ENSG00000211649		IGLV7-46|	-16.0296	0.0001
ENSG00000211893		IGHG2|	-1.60157	0.0032
ENSG00000211899		IGHM|	-1.09468	5.00E-05
ENSG00000213145	1396	CRIP1|cysteine-rich protein 1 (intestinal)	-1.586	5.00E-05
ENSG00000213689	11277	TREX1|three prime repair exonuclease 1	-1.27944	0.00345
ENSG00000213853	2013	EMP2|epithelial membrane protein 2	-1.63523	0.00085
ENSG00000213923	1454///102800317	CSNK1E|casein kinase 1, epsilon///CSNK1E|LOC400927-CSNK1E readthrough	-1.18937	0.00225
ENSG00000214063	7106	TSPAN4|tetraspanin 4	-1.36378	0.00045
ENSG00000218537		MIF-AS1|	-1.59613	0.0005
ENSG00000221968	3995	FADS3|fatty acid desaturase 3	-1.51122	5.00E-05
ENSG00000225783	440823	MIAT|myocardial infarction associated transcript (non-protein coding)	-1.24161	0.00425
ENSG00000226332			-1.49849	0.00195
ENSG00000228300	55009	C19orf24|chromosome 19 open reading frame 24	-2.10381	0.0063
ENSG00000230943	101927686	|uncharacterized LOC101927686	-1.06553	5.00E-05
ENSG00000235173	51236	HGH1|HGH1 homolog	-1.15194	0.001
ENSG00000237989	101928399	|uncharacterized LOC101928399	-1.12929	0.00015
ENSG00000238227	90120	C9orf69|chromosome 9 open reading frame 69	-1.00007	5.00E-05

ENSG00000239857	51608	GET4|golgi to ER traffic protein 4	-1.1015	0.00445
ENSG00000241945	5822///102724159	PWP2|PWP2 periodic tryptophan protein homolog (yeast)///PWP2|periodic tryptophan protein 2 homolog	-1.74723	0.0009
ENSG00000242802	9907	AP5Z1|adaptor-related protein complex 5, zeta 1 subunit	-1.48415	5.00E-05
ENSG00000243156	57553	MICAL3|microtubule associated monooxygenase, calponin and LIM domain containing 3	-1.00651	0.0045
ENSG00000249780			-1.07823	0.0002
ENSG00000254452			-1.35845	0.0001
ENSG00000254559			-1.1664	0.0026
ENSG00000254986	10072	DPP3|dipeptidyl-peptidase 3	-1.10415	0.00015
ENSG00000256007		ARAP1-AS1|	-1.71348	0.004
ENSG00000257156			-1.13237	0.0029
ENSG00000257663			-1.03551	0.0001
ENSG00000261236	23246	BOP1|block of proliferation 1	-1.50606	0.0024
ENSG00000261796	100534599	ISY1-RAB43|ISY1-RAB43 readthrough	-1.10893	0.00035
ENSG00000262049			-1.24797	0.00605
ENSG00000262413			-1.08612	0.0005
ENSG00000267436			-1.09667	0.00075
ENSG00000267519	284454	|uncharacterized LOC284454	-1.17642	0.0003
ENSG00000269858	112398	EGLN2|egl-9 family hypoxia-inducible factor 2	-1.00968	5.00E-05
ENSG00000269968			-1.1508	0.00025
ENSG00000272916			-1.14253	0.0008
ENSG00000273812			-1.43518	0.00015
ENSG00000275074	79873	NUDT18|nudix (nucleoside diphosphate linked moiety X)-type motif 18	-1.35046	5.00E-05
ENSG00000275294			-1.08204	5.00E-05

Table 2-5: Up regulated and down regulated pathway in cells infected with *ankH*.

Up regulated Pathway	P value	Down regulated pathway	P value
Development positive regulation of STK3/4 (Hippo) pathway and negative regulation of YAP/TAZ function	1.338e-9	Transcription, HIF-1 targets	2.822e-15
Transport clathrin coated vesicle cycle	2.291e-9	Immune response, IL-3 signaling via JAK/STAT, p38, JNK, and NFκB	1.745e-14
Apoptosis and survival, FAS signaling cascades	9.334e-9	Immune response, IL-1 signaling pathway	1.270e-11
Immune response, antigen presentation by MHC class I: cross-presentation	5.959e-8	Immune response, IL-10 signaling pathway	1.397e-11
Signal transduction, JNK pathway	6.740e-8	Apoptosis and survival, anti-apoptotic TNFs/NFκB/Bcl-2 pathway	2.725e-11

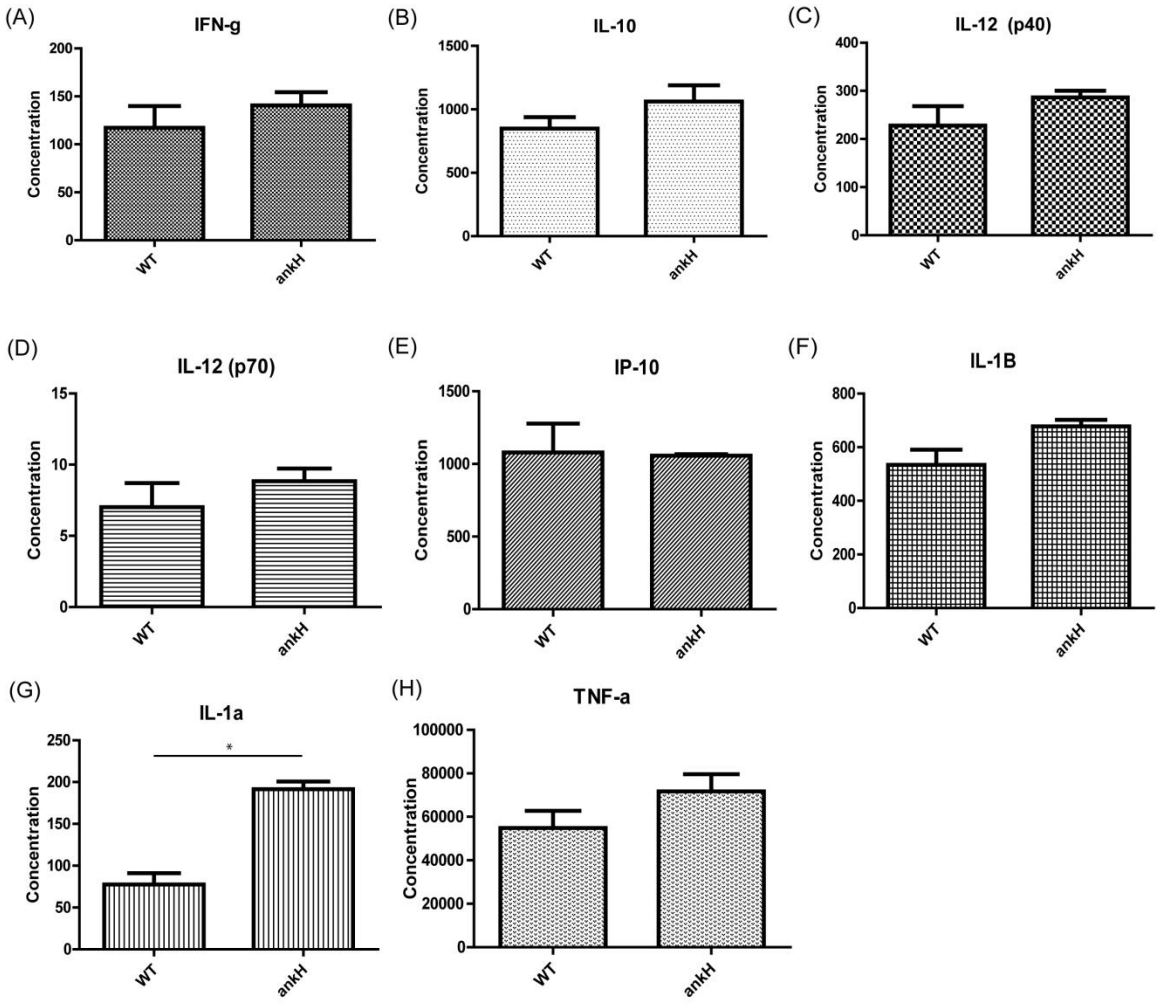

Figure 2-4. Cytokine production in cells infected with WT or ΔankH strains of L. pneumophila. (A-H) HMDMs were infected with either the WT or Δ*ankH* strains of *L. pneumophila* for 6 hours. Supernatants were collected and centrifuged to remove extracellular bacteria and cell debris. Samples were then used for Milliplex assay and cytokine concentrations were determined per assay instructions. The results are representative of one experiment performed in triplicate. Statistical analysis was performed using Student *t*-test where *, $P < 0.05$.

The crystal Structure of AnkH

AnkH is one of a few of the ~330 *Legionella* effectors required for intracellular growth within amoebae hosts and human macrophages [192, 210]. To get more insight into possible cellular function of AnkH, we have determined its three-dimensional crystal structure. AnkH is an α/β fold protein and contains a total of 21 α-helices and seven β-strands (Fig. 2-5A). AnkH consists of 3 domains: N-terminal ankyrin domain (α1-8, red), the middle domain (α10-17 and β3-7, cyan and magenta) and the cap domain (β1-2, α9 and α18-21, wheat [186, 187]. The N-terminal domain contains ankyrin repeats with four helix-turn-helix repeats (α1-α8, residues 1-122) (Fig. 2-5B). The first repeat is somewhat distorted and has shorter α-helices. The ARD is followed by a 4-turn-long helix α9 and an extended β-hairpin (β1-β2, residues 123-162) leading to the middle domain (Fig. 2-5A). This domain (residues 163-361) contains a central 5-stranded antiparallel β-sheet, β3-β7 and extended by helix α12. The β-sheet is flanked by two layers of two helices (inner α11, α16 and outer α10, α17) on one side and two helices α14-α15 on the other side. The C-terminal domain (residues 362-461) contains a five-helix bundle (Fig. 2-5A) and packs tightly together with α9 and the following β-hairpin forming one domain. The N- and C-terminal domains pack end to end into a crescent shape (Fig. 2-5A). The middle domain forms an independent insertion abated to the side of the ARD that is typically functioning as the protein binding surface. The long loops emanating from the ARD, usually involved in protein-protein interactions, face the middle domain.

The inserted middle domain of AnkH has a cysteine protease fold

To gain insight into possible functions of the middle and cap domains we have searched for their structural homologs using the Dali server [244]. The middle (insertion)

domain showed structural similarity to several proteins with cysteine protease fold albeit with relatively low scores. This cysteine protease-like domain (CPLD) is most similar to the outer protein D (XopD, PDB ID: 2OIX) from bacterial plant pathogen *Xanthomonas campestris pv. Vesicatoria* [245-247] (Fig. 2-5C). It also shows similarity to a domain of another *Legionella pneumophila* effector, RavZ [248, 249].

XopD belongs to the ubiquitin-like-specific protease 1 family [250] and is classified within Clan CE in the MEROPS database [251], with the catalytic triad is arranged in the order of histidine, glutamate/aspartate/asparagine and cysteine. Cysteine functions as a nucleophile while histidine serves as a general base and is in turn stabilized by glutamic acid/aspartic acid [250]. The structure-based sequence identity between the aligned regions of CPLD and XopD is only ~12%, nevertheless three β-strands and two α-helices are structurally similar between AnkH and XopD (Fig. 2-5C), with His243, Asp258 and Cys324 of AnkH superposed on the catalytic triad of XopD. The histidine resides on the N-terminal end of the conserved strand within the protease fold (β4 in AnkH, Fig 2-5A). The stabilizing aspartic acid sits at the C-terminal end of the conserved antiparallel strand (β5 in AnkH, Fig 2-5A). The cysteine nucleophile is at the end of a long loop leading to the penultimate helix of the protease fold (Fig. 2-5A). The orientation of these three sidechains in AnkH deviates from the active configuration and a small rearrangement of the triad sidechains has to occur to attain the active state (Fig. 2-5C). The fold of AnkH CPLD was recognized due to very low sequence identity to other cysteine protease and is not yet classified in the peptidase database MEROPS [252], which already includes several other peptidases from the *Legionella* species (data not shown).

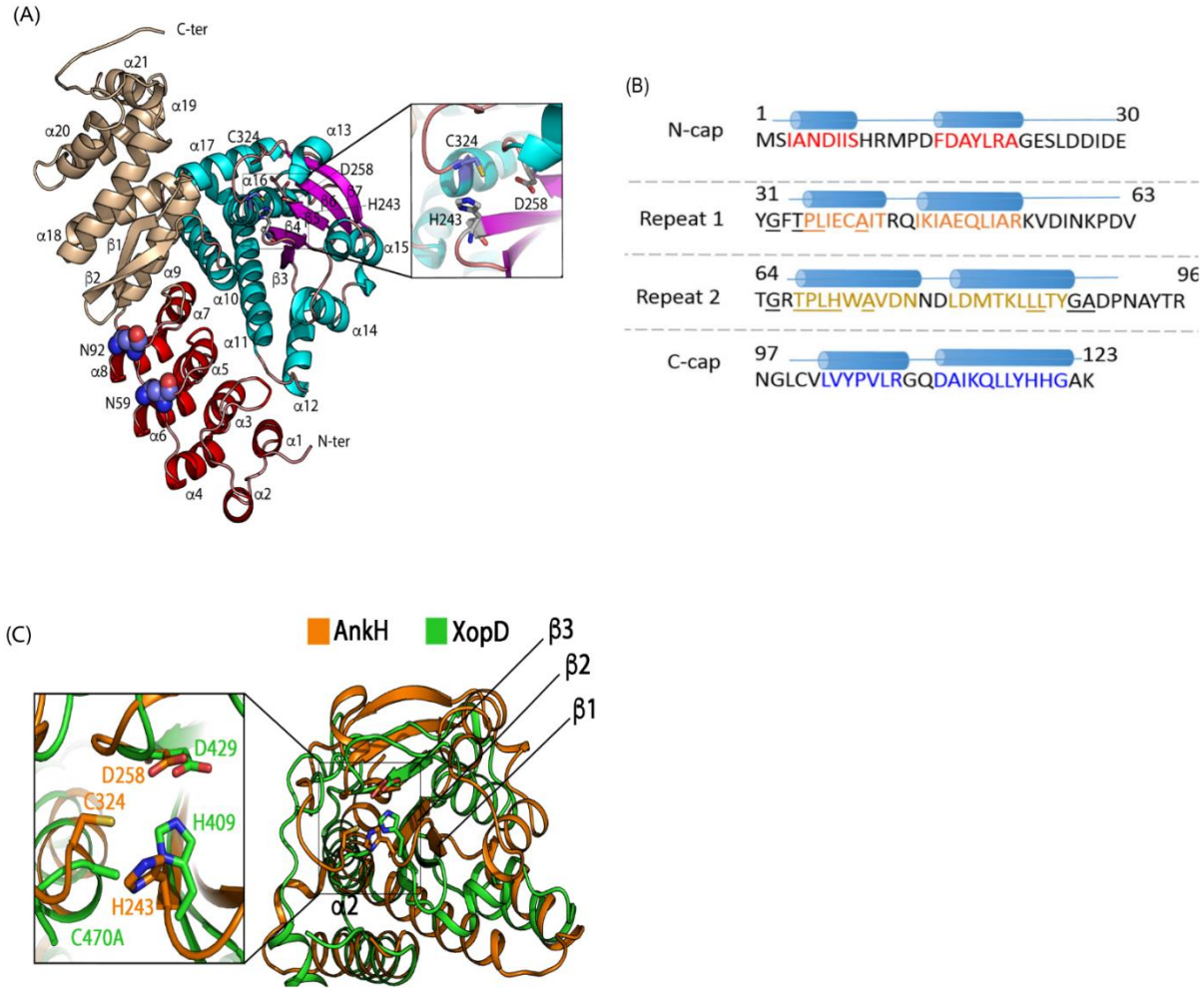

Figure 2-5. The crystal structure of AnkH. (A)AnkH consists of 3 domains: N-terminal ankyrin domain (α1-8, red), the cysteine proteinase-like domain (α10-17 and β3-7, cyan and magenta) and the cap domain (β1-2, α9 and α18-21, wheat). Inset shows the closeup of the putative catalytic triad residues H243, D258 and C324. The HIF hydroxylation sites (N59 and N92) are located within the N-terminal domain and are shown in a sphere representation (blue and red). (B) Primary sequence of ankryin domain. The length of each ankryin repeat was determined using the consensus sequence based on statistical analysis on 4,000 ankryin repeat sequence from the PFAM database as proposed by Mosavi *et al* [217]. Highlighted (colored) letters correspond to α-helices for each domain. The conserved residues are underlined, and the a-helices are shown as cylinders (C) Superposition of AnkH with *Xanthomonas* XopD C470A mutant. Cartoon diagram of superposed AnkH cysteine protease-like domain (residues 163-342, orange) and

Xanthomonas XopD C470A mutant (green, PBD ID:20IX, residues 336-515). The three β-strands and two α-helices that form the core of the domains and overlap well are marked. Inset shows the closeup of the catalytic triad. In AnkH it consists of His243, Asp258 and Cys324 and in XopD these residues are His409, Asp429 and Cys470.

Structure-function of AnkH

The structure of AnkH suggested that it binds cellular target(s) through the β-hairpin loops within ARD domain and has a predicted proteolytic activity (Fig. 2-5A). To better understand the roles of the AnkH domains and to validate its structure, a total of 12 residues were chosen for single substitutions based on their location within a specific domain (Table 2-6 and Fig. 2-5). The substituted residues included residues on the extended β-hairpin loops of ARs (Fig. 2-1 & Fig.2-5B), the putative cysteine protease catalytic triad and two asparagine residues (N59 and N92) that have been reported to undergo asparagine hydroxylation, which impacts protein-protein interactions [179]. Figure 2-6A illustrates each of the ARDs. Figure 2-6B & C illustrate the location of each substitution made within the ARDs. The mutations had no detectable effect on stability of the variant proteins in *L. pneumophila* (Fig. 2-6D) or during transient transfection (Fig. 6E).

HEK293T cells were co-transfected with LARP7 and either native AnkH or AnkH containing substitutions within the β-hairpin loops of the ARDs then immune-precipitated. Our data showed that substitutions of residues within the ARD3, specifically Asn97, diminished LARP7-AnkH interaction (Fig. 2-6F). In contrast, substitution of 30, 31, 33, 63, 64, or 96 resulted in enhanced binding between LARP7 and AnkH (Fig 2-6F).

Table 2-6: Point mutants generated in different domains of AnkH. ANK1,2,3 designate ankyrin repeat 1,2,3.

ANK1	ANK2	ANK3	Asn Hydrox	Cysteine-protease
E30T	V63Y	R96A	N59A	H243D
Y31S	T64E	N97V	N92A	D258A
F33A	V63Y/T64E	R96A/N97V	N59A/N92A	C324S
E30T/Y31S/F33A				H243D/D258A/C324S

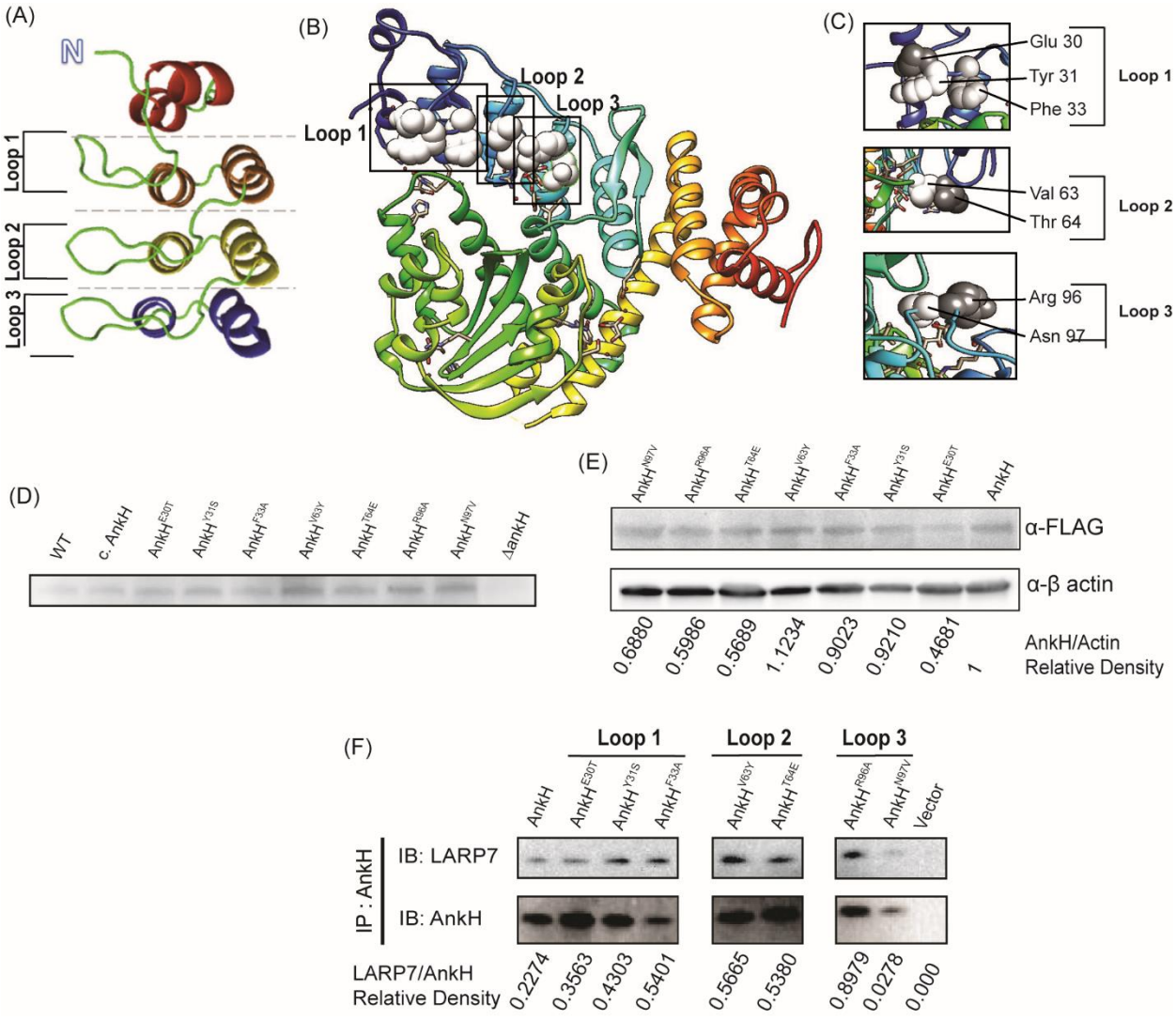

Figure 2-6. Substitutions in ARDs alters binding efficiency of AnkH and LARP7. (A) The ankryin domain of AnkH shown as ribbon diagram. The ankyrin domain consists of four ankryin repeats: N-cap, repeat 1, repeat2 and C-cap. (B&C) Crystal structure of AnkH illustrating different locations within the ARDs where residues were substituted. (D) Bacterial lysates from WT *L. pneumophila* and each of the AnkH substitution mutant strains were tested by immunoblot for AnkH to determine protein stability. Cell lysates were immunoblotted to detect the presence of AnkH using goat α-AnkH (53, 56). Equal number of bacteria were lysed for each strain. (E) HEK293T cells were transiently transfected with 3xFLAG-AnkH or the indicated 3xFLAG-AnkH substitution mutants and c-myc-LARP7. Densitometry was determined with actin ratio. (F) Cell lysates were immunoprecipitated with anti-FLAG antibody, and the co-IP was immunoblotted to detect the presence of AnkH and LARP7. Densitometry of the blots was determined as LARP7 to AnkH ratio. Results are representative of two independent experiments.

In order to determine if the substitutions affected the function of AnkH in intracellular replication of *L. pneumophila*, hMDMs were infected with the WT strain, the Δ*ankH* null mutant, Δ*ankH* mutant complemented with the WT allele of *ankH* or the substitution variants of AnkH. We first determined if the mutated constructs were translocated by the T4SS. One residue from each ANK domain (E30T, V63Y, N97V) was selected for mutation (Figure 2-7A&B). Translocation was determined using Cya-reporter fusions and measurement of cAMP which showed that all three representative mutant constructs were translocated (Figure 2-7A) and produced at equivalent concentrations by *L. pneumophila* (Figure 2-7B). Our data showed that substitution in the β-hairpin loop of ARD3, which led to a reduced binding of LARP7 to AnkH, resulted in reduced intracellular growth of *L. pneumophila* (Fig. 2-8). All other residues selected for substitutions were partially required for various degrees for AnkH function in intracellular replication, since introducing these mutations resulted in a various degrees of partial replication defect compared to the WT strain (Student *t*-test $p < 0.05$) (Fig. 2-8) (Table 2-6). Therefore, we conclude that the ARD, in particular Asn97, cysteine-like

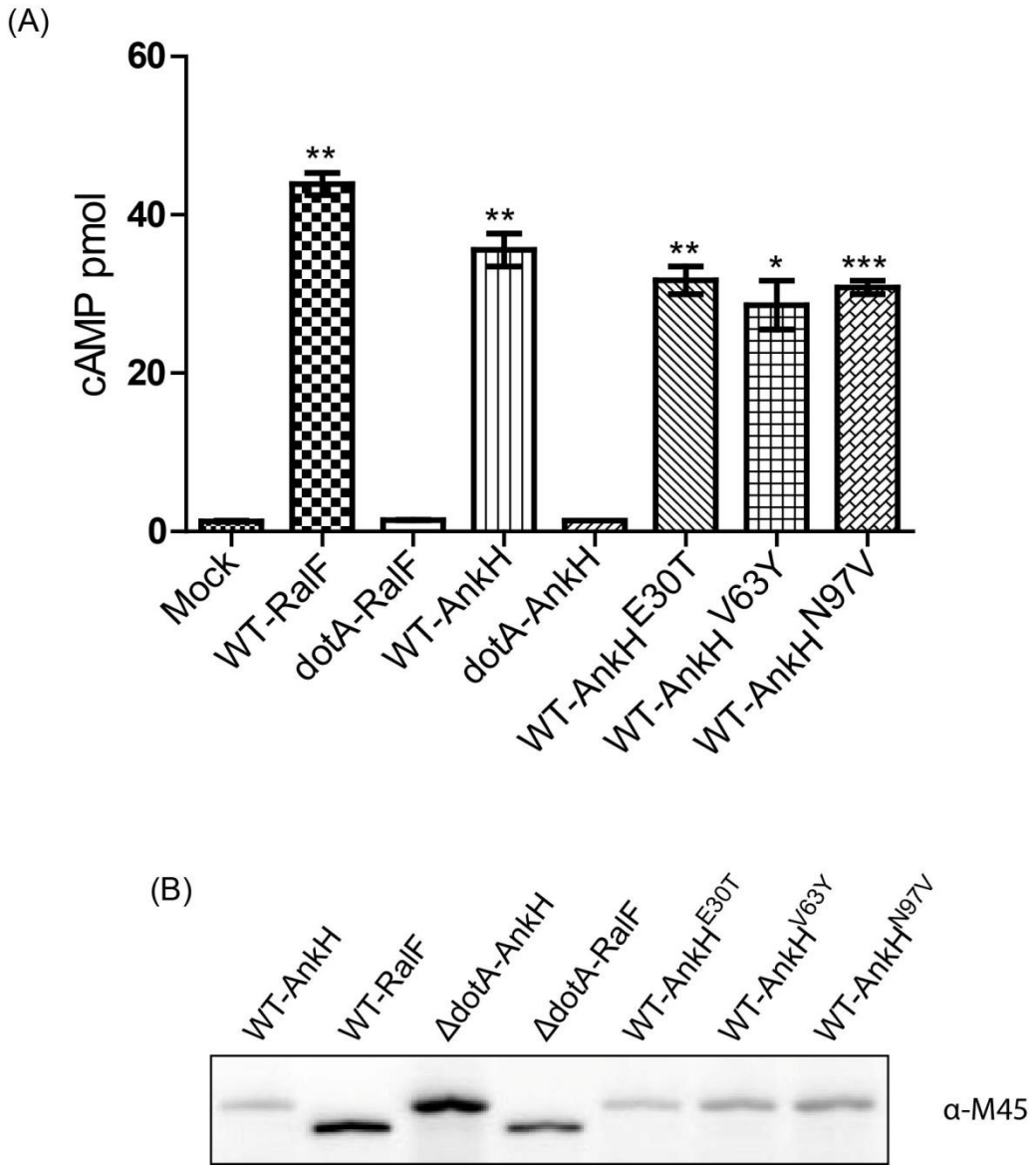

Figure 2-7. Translocation of AnkH ANK domain substitution mutants. (A) U937 cells were infected with WT or dotA strains of *L. pneumophila* harboring Cya contructs of full-length AnkH, RalF, AnkHE30T, AnkHV63Y, AnkHN97V at an MOI of 10 for 1 hour. Cells were lysed and cAMP levels were measured. (B) Bacterial strains used for cAMP assay were used to confirm protein production by *L. pneumophila*. $1X10^6$ bacteria were lysed and used for western blot analysis. The results are representative of experiment performed in triplicate. Statistical analysis was performed using Student *t*-test where *, $P < 0.05$.

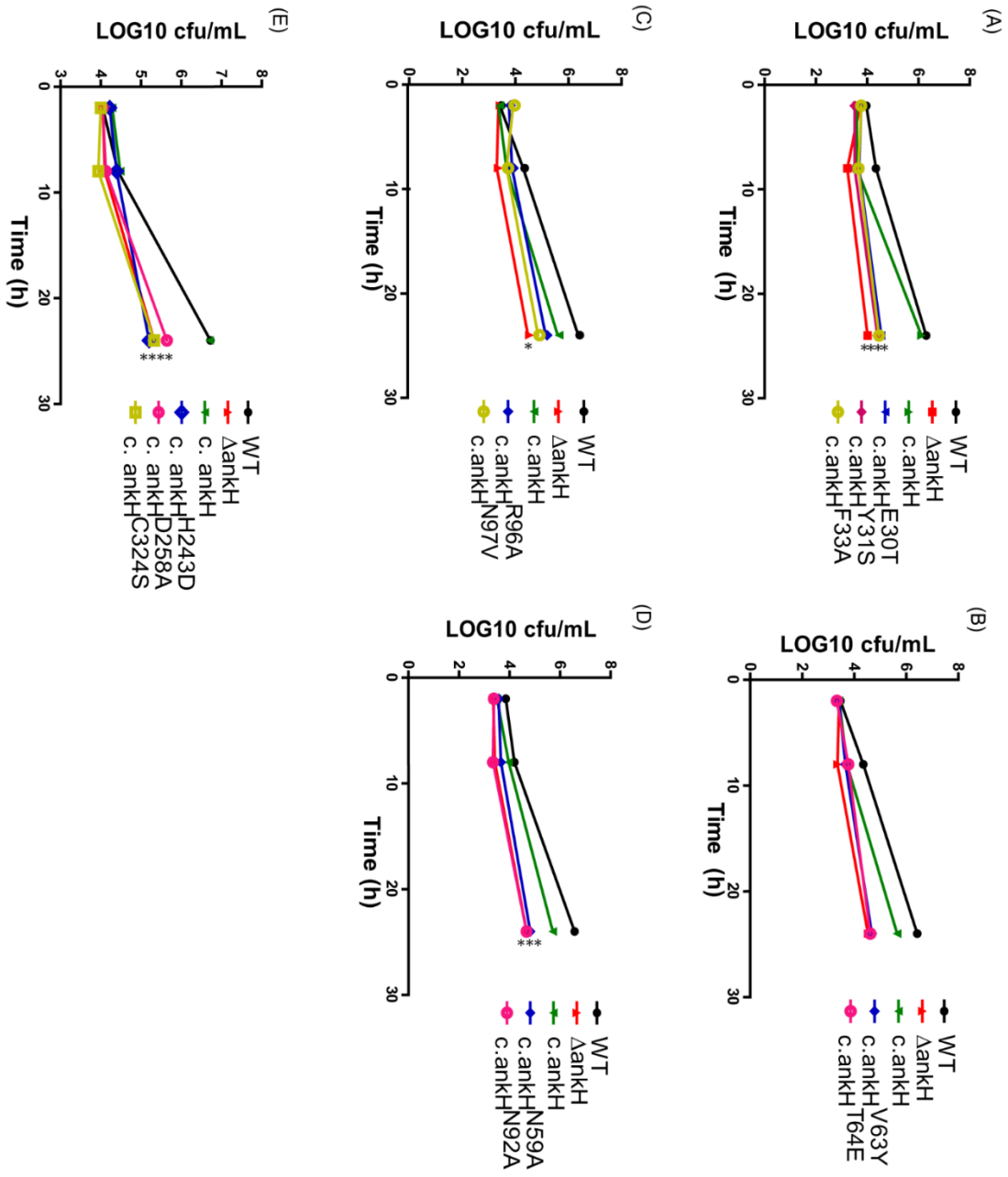

Figure 2-8. Structure-function of AnkH in intracellular growth of *L. pneumophila* within hMDMs. Intracellular growth kinetics were determined for WT strain, the *ankH* mutant, the *ankH* mutant complemented with the WT allele (c.ankH), or with single and multiple substitution variants as indicated. All strains in all the panels were tested using the same WT control. (A) Mutations within first ANK repeat, (B) second ANK repeat, (C) third ANK repeat, (D) asparagine hydroxylation motif and (E) cystine like protease pocket. The results are representative of three independent experiments performed in triplicate. Statistical analysis was performed using Student *t*-test where *, $P < 0.05$.

protease domain, and the asparagine hydroxylation motifs are all required for the function of AnkH in intracellular proliferation of *L. pneumophila* within hMDMs.

Materials and Methods

Bacterial strains and cell culture

L. pneumophila strain AA100/130b (BAA-74; American Type Culture Collection) and the isogeneic mutant's *dotA*, *ankH*, and complemented *ankH* mutants were grown on BCYE agar plates for 3 days at 37°C prior to use in infections, as described previously [54]. *E. coli* strain DH5-α was used for cloning purposes. Human monocyte-derived macrophages (hMDMs) were cultured using RPMI1640 media (Gibco), as described previously [150]. Maintenance of HEK293T cells was performed as previously described [150]. All methods were carried out and approved in accordance to the University of Louisville Institutional Review Board guidelines and blood donors gave informed consent as approved by the University of Louisville Institutional Review Board (IRB # 04.0358).

DNA manipulations

DNA manipulations and restriction enzyme digestions were performed using standard procedures [150, 253]. Restriction enzymes and T4 DNA ligase were purchased from NEB (Madison, WI). Plasmid preparations were performed with the PureLink HiPure Plasmid Maxiprep kit (Invitrogen). Purification of DNA fragments from agarose gels for subcloning was carried out with the QIAquick gel purification kit (Qiagen Inc, Valencia,

CA). Generation of AnkH substitution mutants was achieved using primers listed in Table 2-7 and described previously [150, 186].

Table 2-7. Primers used in this study.

ankH^{E30T}	F	CATATGGTTTTACTCCCCTCATAG
ankH^{E30T}	R	TATCGATATCATCCAAAGATTCCC
ankH^{Y31S}	F	CTGGTTTTACTCCCCTCATAGAGT
ankH^{Y31S}	R	ATTCATCGATATCATCCAA
ankH^{F33A}	F	CTACTCCCCTCATAGAGTGTGCCA
ankH^{F33A}	R	CACCATATTCATCGATATCA
ankH^{V63Y}	F	ACACAGGACGCACTCCATTACATT
ankH^{V63Y}	R	AGTCTGGCTTGTTGATATCCACTT
ankH^{T64E}	F	AAGGACGCACTCCATTACATTGGG
ankH^{T64E}	R	CGACGTCTGGCTTGTTGATA
ankH^{R96A}	F	CTAATGGTCTTTGTGTATTGGTTT
ankH^{R96A}	R	CAGTGTAGGCATTAGGATCAGCGC
ankH^{N97V}	F	TTGGTCTTTGTGTATTGGTTTATC
ankH^{N97V}	R	CACGAGTGTAGGCATTAGGA
ankH^{N59A}	F	CAAGCCAGACGTCACAGGACGC
ankH^{N59A}	R	CGATATCCACTTTTCGAGCAATTAA
ankH^{N92A}	F	TTGCCTACACTCGTAATGGTCTT
ankH^{N59A}	R	CAGGATCAGCGCCGTAGGTTAA
ankH^{H243D}	F	AATGCCTTATGCTTTGTC
ankH^{H243D}	R	GCCACGACTCGCCGCAGG
ankH^{D258A}	F	CCAGGGGTGAAAATAGCTTACAAG
ankH^{H243D}	R	CAATTTTTGCCCACCACTGGTGAT
ankH^{C324S}	F	AGTTCGTGGGCTAATGTG
ankH^{C324S}	R	ATTTCCACTAATTTGAGA

*All primers are 5'-phosphorylated. Orientation: F, forward; R, reverse

Translocation Assay

Legionella pneumophila strains AA100-Ralf-Cya, Δ*dotA*-RalF-Cya, AA100-AnkH-Cya, Δ*dotA*-RalF-Cya, AA100-AnkH^{E30T}-Cya, AA100-AnkH^{V63Y}-Cya, and AA100-AnkH^{N97V}-Cya were grown for on BCYE for 3 days at 37°C prior to infection. U937 cells were plated in 12 well plates at a concentration of 2 X 10^6 in triplicates and infected with *L. pneumophila* at an MOI of 10 for 1 hour. Cells were lysed using HCl + 0.1% Triton-X.

Levels of cAMP in cell lysates was analyzed using Direct cAMP ELISA (Enzo Life Sciences) kit per instructions.

Transfection of HEK293T cell

HEK293T cells were grown to ~70% confluent and plated onto poly-L-lysine-treated 24 well plates. Following 24 h of incubation, HEK293T cell monolayers were transfected with ~2 µg of plasmid DNA/well by using polyethylenimine (Polysciences) and OptiMem (Gibco) for 24 h, as described previously [89, 186]. The c-myc-LARP7 plasmid was a gift from B. Matija Peterlin, University of California, San Francisco.

Confocal laser scanning microscopy

Processing of transfected cells for confocal microscopy was performed as we described previously. Briefly, monolayers were permeabilized and fixed using 100% methanol held at -20°C for 5 min, and then blocked and labeled with mouse-anti-FLAG (1/200 dilution, Sigma, in 3% BSA-PBS), and rabbit-anti-Myc (1/200 dilution, ProteinTech, in 3% BSA-PBS). Cells were counter-labelled with Alexa-Fluor 488 anti-mouse antibody (1/4000 dilution, Invitrogen, 3% BSA-PBS), Alexa-Fluor 555 anti-rabbit antibody (1/4000 dilution, Invitrogen) and DAPI to stain the nuclei. Monolayers were examined by confocal microscopy. A total of 100 cells for each replicate were counted for presence or absence of localization.

Intracellular Replication

The wild type strain and the isogenic mutants, *dotA* and *ankH*, and the complemented *ankH* mutants were grown on BCYE for 3 days at 37°C prior to infection and used to infect hMDMs. A total of 1 X 10^5 host cells (hMDMs) per well were plated in 96 well

plates and infected with *L. pneumophila* at an MOI of 10 for 1 h and then treated for 1 h with gentamicin to kill remaining extracellular bacteria. Over a 24 h time course, the host cells were lysed with sterile water and *L. pneumophila* CFUs were determined by plating serial dilutions onto BCYE agar. Experiments were performed in triplicate.

Yeast two-hybrid (Y2H) analysis

The Matchmaker Gold Two-Hybrid system (Clontech) was used to screen host proteins that interact with the AnkH protein per manufactures instructions. Full length AnkH coding sequence was amplified, sequenced and cloned into the pGBKT7 bait vector (Clontech) and transformed into the AH109 yeast strain (Clontech). A normalized universal human cDNA library in pGADT7 was purchased (Clontech) to use as potential prey targets. The library and bait containing AH109 were mated and resulting colonies were screened per manufactures instructions. Plasmids from positive clones were isolated using yeast lysis buffer and glass beads. Isolated prey plasmid and bait plasmid were used to co-transform the AH109 yeast strain. Transformants were selected by growing the yeast on SD medium lacking His, Leu and Trp (SD-His/-Leu/-Trp) (Clontech). Positive colonies were then transferred to SD-Ade/-His/-Leu/-Trp plates containing 5-bromo-4-chloro-3-indoxyl-α-D-galactopyranoside (X-α-gal) (GoldBio). Blue colonies were selected for plasmid isolation. Isolated plasmids were then sequenced to determine the human genes.

In vivo Co-immunoprecipitation

HEK293T cells were transfected with 3XFLAG-AnkH, BAP, and c-myc-LARP7 for 24 h and collected in lysis buffer, as described previously [177, 254]. FLAG-tagged and myc-

tagged proteins were immunoprecipitated by using anti-FLAG M2 magnetic beads (Sigma) or SureBeads Protein G Magnetic Beads (BioRad) crosslinked with anti-myc antibody (ProteinTech).

Antibodies and western blot analysis

Legionella pneumophila strains were lysed using B-PER (Thermo Scientific) and heated at 99°C for 5 minutes in sample buffer. 1 X 10^6 bacteria were loaded per lane and separated by 10.4 to 15% SDS-PAGE (BioRad), and transferred onto a polyvinylidene difluoride (PVDF) (BioRad) membrane, as described previously [254] Immunoprecipitated proteins were heated at 99°C for 5 minutes in sample buffer, separated by 10.4 to 15% SDS-PAGE (BioRad), and transferred onto a polyvinylidene difluoride (PVDF) (BioRad) membrane, as described previously [254]. Anti-Flag (Sigma) used at 1:1000 dilution, anti-myc (60003-2-Ig) (ProteinTech) used at 1:1000 were incubated overnight in 8% milk at 4°C overnight. Anti-LaminB (13435) (Cell Signaling) was used at 1:1000 dilution. Anti-HEXIM1 (15676-1-AP), anti-LARP7 (17067-1-AP) and anti-MePCE (14917-1-AP) were purchased from ProteinTech and used at a 1:500 dilution. Anti-CDK9 (sc-13130) was purchased from Santa Cruz and used at a dilution of 1:200. Anti-CyclinT1 (sc-271348) was purchased from Santa Cruz and used at a dilution of 1:100. Goat Anti-AnkH antiserum was produced at Cocalico Biologics and was used at a dilution of 1:100 [186, 187].Cya-hybrids were detected using monoclonal M-45 antibody at a dilution of 1:50 [53].

RNA Isolation, Reverse transcription and Real-Time PCR

Total RNA was extracted using Trizol reagent (Invitrogen). cDNA synthesis was performed with 1 μg of total RNA using iScript™ cDNA Synthesis kit (Bio-Rad) according to the manufacturer's instructions. Endogenous mRNA levels were measured by real-time PCR analysis based on SYBR Green detection (Fermentas) with the Bio-Rad MiniOpticon real-time PCR system.

RNA Seq

Libraries were prepared using the TruSeq Stranded mRNA LT Sample prep kit Set A or Set B with poly-A enrichment (Illumina). One microgram of sample (in a volume of 50μl) were treated with RNA purification beads and denatured for 5 minutes at 65°C. Then the supernatant was discarded, and the beads were washed with bead wash buffer. Captured polyadenylated RNA was eluted using Elution buffer at 80°C for 2 min. mRNA is further purified in a second bead clean-up, as well as fragmented and primed during elution by adding 19.5μg of Elute, Prime, Fragment High Mix to the beads and incubating the samples for 8 minutes at 94°C. After fragmentation, 17μl of supernatant is removed from the beads and we proceeded immediately to synthesize first strand cDNA. Following the protocol, 8μl of First Strand Synthesis Mix Act D and SuperScript II mix (Illumina) was added to each sample and heated on a thermocycler using preprogramed thermal conditions. Once the reaction finished and reached 4°C, we immediately proceeded for second strand cDNA synthesis.

Diluted end repair control and Second Strand Marking Mix were added, mixed well and incubated in a pre-heated thermocycler at 16°C for one hour. The DNA was purified

using Agencourt AMPure XP Beads (Beckman). Finally, samples were eluted with resuspension buffer and 15µl of elute was collected and stored at -20°C.

A-Tailing control and A-Tailing mix were added to the purified samples and the samples were incubated on the preprogrammed thermal cycler. Once the incubation is done, we proceeded immediately to ligate adapters. Diluted ligation control, Ligation Mix and barcodes were added and incubated in a pre-heated thermocycler at 30°C for 10 minutes. Stop Ligation Buffer was immediately added to each sample and mixed well. Then the ligated samples were purified using Agencourt AMPure XP Beads. We eluted with 50µl of resuspension buffer and the elute was again purified for a second time using Agencourt AMPure XP Beads. Afterwards, the final elution, consisting of 20µl of the elute was collected and used for DNA enrichment. Samples were barcoded with Illumina TruSeq Adapters as listed Table 2-8. A complete list of the barcode sequences can be obtained from the Illumina support site (http://support.illumina.com/dam/illumina-support/documentation/chemistry_documentation/experiment-design/illumnia-adapter-sequences_1000000002694-01.pdf).

Table 2-8: Sample and Barcode Information

No.	Sample-BMDM	Barcodes Used
1	Control_1	2
2	Control_2	4
3	Control_3	5
4	Wildtype_1	6
5	Wildtype_2	7
6	Wildtype_3	12
10	ankH_1	16
11	ankH_2	18
12	ankH_3	19

PCR Primer Cocktail Mix and PCR Master Mix were added to the samples and incubated on a preprogrammed thermal cycler. Then the samples were purified using Agencourt AMPure XP Beads. Finally, 30µl of eluted library was collected and stored -20°C. Libraries were validated by quality where size, purity, and semi quantitation was performed on an Agilent Bioanalyzer using the Agilent DNA 1000 Kit. The final fragment size for all the samples was approximately 300bp which is expected according to the protocol. Libraries were also validated by quantity. Sequencing library quantitation was performed by qPCR using he KAPA library Quantitation Kit (KAPA Biosystems) for Illumina Platforms. The standard curve method was used for quantitation using 1-5 DNA standards that came with the kit.

Ten microliters of sample was transferred from the wells to a new MIDI plate. We then normalized the concentration of the libraries to 10nM using Tris-HCl 10mM, pH 8.5 with 0.1% Tween 20. Five microliters of each sample was then transferred to be pooled into a new LowBind 1.5ml micro centrifuge tube for a total volume of 60µl pooled 10nM library. Then, 4nM dilution was made from the 10nM pooled library by diluted with Tris-HCl 10mM, pH 8.5 with 0.1% Tween 20.

A total volume of 1.3ml of 1.8pM denatured library is needed for sequencing using v2 kit. Pooled 4nM library was denatured by mixing with diluted NaOH and incubated at room temperature for 5 minutes. Two hundred millimolar Tris HVl, pH 7.0 was then added. The reaction mixture was diluted to 20pM using a pre-chilled Hybridization buffer. Twenty picomolar denatured library was further diluted to 1.8pM using the same Hybridization buffer. Before loading onto the reagent cartridge, 1.3 µl of denatured 20pM

Phix control was added to the 1299µl of denatured 1.8pM library to a total volume of 1.3 ml for the sequencing run.

Sequencing was performed on the University of Louisville Center for Genetics and Molecular Medicines (CGeMM) Illumina NextSeq 500 using the NextSeq 500/550 1x75 cycle High Output Kit v2.

Milliplex Assay

Human monocyte derived macrophages were plated at a concentration of $2X10^6$ in 12 well plates. Cells were infected with either WT or Δ*ankH* strain of *L. pneumophila* for 6 hours at an MOI of 10. Cell supernatants were collected and used for assay. Milliplex assays (Millipore) were performed according to the manufacturer's instruction. Standards or culture supernatant samples were mixed with antibody-bound magnetic beads, and incubated overnight at 4 °C. Beads were washed and then incubated with the biotinylated detection antibody for one hour at room temperature. The beads were incubated with phycoerythrin-labeled streptavidin for thirty minutes at room temperature and the median fluorescent intensities were quantified with a Bio-plex 200 analyzer and analyzed with Bio-plex Manager 6.0 software. All samples were measured in duplicate.

RNAi Knockdown

Human LARP7 siRNA Lentivector against four LARP7 target sequences and scrambled siRNA GFP Lentivector were used with pLEnti-P2A, pLenti-P2B and Lentifectin to produce lentiviral particles per manufactures protocol (Applied Biological Materials, Inc). Lentiviral particles were mixed with complete RPMI (Corning) containing 8 µg/ml

polybrene (Milipore). Virus and media mixture was added to wells at 50 µL mixture per 1 mL of cells and incubated for 24 h.

Cloning ankH and protein expression

The ankH gene (Uniprot: Q5ZT65) from Legionella pneumophila strain Philadelphia 1 was cloned into vector pMCSG7, a derivative of vector pET-21a adapted for ligation-independent cloning **(**PMID: 18988021). This plasmid was then transformed into BL21 (DE3). The expressed protein contained a TEV-cleavable 6X-histidine tag at the N-terminus. For large-scale expression, a 15 mL overnight culture in LB was inoculated into 1 L of terrific broth medium (Bio Basic Inc. Markham, Ontario). The inoculated culture was grown at 37^{o}C and was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside when OD_{600} reached 0.6 and the temperature reduced to 18^{o}C for overnight growth. The cells were harvested by centrifugation at rpm of 9,000 x g for 15 min.

For expression of the Se-methionine derivative, the cell pellet from 100 mL of overnight culture grown in LB media was inoculated into 1 L of M9 minimal media. After shaking at 37^{o}C until OD_{600} reached 0.6, a mixture of L-amino acids (100 mg of lysine, phenylalanine, and threonine; 50 mg of isoleucine, leucine, and valine) and 60 mg of Se-methionine were added to the culture. Protein expression was induced with by adding 1 mM of Isopropyl β-D-1-thiogalactopyranoside after 15 minutes. The induced culture grew overnight at 18^{o}C and the cells were harvested by centrifugation at rpm of 9,000 x g for 15 min**.**

Protein purification

The cell pellet was re-suspended in lysis buffer (50 mM Tris-HCl buffer pH 8.0, 10% glycerol, and 0.1% Triton X). The cells were lysed in a cell disruptor (Constant Systems Ltd). The cell debris was removed by centrifugation at rpm of 31,000 x g for 30 min. The resulting supernatant was applied to a 3 mL TALON cobalt metal-affinity column (Clontech). The column was washed with 5 column volumes of standard buffer (20 mM Tris pH 8.0 and 50 mM NaCl). A step gradient containing 100 mM and 200 mM imidazole in standard buffer was used to elute the His-tagged protein. Fractions containing AnkH were pooled and loaded on a Superdex 200 10/300 GL column (GE Healthcare) equilibrated with crystallization buffer (15 mM Tris-HCl pH 8.0 and 100 mM NaCl). AnkH-containing fractions were pooled and concentrated to 5 mg/mL in a Millipore centrifugal filter with a molecular weight cut-off of 10,000 Da for-crystallization trials. The concentration was measured using the Nanodrop UV Spectrophotometer (Themo Scientific) using extinction coefficient of 70,250 for AnkH calculated by the ProtParam [255].

Protein crystallization and data collection

Initial crystals were obtained by screening and optimized using the 24-well plate format. The best crystals were obtained by the hanging-drop method by mixing 1 μL of protein solution and 1 μL of reservoir solution containing 1.0 M ammonium tartrate dibasic, pH 7.0. The drop was incubated over 0.5 mL reservoir solution. The crystals were cryoprotected in solution containing 70% of reservoir solution and 30% glycerol. Crystals were flash cooled in liquid nitrogen and diffraction data collected at the 08ID and 08BM

beamlines at the Canadian Light Source. Data were processed and scaled with XDS. The same procedure was followed for the Se-methionine labeled derivative.

Structure determination

The native and SeMet dataset were indexed, integrated and scaled using Program HKL3000 [256]. Experimental phases were obtained by single-wavelength anomalous dispersion (SAD) method and the structure was solved using Program HKL3000. The auto-built model from HKL3000 was ~90% complete and the remaining 10% of the molecule was built manually using program Coot (PMID: 20383002). The refinement was done using program suite Phenix [257]. The model contains residues 1-461 and was refined to R_{work}=0.172 and R_{free}=0.210. The geometry was validated with the program MolProbity [258]. The pertinent details of data collection and refinement are listed in Table 2-9. The coordinates and structure factors were deposited with the Protein Data Bank with the code 6MCA. Crystal structure was modeled using Chimera software (UCSF) and structure similarity to other peptidases was determined using peptidase database MEROPS [252].

Table 2-9. Data collection and refinement

	SeMet AnkH	**Native AnkH**
Data collection statistics		
Space group	$P6_5$ 2 2	$P6_5$ 2 2
a,b,c (Å), γ (º)	100.3, 100.3, 266.6, 120	102.1, 102.1, 266.0, 120
Wavelength (Å)	0.9788	0.9795
Resolution (Å)	50-2.9 (2.95-2.90)	51.1-2.45 (2.49-2.45)
Total Reflections	930045	643977
Unique reflections	18541	30933

R_{meas}	0.117 (0.882)	0.082 (0.855)
Completeness (%)	96.8 (94.7)	97.3 (94.4)
Redundancy	50.2 (49.8)	20.8 (21.3)
I/σ(*I*)	49.3 (6.0)	48.1 (6.1)
Wilson B ($Å^2$)	47.0	32.4
Refinement statistics		
R_{cryst} [d] / R_{free} [e] (%)		0.172 / 0.210
Rmsd on bonds (Å)		0.004
Rmsd on angles (º)		0.601
Favored (%)		98.25
Allowed (%)		1.75
PDB code		6MCA

Statistical analysis

All experiments were performed at least three independent biological repeats, and the data shown are representatives of one experiment. To analyze for statistically significant differences between three sets of data, the two-tailed Student's *t*-test was used, and the *p*-value was obtained.

Data Availability

CHAPTER 3:

DISCUSSION

An important step for the survival and replication of intracellular pathogens after infection of a host cell is to create an environment that supports the pathogen life cycle. Establishment of this environment is commonly accomplished through effector proteins [137, 138]. *Legionella* is a unique genus in that it codes for roughly 18,000 effectors proteins, many of which contain eukaryotic like protein domains and motifs. Among the 18,000 effector proteins, AnkH is the only effector conserved among all *Legionella* species, as well as other pathogens that harbor the Dot/Icm secretion system [137, 138]. While many *L. pneumophila* effectors are dispensable for intracellular growth of the pathogen in macrophages [156, 210, 259, 260], the AnkH effector plays an important role in intracellular growth of *L. pneumophila* within amoeba hosts and within macrophages [186, 187]. In addition, the high conservation of AnkH among many pathogenic obligate and facultative intracellular species of bacteria [137, 138] suggests it has a role in modulating an evolutionarily conserved eukaryotic process exploited by various obligate and facultative intracellular pathogens that translocate the AnkH effector by the Dot/Icm secretion system. Blast searches [261] with the nucleotide sequence of AnkH shows that in addition to various *Legionella* species homologous proteins are also found in Gammaproteobacteria species, *Coxiella* species, *Candidatus berkiella*, *Rickettsia* species, *Aquicella*, and *Legionella micdadei* [259].

AnkH is one of the many *L. pneumophila* effectors that contain eukaryotic like protein domains. AnkH contains ankyrin repeat domains (ARD) which are the most commonly found eukaryotic like protein domains found among *Legionella* effectors. The ARDs are capable of binding to multiple protein partners. Due to this possibility, it is likely that AnkH has multiple host interacting partners. We confirmed the host LARP7 is

an interacting partner for AnkH in HEK293T cells. LARP7 is a component of the 7SK snRNP complex which controls pausing of RNA polymerase II at the initiation of transcriptional elongation (see model in Fig. 3-1) [227-229]. Formation of the 7SK snRNP core complex (7SK, LARP7 and MePCE) enables recruitment of the P-TEFb and HEXIM1/2 to the complex [228, 231, 234-237]. Binding and sequestration of P-TEFb within the 7SK snRNP complex results in inhibition of its kinase activity and continued pause in Pol II transcription elongation [230, 233, 238, 239]. Various stimuli trigger the release of P-TEFb from the 7SK snRNP complex, leading to activation of its kinase activity, which is responsible for phosphorylating RNA pol II. This phosphorylation event ends the paused state of RNA pol II leading to productive transcriptional elongation [240, 241]. While LARP7 immunoprecipitates with AnkH, other components of the 7SK snRNP complex were not able to consistently be immunoprecipitated with AnkH, indicating that AnkH does not interact with a fully formed and functional 7SK snRNP complex.

The cellular localization of LARP7 corresponds to its function. Transcription takes place in the nucleus of cells and that is where LARP7 and the 7SK snRNP complex are found. Pathogenic effectors that modulate host transcription machinery are limited and the manipulation of the host 7SK snRNP complex via LARP7-AnkH interaction identifies a novel effector mechanism for host transcription control during infection. However, it is not known whether the interaction between AnkH and LARP7 and potentially other host targets has evolved during interaction of *L. penumophila* with various protist hosts in the aquatic environment to modulate amoeba hosts-specific gene transcription [137, 138, 210]. LARP7 is conserved in both human cells and amoeba.

Therefore, it is highly possible that some of the transcriptional activity impacted by the interaction of AnkH with LARP7 and other host targets in human macrophages may simply be an evolutionary accident [137, 138, 210]. Since knockdown of LARP7 results in a significant decrease in the intracellular replication of both WT and Δ*ankH* null mutant of *L. pneumophila,* it is likely that the AnkH-LARP7 interaction promotes transcription of genes involved in permissiveness to *L. pneumophila* in evolutionarily distant hosts. It was unexpected that LARP7 knockdown caused a significant decrease in intracellular replication of Δ*ankH* mutant. This could be explained by the hypothesis that AnkH does not interact with all LARP7 available within a host cell, which could create a balance between pause of transcription elongation and relief of the pause in elongation that creates a favorable environment for *L. pneumophila* replication. When AnkH is deleted and LARP7 is knocked down there is no longer a transcriptional balance. This may result in the decrease in replication as the result of unchecked transcription in hose cells leading to the alteration of many processes involved in permissiveness of the host cell to *L. pneumophila*.

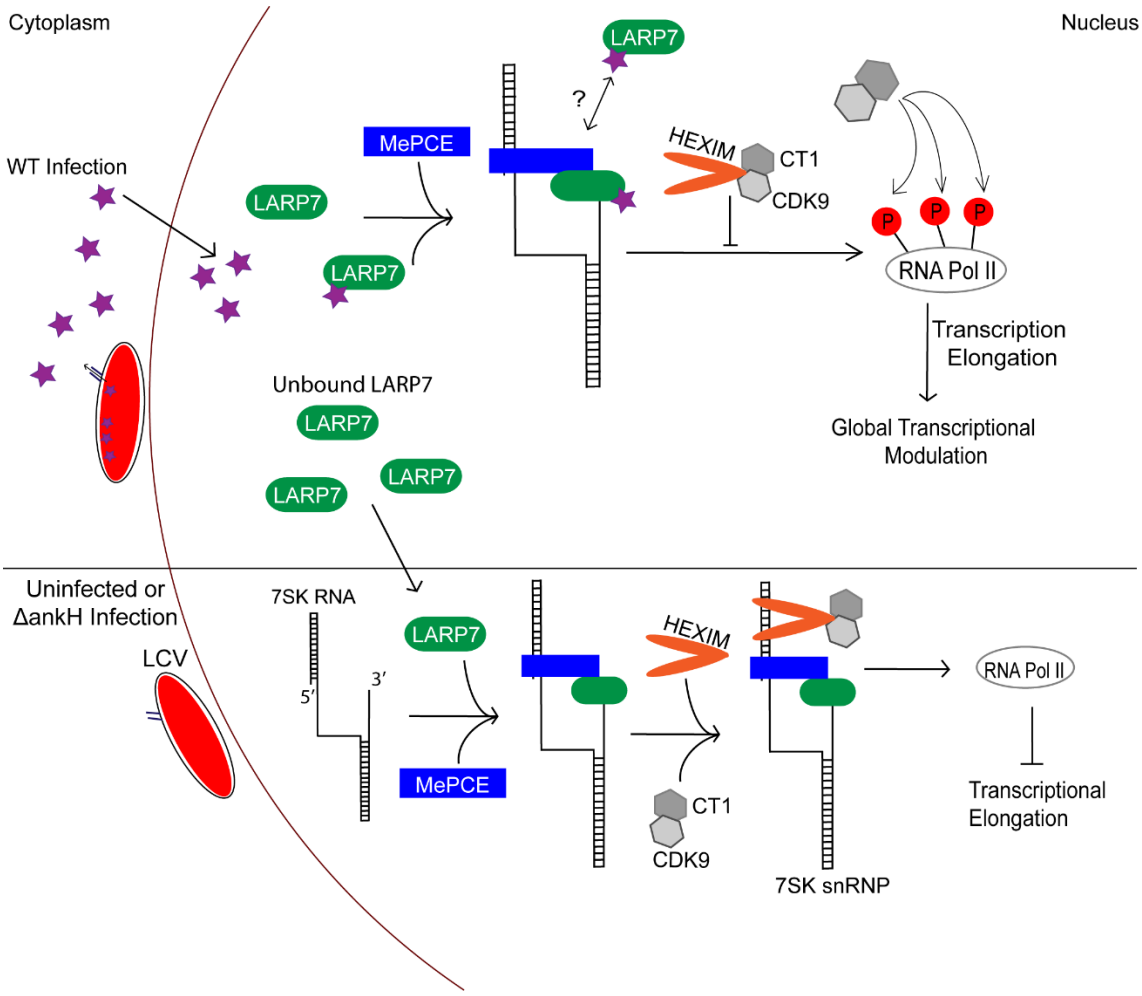

Figure 3-1. Working Model of AnkH-LARP7 Interaction. In un-infected cells or during Δ*ankH* mutant infection of HEK293T cells, formation of the 7SK snRNP begins when the 5' methyl capping enzyme (MePCE) and LARP7 are recruited to the 7SK snRNA forming the core of the 7SK snRNP. After core formation, the HEXIM 1/2 dimers as well as the P-TEFb (Cdk9 & CyclinT1) kinase are recruited to complete the 7SK snRNP complex, which prevents transcription elongation by holding RNA Polymerase II in a paused state. During infection with WT *L. pneumophila*, AnkH is trafficked to the nucleus where it interacts with a portion of available LARP7 in the cell. The interaction between AnkH and LARP7 results in a partial inhibition of the 7SK snRNP complex function leading to enhanced transcriptional elongation by blocking the recruitment of HEXIM1/2 and P-TEFb. The remaining LARP7 present in the cell (fraction that does not interact with AnkH) is available to interact with other components of the 7SK snRNP complex to pause transcription elongation by preventing P-TEFb from phosphorylating RNA Polymerase 2 keeping the polymerase in a paused state. This balance between the pause and relief of the pause in transcriptional elongation results in

transcriptional reprogramming within host cell that enhance permissiveness to *L. pneumophila* infection. There are likely other unidentified substrates of AnkH that could aid in this process or could act independently of the interaction with LARP7. The effect on amoeba host transcription by AnkH maybe different from human macrophages.

Our data indicate LARP7-AnkH interaction impedes 7SKsnRNP complex formation leading to transcriptional elongation of certain genes by Pol II resulting in host global transcriptional reprogramming. Translocation of AnkH into the host cell results in up regulation of pathways regulating transcription and immune responses in the host cell. However, in the absence of AnkH there is an upregulation in pathways involved in vesicular trafficking, autophagy and apoptosis. Due to the up regulation in immune response pathways, a series of cytokine levels were measured in response to infection. Ten cytokines levels were measured but only 8 were at at detectable levels. The multiplex data showed that of the cytokines tested, IL-1α was the only one that had levels that were significantly higher in cells infected with the Δ*ankH* compared to the cells infected with the WT strain. IL-1α was one of the cytokines identified as differentially regulated in the absence of AnkH and these multiplex data support those findings. These observations support our finding for the role of AnkH-LARP7 interaction in modulating function of the 7SK snRNP complex in human macrophages but the effect of AnkH on host global transcription is likely impacted by interaction of AnkH with other host targets.

Although AnkH had previously been studied, little was known about the crystal structure. The crystal structure revealed that AnkH contains four ARDs, two of which contain an asparagine hydroxylation motif located on the outer surface of the ARD domain. The crystal structure also revealed a cysteine-like protease pocket which had

previously not been detected based on secondary structure predictions. The ARDs are involved in protein-protein interactions by acting as a scaffold for proteins to bind. ARD-containing proteins can typically bind to one or more targets [54, 181]. The ARD domains contain multiple ankyrin repeats that form crescent-like structures and contact their binding partners on the concave side that is formed from the inner short helices and the long β-hairpin/loop regions connecting the ankyrin repeats [182]. Several residues on the putative target binding side of AnkH ARD that are located on the tips of the inter-repeat loops are required for the function of AnkH in intracellular replication of *L. pneumophila*. These sidechains are exposed to the solvent and aside from Tyr31 and Asn97, are not in contact with the cysteine protease-like domain. Therefore, mutation of these residues likely disrupts the ability of AnkH to interact with LARP7 or other specific host targets. Data has shown that residues within the β hairpin loops of the ARDs are involved in binding to substrates. Our data are consistent with these findings as we have shown that substitution in β hairpin loops of the ARD3, particularly at residue 97, results in reduced binding of AnkH to LARP7, indicating that this loop likely is the loop that is interacting with the LARP7 component of the 7SK snRNP complex. Not only do these mutations affect binding of AnkH and LARP7, but the residues within the ARDs also resulted in a defect in intracellular replication within human macrophages similar to the Δ*ankH* indicating that each repeat is required for the function of AnkH.

We have previously shown that AnkH is hydroxylated at N59 [179]. We have also shown that the host FIH asparagine hydroxylase localizes to the LCV and is involved in hydroxylating another *L. pneumophila* effector, AnkB [179]. Asparagine hydroxylation of AnkB is also required for the function of the AnkB effector in intracellular replication

of *L. pneumophila* [179]. The asparagine hydroxylation motif is commonly found in ARDs and serve as target sequences for the FIH asparagine hydroxylase [262, 263], which is responsible for hydroxylating an asparagine residue within this motif [179]. This hydroxylation can act as a molecular switch for protein-protein interactions by either inhibiting or strengthening the interaction [263-265]. The N59 and N92 residues of AnkH are located at the beginning of the loop connecting two neighboring ARDs. Mutation of either of these residues results in a significant decrease in intracellular replication of *L.pneumophila,* indicating that the asparagine hydroxylation motifs are important for the function of AnkH in the intracellular replication. A possible explanation for the role of this modification is provided by the structure of the ankyrin domain of the mouse notch 1 with this modification (PDB ID: 2QC9) [266]. The FIH-hydroxylated asparagine is located at a sharp bend of the backbone and hydrogen bonds through the added hydroxyl with the aspartic acid sidechain two residues back and located at the other corner of the bend. It has been suggested that this additional hydrogen bond might help to stabilize the loop in the ARD [266]. Equivalent aspartic acids are found in AnkH at positions 57 and 90, two back from the asparagines. Therefore, a similar possibility of stabilization of the inter-ARD loops has to be considered for AnkH as a means to strengthen the interaction with its cellular target.

The crystal structure also revealed a cysteine-like protease domain. The function of this domain is currently unknown. Our data show that the predicted protease catalytic triad is essential for the function of AnkH, but we were not able to detect protease activity *in vitro* for AnkH purified from *E. coli* (unpublished data) or cleavage of the interacting partner LARP7. We speculate that the lack of a detectable protease activity *in*

vitro is likely due to the closed nature of the catalytic pocket of purified AnkH, suggesting a requirement of its binding to a substrate *in vivo* to potentially open the pocket for catalysis. In most cases, homologs of AnkH contain all domains and in some cases the C-terminal domain is partially or fully missing. All these homologs conserve the His-Asp-Cys catalytic triad residues, which are embedded in conserved patterns: rG**H**a, **D/N**Rg and GN**C**SWANV that is preserved down to ~30% sequence identity with AnkH Cysteine-like protease domain. This would indicate that the cysteine-like protease domain is important for the function of AnkH since it is conserved in AnkH homologs.

In summary, AnkH is targeted to the nucleus where it interacts with LARP7 and likely other host targets, leading to reprograming of host transcription to promote intracellular bacterial growth. This is mediated, at least in part, by the effect of AnkH-LARP7 interaction and abolishment of interaction of LARP7 with critical subunits of the 7SK snRNP complex essential for its negative transcriptional elongation, leading to host global transcriptional reprogramming. The conservation of AnkH in intracellular pathogens harboring the Dot/Icm T4SS and its involvement in a conserved pathway supports AnkH-LARP7 interaction and its partial effect on reprogramming global host transcription, which is likely impacted by interaction of AnkH with other host targets. It is most likely the AnkH-dependent host transcriptional reprogramming to have unique consequences in various protist hosts compared to human macrophages. The crystal structure of AnkH shows it contains an ARD with four ankyrin repeats containing two asparagine hydroxylation motifs, a cysteine protease-like domain and a C-terminal domain of unknown function. Critical residues in the ARD and the cysteine protease-like

domains identified from the structure are shown to be required for AnkH-LARP7 interaction and function of AnkH in intracellular replication.

CHAPTER 4:

CONCLUSIONS AND FUTURE DIRECTIONS

No human proteins had previously been identified as interacting partners for AnkH. In addition, no pathogenic factor has ever been shown to interact with a host transcriptional complex. We show that the human LARP7 protein is an interacting partner of AnkH. LARP7 is a component of a transcription regulatory complex, the 7SK snRNP. After identifying LARP7 as an AnkH interacting partner we wanted to determine if other components of the 7SK snRNP complex were associated with AnkH. We were not able to detect other components of the 7SK snRNP complex during immunoprecipitation. Native protein levels were measured during immunoprecipitation of AnkH, so these proteins may not occur in high abundance which would make them harder to detect. One way we could further test if there were any interactions between the complex and AnkH would be to over express the different complex components and perform co-immunoprecipitation to determine if they are capable of interacting. It would also be beneficial to show that the interaction of AnkH and LARP7 is occurring during infection conditions and not just during ectopic expression of AnkH. An antibody suitable for detecting AnkH via immunofluorescence does not currently exist so it would need to be made in order to perform immunofluorescence during infection.

We show that AnkH has a subcellular localization to both the nucleus and the cytoplasm of transfected cells. AnkH does not contain a known nuclear localization signal (NLS) so the method by which it is being localized to the nucleus remains unknown. Through better immunofluorescence, with both more tags and using an AnkH specific antibody, it could more definitely be shown that AnkH localizes to the cell nucleus. The localization could also be tested in the presence of LARP7 knock down to determine if that alters the cellular localization of AnkH. There are multiple different

classes of NLS. It is possible that AnkH harbors a modified signal sequence that aids in transport to the nucleus of cells. There may also be an uncharacterized NLS located within AnkH which is what is responsible for localization. AnkH may complex with a host factor harboring an NLS, which would enable its nuclear localization.

Knockdown of LARP7 resulted in a significant defect in intracellular replication. We hypothesize that this is a result of the shutdown of transcriptional regulation within the host cell which creates an environment that negatively affects bacterial proliferation. It is possible that during infection AnkH interacts with a portion of the pool of LARP7 present within the host cell. This would leave room for partial transcriptional regulation through the 7SK snRNP complex. We hypothesize that a favorable environment for *L. pneumophila* is achieved by creating a balance between functional and non-functional 7SK snRNP transcriptional regulation through the LARP7-AnkH interaction. This theory could be tested using varying methods. One way to test this would be to determine the stability of the 7SK snRNA during infection, since LARP7 is required for formation of the complex and the 7SK snRNA is degraded when not part of the 7SK snRNP complex. This could be achieved by using RT-PCR to measure the amount of 7SK snRNA present during infection compared to uninfected cells or by using northern blots to test the stability of the 7SK snRNA. RNA-FISH could also be used to visualize 7SK snRNP location and concentration as well. To test if the 7SK snRNP complex is functioning, P-TEFb is responsible for phosphorylating RNA pol 2 but when it is sequestered in the 7SK snRNP complex it is not capable of performing its kinase activity. As a result, measuring the phosphorylation of RNA pol 2 would be an indicator of P-TEFb function which could show whether the 7SK snRNP complex is properly functioning. To test partial shutdown

of transcription as a result of AnkH the total RNA from infected cells could be compared to the total RNA from uninfected cells to quantitate transcription. Radiolabeled nucleotides could also be used to determine transcription levels by measuring their integration in mRNA. Another method could be to knockdown LARP7 and perform RNASeq. The 7SK snRNP complex does not function in the absence of LARP7 and comparing the transcriptome of RNAi treated cells to the transcriptome of infected cells could aid in determining if infection with *L. pneumophila* is indeed causing a partial shutdown in transcription.

The crystal structure of AnkH revealed that AnkH contains four ankyrin domains, a cysteine-like protease domain, two asparagine hydroxylation motifs and a CAP domain. Through point mutation of specific residues within the ANK domains, asparagine hydroxylation motif and the cysteine-like protease domain, we determined that all are important for the function of AnkH because mutation resulted in a decrease in intracellular proliferation. We also show that the ANK domains are required for AnkH-LARP7 interaction through co-immunoprecipitation of the point mutations and LARP7. Asn97, in particular, is likely important due to the observation that less LARP7 was able to be pulled down when this residue was mutated based on the amount of AnkH also present. It is also possible that loops 1 and/or 2 play a role in the interaction of AnkH and LARP7 as well. The effect the asparagine hydroxylation sites and the cysteine-like protease domain on the interaction between LARP7 and AnkH is unknown. This could be tested by performing co-immunoprecipitation of the AnkH Asn hydroxylation and cysteine-like protease domain mutants with LARP7. The ARD mutants would also need to be tested with other confirmed interacting partners to determine which loops are

responsible for the interaction between AnkH and its host targets. Another way the substitution mutants could be tested for their effect on AnkH interactions would be to utilize the yeast 2-hybrid system. The prey proteins have already been determined. The substitution mutants would have to be created in the AnkH bait construct and then used with the prey targets for yeast mating. If a mutation does alter AnkH binding to a target, then we would see results indicative of no interaction using this system.

The function of AnkH during infection of different protozoa hosts has not been elucidated. Further characterization of the host target and function of AnkH in amoeba would determine whether the function of AnkH is host specific and whether the response seen in human macrophages is an accidental response to an amoeba-adapted effector. Amoeba contain a LARP7 homolog. Therefore, the LARP7-AnkH interaction may also occur in amoeba but the affected cellular pathways in amoeba may be different from human macrophages. One way to test this would be to perform RNASeq on infected amoebae and compare the results to the human macrophage transcriptome results. This would help identify which pathways are affected in each host and to determine where the differences are. Similar pathways affected could be pursed further as they would likely be more indicative of the true function of AnkH.

While we have some answers as to the function of AnkH, it is still unclear how AnkH enables the survival and robust intracellular replication of *L. pneumophila* within target host cells. In order to more fully answer this question, the genes identified by RNASeq as being differentially regulated in the absence of AnkH and numerous pathways were identified by this screen, each one gives more insight into the function of AnkH and each would need to be further explored to determine how AnkH is altering

these pathways. Determining what other proteins identified in the yeast 2-hybrid screen will aid in answering this question since each will likely act on parts of different pathways.

REFERENCES

1. Honigsbaum, M., *Legionnaires' disease: revisiting the puzzle of the century.* The Lancet, 2016. **388**(10043): p. 456-457.
2. McDade, J.E., et al., *Legionnaires' disease: isolation of a bacterium and demonstration of its role in other respiratory disease.* N.Engl.J.Med., 1977. **297**: p. 1197-1203.
3. Brenner, D.J., A.G. Steigerwalt, and J.E. McDade, *Classification of the Legionnaires' disease bacterium: Legionella pneumophila, genus novum, species nova, of the family Legionellaceae, familia nova.* Ann Intern Med, 1979. **90**(4): p. 656-8.
4. Fields, B.S., R.F. Benson, and R.E. Besser, *Legionella and Legionnaires' disease: 25 years of investigation.* Clin Microbiol Rev, 2002. **15**(3): p. 506-26.
5. Newton, H.J., et al., *Molecular pathogenesis of infections caused by Legionella pneumophila.* Clin Microbiol Rev, 2010. **23**(2): p. 274-98.
6. Whiley, H. and R. Bentham, *Legionella longbeachae and legionellosis.* Emerg Infect Dis, 2011. **17**(4): p. 579-83.
7. Beaute, J., P. Zucs, and B. de Jong, *Legionnaires disease in Europe, 2009-2010.* Euro Surveill, 2013. **18**(10): p. 20417.
8. Yu, V.L., et al., *Distribution of Legionella species and serogroups isolated by culture in patients with sporadic community-acquired legionellosis: an international collaborative survey.* J Infect Dis, 2002. **186**(1): p. 127-8.
9. Burillo, A., M.L. Pedro-Botet, and E. Bouza, *Microbiology and Epidemiology of Legionnaire's Disease.* Infectious Disease Clinics of North America, 2017. **31**(1): p. 7-27.
10. Gump, D.W. and M. Keegan, *Pulmonary infections due to Legionella in immunocompromised patients.* Semin.Respir.Infect, 1986. **1**: p. 151-159.
11. Khweek, A.A. and A. Amer, *Replication of Legionella Pneumophila in Human Cells: Why are We Susceptible?* Front Microbiol, 2010. **1**: p. 133.
12. Dooling, K.L., et al., *Active Bacterial Core Surveillance for Legionellosis - United States, 2011-2013.* MMWR Morb Mortal Wkly Rep, 2015. **64**(42): p. 1190-3.
13. Marston, B.J., H.B. Lipman, and R.F. Breiman, *Surveillance for Legionnaires' disease. Risk factors for morbidity and mortality.* Arch Intern Med, 1994. **154**(21): p. 2417-22.
14. Bopp, L.H., et al., *Activities of tigecycline and comparators against Legionella pneumophila and Legionella micdadei extracellularly and in human monocyte-*

derived macrophages. Diagnostic Microbiology and Infectious Disease, 2011. **69**(1): p. 86-93.
15. Mandell, L.A., et al., *Infectious Diseases Society of America/American Thoracic Society Consensus Guidelines on the Management of Community-Acquired Pneumonia in Adults.* Clinical Infectious Diseases, 2007. **44**(Supplement_2): p. S27-S72.
16. Pedro-Botet, L. and V.L. Yu, *Legionella: macrolides or quinolones?* Clinical Microbiology and Infection, 2006. **12**: p. 25-30.
17. Phin, N., et al., *Epidemiology and clinical management of Legionnaires' disease.* Lancet Infect Dis, 2014. **14**(10): p. 1011-21.
18. Control, C.F.D. *Surveillance and Reporting* 2018 2/7/2018; Available from: https://www.cdc.gov/legionella/surv-reporting.html.
19. Leland, D.S. and R.B. Kohler, *Evaluation of the L-CLONE Legionella pneumophila serogroup 1 urine antigen latex test.* J.Clin.Microbiol., 1991. **29**: p. 2220-2223.
20. Mangiafico, J.A., K.W. Hedlund, and A.R. Knott, *Rapid and sensitive method for quantitation of Legionella pneumophila serogroup 1 antigen from human urine.* J.Clin.Microbiol., 1981. **13**: p. 843-845.
21. Mao, H., *[Detection of Legionella antigens in the urine by a biotin-avidin enzyme-linked immunosorbent assay and coagglutination test].* Chung.Hua.Liu.Hsing.Ping.Hsueh.Tsa.Chih., 1988. **9**: p. 103-106.
22. Berendt, R.F., et al., *Dose-response of guinea pigs experimentally infected with aerosols of Legionella pneumophila.* J.Infect.Dis., 1980. **141**: p. 186-192.
23. Davis, G.S., et al., *Legionnaires' pneumonia after aerosol exposure in guinea pigs and rats.* Am.Rev.Respir.Dis., 1982. **126**: p. 1050-1057.
24. Hambleton, P., et al., *Survival of virulent Legionella pneumophila in aerosols.* J.Hyg.(Lond), 1983. **90**: p. 451-460.
25. Muder, R.R., V.L. Yu, and A.H. Woo, *Mode of transmission of Legionella pneumophila. A critical review.* Arch.Intern.Med., 1986. **146**: p. 1607-1612.
26. Borges, V., et al., *Legionella pneumophila strain associated with the first evidence of person-to-person transmission of Legionnaires' disease: a unique mosaic genetic backbone.* Scientific Reports, 2016. **6**: p. 26261.
27. Correia, A.M., et al., *Probable Person-to-Person Transmission of Legionnaires' Disease.* N Engl J Med, 2016. **374**(5): p. 497-8.
28. Yu, V.L., et al., *Lack of evidence for person-to-person transmission of Legionnaires' disease.* J.Infect.Dis., 1983. **147**: p. 362.
29. Llewellyn, A.C., et al., *Distribution of Legionella and bacterial community composition among regionally diverse US cooling towers.* PLOS ONE, 2017. **12**(12): p. e0189937.
30. Biurrun, A., et al., *Treatment of a Legionella pneumophila -Colonized Water Distribution System Using Copper-Silver Ionization and Continuous Chlorination* •. Vol. 20. 1999. 426-8.
31. Darelid J, B.S., Jacobson K, Lofgren S, *The presence of a specific genotype of Legionella pneumophila serogroup 1 in a hospital and municipal water*

distribution system over a 12-year period Scand J Infect Dis, 2004. **36**(6-7): p. 417-423.
32. Kusnetsov, J., et al., *Copper and silver ions more effective against legionellae than against mycobacteria in a hospital warm water system.* Water Res, 2001. **35**(17): p. 4217-25.
33. Muraca, P., J.E. Stout, and V.L. Yu, *Comparative assessment of chlorine, heat, ozone, and UV light for killing Legionella pneumophila within a model plumbing system.* Appl.Environ.Microbiol., 1987. **53**: p. 447-453.
34. Schwake, D.O., A. Alum, and M. Abbaszadegan, *Impact of Environmental Factors on Legionella Populations in Drinking Water.* Pathogens, 2015. **4**(2): p. 269-282.
35. Breiman, R.F., et al., *Association of shower use with Legionnaires' disease: possible role of amoebae.* JAMA, 1990. **263**: p. 2924-2926.
36. Fields, B.S., et al., *Intracellular multiplication of Legionella pneumophila in amoebae isolated from hospital hot water tanks.* Curr.Microbiol., 1989. **18**: p. 131-137.
37. Fields, B.S., et al., *Proliferation of Legionella pneumophila as an intracellular parasite of the ciliated protozoan Tetrahymena pyriformis.* Appl.Environ.Microbiol., 1984. **47**: p. 467-471.
38. Hagele, S., et al., *Dictyostelium discoideum: a new host model system for intracellular pathogens of the genus Legionella.* Cell Microbiol, 2000. **2**(2): p. 165-71.
39. Newsome, A.L., et al., *Interactions between Naegleria fowleri and Legionella pneumophila.* Infect.Immun., 1985. **50**: p. 449-452.
40. Rowbotham, T.J., *Preliminary report on the pathogenicity of Legionella pneumophila for freshwater and soil amoebae.* J.Clin.Pathol., 1980. **33**: p. 1179-1183.
41. Rowbotham, T.J., *Current views on the relationships between amoebae, legionellae and man.* Isr.J.Med.Sci., 1986. **22**: p. 678-689.
42. Smith-Somerville, H.E., et al., *Survival of Legionella pneumophila in the cold-water ciliate Tetrahymena vorax.* Appl.Environ.Microbiol., 1991. **57**: p. 2742-2749.
43. Solomon, J.M., et al., *Intracellular growth of Legionella pneumophila in Dictyostelium discoideum, a system for genetic analysis of host-pathogen interactions.* Infect Immun, 2000. **68**(5): p. 2939-47.
44. Tyndall, R.L. and E.L. Domingue, *Cocultivation of Legionella pneumophila and free-living amoebae.* Appl.Environ.Microbiol., 1982. **44**: p. 954-959.
45. Kikuhara, H., et al., *Intracellular multiplication of Legionella pneumophila in Tetrahymena thermophila.* J UOEH, 1994. **16**(4): p. 263-75.
46. Tyson, J.Y., P. Vargas, and N.P. Cianciotto, *The novel Legionella pneumophila type II secretion substrate NttC contributes to infection of amoebae Hartmannella vermiformis and Willaertia magna.* Microbiology, 2014. **160**(Pt 12): p. 2732-2744.

47. Rasch, J., et al., *Legionella-protozoa-nematode interactions in aquatic biofilms and influence of Mip on Caenorhabditis elegans colonization.* International Journal of Medical Microbiology, 2016. **306**(6): p. 443-451.
48. Watanabe, K., et al., *Ciliate Paramecium is a natural reservoir of Legionella pneumophila.* Scientific Reports, 2016. **6**: p. 24322.
49. Scheikl, U., et al., *Free-living amoebae and their associated bacteria in Austrian cooling towers: a 1-year routine screening.* Parasitol Res, 2016. **115**(9): p. 3365-74.
50. Garcia, M.T., et al., *Acanthamoeba polyphaga resuscitates viable non-culturable Legionella pneumophila after disinfection.* Environ Microbiol, 2007. **9**(5): p. 1267-77.
51. Molmeret, M., et al., *Amoebae as training grounds for intracellular bacterial pathogens.* Appl. Environ. Microbiol., 2005. **71**: p. 20-28.
52. Berk, S.G., et al., *Production of respirable vesicles containing live Legionella pneumophila cells by two Acanthamoeba spp.* Appl.Environ.Microbiol., 1998. **64**: p. 279-286.
53. Cirillo, J.D., L.S. Tompkins, and S. Falkow, *Growth of Legionella pneumophila in Acanthamoeba castellanii enhances invasion.* Infect.Immun., 1994. **62**: p. 3254-3261.
54. Al-Khodor, S., et al., *A Dot/Icm-translocated ankyrin protein of Legionella pneumophila is required for intracellular proliferation within human macrophages and protozoa.* Mol Microbiol, 2008. **70**(4): p. 908-23.
55. Al-Khodor, S. and Y. Abu Kwaik, *Triggering Ras signalling by intracellular Francisella tularensis through recruitment of PKCalpha and betaI to the SOS2/GrB2 complex is essential for bacterial proliferation in the cytosol.* Cell Microbiol, 2010. **12**(11): p. 1604-21.
56. Al-Khodor, S., T. Al-Quadan, and Y. Abu Kwaik, *Temporal and differential regulation of expression of the eukaryotic-like ankyrin effectors of Legionella pneumophila.* Environmental Microbiology Reports, 2010. **2**(5): p. 677-684.
57. Price, C.T.D. and Y. Abu Kwaik, *Amoebae and Mammals Deliver Protein-Rich Atkins Diet Meals to Legionella Cells-Legionella pneumophila promotes proteasome-mediated protein degradation of its diverse hosts to meet its specific carbon and energy needs.* Microbe, 2012. **7**(11): p. 506.
58. Swanson, M.S. and B.K. Hammer, *Legionella pneumophila pathogesesis: a fateful journey from amoebae to macrophages.* Annu Rev Microbiol, 2000. **54**: p. 567-613.
59. Abu Kwaik, Y., et al., *Transcriptional regulation of the macrophage-induced gene (gspA) of Legionella pneumophila and phenotypic characterization of a null mutant.* Mol.Microbiol., 1997. **24**: p. 629-642.
60. Barker, J., et al., *Relationships between Legionella pneumophila and Acanthamoebae polyphaga: physiological status and susceptibility to chemical inactivation.* Appl.Environ.Microbiol., 1992. **58**: p. 2420-2425.

61. Barker, J., P.A. Lambert, and M.R.W. Brown, *Influence of intra-amoebic and other growth conditions on the surface properties of Legionella pneumophila.* Infect.Immun., 1993. **61**: p. 3503-3510.
62. Barker, J., H. Scaife, and M.R.W. Brown, *Intraphagocytic growth induces an antibiotic-resistant phenotype of Legionella pneumophila.* Antimicrob.Agents Chemother., 1995. **39**: p. 2684-2688.
63. Fields, B.S., et al., *Characterization of an axenic strain of Hartmannella vermiformis obtained from an investigation of nosocomial legionellosis.* J.Protozool., 1990. **37**: p. 581-583.
64. Ensminger, A.W. and R.R. Isberg, *Legionella pneumophila Dot/Icm translocated substrates: a sum of parts.* Curr Opin Microbiol, 2009. **12**(1): p. 67-73.
65. Abu Kwaik, Y., et al., *Invasion of protozoa by Legionella pneumophila and its role in bacterial ecology and pathogenesis.* Appl.Environ.Microbiol., 1998. **64**: p. 3127-3133.
66. Bandyopadhyay, P., et al., *Icm/dot-independent entry of Legionella pneumophila into amoeba and macrophage hosts.* Infect Immun, 2004. **72**(8): p. 4541-51.
67. Segal, G. and H.A. Shuman, *Legionella pneumophila utilizes the same genes to multiply within Acanthamoeba castellanii and human macrophages.* Infect.Immun., 1999. **67**: p. 2117-2124.
68. O'Brein, S.J. and R.S. Bhopal, *Legionnaires' disease: the infective dose paradox.* Lancet, 1993. **342**: p. 5-6.
69. Steinert, M., et al., *Resuscitation of viable but nonculturable Legionella pneumophila Philadelphia JR32 by Acanthamoeba castellanii.* Appl.Environ.Microbiol., 1997. **63**: p. 2047-2053.
70. Adeleke, A., et al., *Legionella-like amoebal pathogens-phylogenetic status and possible role in respiratory disease.* Emerg.Infect.Dis, 1996. **2**: p. 225-229.
71. Marrie, T.J., et al., *Legionella-like and other amoebal pathogens as agents of community-acquired pneumonia.* Emerg Infect Dis, 2001. **7**(6): p. 1026-9.
72. Bozue, J.A. and W. Johnson, *Interaction of Legionella pneumophila with Acanthamoeba catellanii: uptake by coiling phagocytosis and inhibition of phagosome-lysosome fusion.* Infect.Immun., 1996. **64**: p. 668-673.
73. Horwitz, M.A., *Phagocytosis of the Legionnaires' disease bacterium (Legionella pneumophila) occurs by a novel mechanism: engulfment within a pseudopod coil.* Cell, 1984. **36**: p. 27-33.
74. Swanson, M.S. and I. Sturgill-Koszycki, *Exploitation of macrophages as a replication niche by Legionella pneumophila [see comments].* Trends Microbiol, 2000. **8**(2): p. 47-9.
75. Molofsky, A.B. and M.S. Swanson, *Differentiate to thrive: lessons from the Legionella pneumophila life cycle.* Mol Microbiol, 2004. **53**(1): p. 29-40.
76. Rasis, M. and G. Segal, *The LetA-RsmYZ-CsrA regulatory cascade, together with RpoS and PmrA, post-transcriptionally regulates stationary phase activation of Legionella pneumophila Icm/Dot effectors.* Mol Microbiol, 2009.
77. Gao, L.-Y. and Y. Abu Kwaik, *Activation of caspase-3 in Legionella pneumophila-induced apoptosis in macrophages.* Infect.Immun., 1999. **67**(9): p. 4886-4894.

78. Gao, L.-Y. and Y. Abu Kwaik, *Apoptosis in macrophages and alveolar epithelial cells during early stages of infection by Legionella pneumophila and its role in cytopathogenicity.* Infect.Immun., 1999. **67**: p. 862-870.
79. Hägele, S., J. Hacker, and B.C. Brand, *Legionella pneumophila kills human phagocytes but not protozoan host cells by inducing apoptotic cell death.* FEMS Microbiol.Lett., 1998. **169**: p. 51-58.
80. Richards, A.M., et al., *Cellular microbiology and molecular ecology of Legionella-amoeba interaction.* Virulence, 2013. **4**(4): p. 307-14.
81. Amer, A.O. and M.S. Swanson, *A phagosome of one's own: a microbial guide to life in the macrophage.* Curr Opin Microbiol, 2002. **5**(1): p. 56-61.
82. Molmeret, M., et al., *Cell biology of the intracellular infection by Legionella pneumophila.* Microbes Infect, 2004. **6**(1): p. 129-39.
83. Molmeret, M., et al., *Temporal and spatial trigger of post-exponential virulence-associated regulatory cascades by Legionella pneumophila after bacterial escape into the host cell cytosol.* Environ Microbiol, 2010. **12**(3): p. 704-15.
84. Al-Khodor, S., et al., *The PmrA/PmrB two-component system of Legionella pneumophila is a global regulator required for intracellular replication within macrophages and protozoa.* Infect Immun, 2009. **77**(1): p. 374-86.
85. Bruggemann, H., et al., *Virulence strategies for infecting phagocytes deduced from the in vivo transcriptional program of Legionella pneumophila.* Cell Microbiol, 2006. **8**(8): p. 1228-40.
86. Byrne, B. and M.S. Swanson, *Expression of Legionella pneumophila virulence traits in response to growth conditions.* Infect.Immun., 1998. **66**: p. 3029-3034.
87. Molofsky, A.B. and M.S. Swanson, *Legionella pneumophila CsrA is a pivotal repressor of transmission traits and activator of replication.* Mol Microbiol, 2003. **50**(2): p. 445-61.
88. Faucher, S.P., C.A. Mueller, and H.A. Shuman, *Legionella Pneumophila Transcriptome during Intracellular Multiplication in Human Macrophages.* Front Microbiol, 2011. **2**: p. 60.
89. Al-Quadan, T., C. Price, and Y. Abu Kwaik, *Exploitation of evolutionarily conserved amoeba and mammalian processes by Legionella.* Trends Microbiol, 2012. **20**: p. 299-306, DOI: 10.1016/j.tim.2012.03.005.
90. McNealy, T.L., et al., *The Hfq homolog in Legionella pneumophila demonstrates regulation by LetA and RpoS and interacts with the global regulator CsrA.* J Bacteriol, 2005. **187**(4): p. 1527-32.
91. Oliva, G., et al., *A Unique cis-Encoded Small Noncoding RNA Is Regulating Legionella pneumophila Hfq Expression in a Life Cycle-Dependent Manner.* MBio, 2017. **8**(1).
92. Edwards, R.L., et al., *The Legionella pneumophila LetA/LetS two-component system exhibits rheostat-like behavior.* Infect Immun, 2010. **78**(6): p. 2571-83.
93. Hovel-Miner, G., et al., *SigmaS controls multiple pathways associated with intracellular multiplication of Legionella pneumophila.* J Bacteriol, 2009. **191**(8): p. 2461-73.

94. Sahr, T., et al., *Two small ncRNAs jointly govern virulence and transmission in Legionella pneumophila.* Mol Microbiol, 2009. **72**(3): p. 741-62.
95. Bitar, D.M., M. Molmeret, and Y. Abu Kwaik, *Molecular and cell biology of Legionella pneumophila.* Int J Med Microbiol, 2004. **293**(7-8): p. 519-27.
96. Dalebroux, Z.D., et al., *ppGpp conjures bacterial virulence.* Microbiol Mol Biol Rev, 2010. **74**(2): p. 171-99.
97. Haugen, S.P., W. Ross, and R.L. Gourse, *Advances in bacterial promoter recognition and its control by factors that do not bind DNA.* Nature reviews. Microbiology, 2008. **6**(7): p. 507-519.
98. Filloux, A., A. Hachani, and S. Bleves, *The bacterial type VI secretion machine: yet another player for protein transport across membranes.* Microbiology, 2008. **154**(6): p. 1570-1583.
99. Cianciotto, N.P., *Type II secretion: a protein secretion system for all seasons.* Trends Microbiol, 2005. **13**(12): p. 581-8.
100. Hales, L.M. and H.A. Shuman, *Legionella pneumophila contains a type II general secretion pathway required for growth in amoebae as well as for secretion of the Msp protease.* Infect.Immun., 1999. **67**: p. 3662-3666.
101. Hubber, A. and C.R. Roy, *Modulation of host cell function by Legionella pneumophila type IV effectors.* Annu Rev Cell Dev Biol, 2010. **26**: p. 261-83.
102. Cianciotto, N.P. and R.C. White, *The Expanding Role of Type II Secretion in Bacterial Pathogenesis and Beyond.* Infection and Immunity, 2017.
103. Peabody, C.R., et al., *Type II protein secretion and its relationship to bacterial type IV pili and archaeal flagella.* Microbiology, 2003. **149**(Pt 11): p. 3051-72.
104. Tyson, J.Y., et al., *Multiple Legionella pneumophila Type II secretion substrates, including a novel protein, contribute to differential infection of the amoebae Acanthamoeba castellanii, Hartmannella vermiformis, and Naegleria lovaniensis.* Infect Immun, 2013. **81**(5): p. 1399-410.
105. Rossier, O., J. Dao, and N.P. Cianciotto, *The type II secretion system of Legionella pneumophila elaborates two aminopeptidases, as well as a metalloprotease that contributes to differential infection among protozoan hosts.* Appl Environ Microbiol, 2008. **74**(3): p. 753-61.
106. Rossier, O., S.R. Starkenburg, and N.P. Cianciotto, *Legionella pneumophila type II protein secretion promotes virulence in the A/J mouse model of Legionnaires' disease pneumonia.* Infect Immun, 2004. **72**(1): p. 310-21.
107. Söderberg, M.A., et al., *Importance of Type II Secretion for Survival of Legionella pneumophila in Tap Water and in Amoebae at Low Temperatures.* Applied and Environmental Microbiology, 2008. **74**(17): p. 5583.
108. White, R.C. and N.P. Cianciotto, *Type II Secretion Is Necessary for Optimal Association of the Legionella-Containing Vacuole with Macrophage Rab1B but Enhances Intracellular Replication Mainly by Rab1B-Independent Mechanisms.* Infection and Immunity, 2016. **84**(12): p. 3313-3327.

109. Liles, M.R., P.H. Edelstein, and N.P. Cianciotto, *The prepilin peptidase is required for protein secretion by and the virulence of the intracellular pathogen Legionella pneumophila.* Mol.Microbiol., 1999. **31**: p. 959-970.
110. Rossier, O. and N.P. Cianciotto, *Type II protein secretion is a subset of the PilD-dependent processes that facilitate intracellular infection by Legionella pneumophila.* Infect Immun, 2001. **69**(4): p. 2092-8.
111. Rossier, O., J. Dao, and N.P. Cianciotto, *A type II secreted RNase of Legionella pneumophila facilitates optimal intracellular infection of Hartmannella vermiformis.* Microbiology (Reading, England), 2009. **155**(Pt 3): p. 882-890.
112. Söderberg, M.A. and N.P. Cianciotto, *A Legionella pneumophila Peptidyl-Prolyl cis-trans Isomerase Present in Culture Supernatants Is Necessary for Optimal Growth at Low Temperatures.* Applied and Environmental Microbiology, 2008. **74**(5): p. 1634-1638.
113. Duncan, C., et al., *Lcl of Legionella pneumophila Is an Immunogenic GAG Binding Adhesin That Promotes Interactions with Lung Epithelial Cells and Plays a Crucial Role in Biofilm Formation.* Infection and Immunity, 2011. **79**(6): p. 2168.
114. Johnston, C.W., et al., *Informatic analysis reveals Legionella as a source of novel natural products.* Synthetic and Systems Biotechnology, 2016. **1**(2): p. 130-136.
115. Stewart, C.R., D.M. Burnside, and N.P. Cianciotto, *The Surfactant of Legionella pneumophila Is Secreted in a TolC-Dependent Manner and Is Antagonistic toward Other Legionella Species.* Journal of Bacteriology, 2011. **193**(21): p. 5971.
116. Stewart, C.R., O. Rossier, and N.P. Cianciotto, *Surface translocation by Legionella pneumophila: a form of sliding motility that is dependent upon type II protein secretion.* J Bacteriol, 2009. **191**(5): p. 1537-46.
117. DebRoy, S., et al., *Legionella pneumophila type II secretome reveals unique exoproteins and a chitinase that promotes bacterial persistence in the lung.* Proc Natl Acad Sci U S A, 2006. **103**(50): p. 19146-51.
118. Mallama, C.A., K. McCoy-Simandle, and N.P. Cianciotto, *The Type II Secretion System of Legionella pneumophila Dampens the MyD88 and Toll-Like Receptor 2 Signaling Pathway in Infected Human Macrophages.* Infection and Immunity, 2017. **85**(4).
119. White, R.C. and N.P. Cianciotto, *Type II Secretion Is Necessary for Optimal Association of the Legionella-Containing Vacuole with Macrophage Rab1B but Enhances Intracellular Replication Mainly by Rab1B-Independent Mechanisms.* Infect Immun, 2016. **84**(12): p. 3313-3327.
120. McCoy-Simandle, K., et al., *Legionella pneumophila Type II Secretion Dampens the Cytokine Response of Infected Macrophages and Epithelia.* Infection and Immunity, 2011. **79**(5): p. 1984-1997.

121. Truchan, H.K., et al., *Type II Secretion Substrates of Legionella pneumophila Translocate Out of the Pathogen-Occupied Vacuole via a Semipermeable Membrane.* MBio, 2017. **8**(3).
122. Prashar, A. and M.R. Terebiznik, *Legionella pneumophila: homeward bound away from the phagosome.* Curr Opin Microbiol, 2015. **23**: p. 86-93.
123. Andrews, H.L., J.P. Vogel, and R.R. Isberg, *Identification of linked Legionella pneumophila genes essential for intracellular growth and evasion of the endocytic pathway.* Infect.Immun., 1998. **66**: p. 950-958.
124. Berger, K.H. and R.R. Isberg, *Two distinct defects in intracellular growth complemented by a single genetic locus in Legionella pneumophila.* Mol.Microbiol., 1993. **7**: p. 7-19.
125. Berger, K.H., J. Merriam, and R.R. Isberg, *Altered intracellular targeting properties associated with mutations in the Legionella pneumophila dotA gene.* Mol.Microbiol., 1994. **14**: p. 809-822.
126. Brand, B., A.B. Sadosky, and H.A. Shuman, *The Legionella pneumophila icm locus: a set of genes required for intracellular multiplication in human macrophages.* Mol.Microbiol., 1994. **14**: p. 797-808.
127. Purcell, M. and H.A. Shuman, *The Legionella pneumophila icmGCDJBF genes are required for killing of human macrophages.* Infect Immun, 1998. **66**: p. 2245-55.
128. Segal, G. and H.A. Shuman, *Characterization of a new region required for macrophage killing by Legionella pneumophila.* Infect Immun, 1997. **65**(12): p. 5057-66.
129. Vogel, J.P., et al., *Conjugative transfer by the virulence system of Legionella pneumophila.* Science, 1998. **279**: p. 873-876.
130. Nagai, H. and T. Kubori, *Type IVB Secretion Systems of Legionella and Other Gram-Negative Bacteria.* Front Microbiol, 2011. **2**: p. 136.
131. Sexton, J.A. and J.P. Vogel, *Type IVB secretion by intracellular pathogens.* Traffic, 2002. **3**(3): p. 178-85.
132. Sutherland, M.C., et al., *The Legionella IcmSW complex directly interacts with DotL to mediate translocation of adaptor-dependent substrates.* PLoS Pathog, 2012. **8**(9): p. e1002910.
133. Meir, A., et al., *Legionella DotM structure reveals a role in effector recruiting to the Type 4B secretion system.* Nat Commun, 2018. **9**(1): p. 507.
134. Ninio, S., et al., *The Legionella IcmS-IcmW protein complex is important for Dot/Icm-mediated protein translocation.* Mol Microbiol, 2005. **55**(3): p. 912-26.
135. Vincent, C.D., et al., *Identification of the DotL coupling protein subcomplex of the Legionella Dot/Icm type IV secretion system.* Mol Microbiol, 2012. **85**(2): p. 378-91.
136. Cambronne, E.D. and C.R. Roy, *The Legionella pneumophila IcmSW complex interacts with multiple Dot/Icm effectors to facilitate type IV translocation.* PLoS Pathog, 2007. **3**(12): p. e188.
137. Burstein, D., et al., *Genomic analysis of 38 Legionella species identifies large and diverse effector repertoires.* Nat Genet, 2016. **48**(2): p. 167-75.

138. Gomez-Valero, L., et al., *More than 18,000 effectors in the Legionella genus genome provide multiple, independent combinations for replication in human cells.* Proceedings of the National Academy of Sciences, 2019. **116**(6): p. 2265-2273.
139. Lifshitz, Z., et al., *Computational modeling and experimental validation of the Legionella and Coxiella virulence-related type-IVB secretion signal.* Proc Natl Acad Sci U S A, 2013. **110**(8): p. E707-15.
140. Zhu, W., et al., *Comprehensive identification of protein substrates of the Dot/Icm type IV transporter of Legionella pneumophila.* PLoS One, 2011. **6**(3): p. e17638.
141. Crabill, E., et al., *Dot/Icm-translocated proteins important for biogenesis of the Coxiella burnetii-containing vacuole identified by screening of an effector mutant sub-library.* Infection and immunity, 2018: p. IAI. 00758-17.
142. Bruckert, W.M., C.T. Price, and Y. Abu Kwaik, *Rapid nutritional remodeling of the host cell upon attachment of Legionella pneumophila.* Infect Immun, 2014. **82**(1): p. 72-82.
143. Nagai, H., et al., *A C-terminal translocation signal required for Dot/Icm-dependent delivery of the Legionella RalF protein to host cells.* Proc Natl Acad Sci U S A, 2005. **102**(3): p. 826-31.
144. Nagai, H., et al., *A bacterial guanine nucleotide exchange factor activates ARF on Legionella phagosomes.* Science, 2002. **295**(5555): p. 679-82.
145. Belyi, Y., et al., *Legionella pneumophila glucosyltransferase inhibits host elongation factor 1A.* Proc Natl Acad Sci U S A, 2006. **103**(45): p. 16953-8.
146. Fontana, M.F., et al., *Secreted bacterial effectors that inhibit host protein synthesis are critical for induction of the innate immune response to virulent Legionella pneumophila.* PLoS Pathog, 2011. **7**(2): p. e1001289.
147. Kubori, T., et al., *Legionella metaeffector exploits host proteasome to temporally regulate cognate effector.* PLoS Pathog, 2010. **6**(12): p. e1001216.
148. Laguna, R.K., et al., *A Legionella pneumophila-translocated substrate that is required for growth within macrophages and protection from host cell death.* Proc Natl Acad Sci U S A, 2006. **103**(49): p. 18745-50.
149. Neunuebel, M.R., et al., *Legionella pneumophila LidA affects nucleotide binding and activity of the host GTPase Rab1.* J Bacteriol, 2012. **194**(6): p. 1389-400.
150. Price, C.T., et al., *Molecular mimicry by an F-box effector of Legionella pneumophila hijacks a conserved polyubiquitination machinery within macrophages and protozoa.* PLoS Pathog, 2009. **5**(12): p. e1000704.
151. Tan, Y., R.J. Arnold, and Z.Q. Luo, *Legionella pneumophila regulates the small GTPase Rab1 activity by reversible phosphorylcholination.* Proc Natl Acad Sci U S A, 2011. **108**(52): p. 21212-7.
152. Jules, M. and C. Buchrieser, *Legionella pneumophila adaptation to intracellular life and the host response: clues from genomics and transcriptomics.* FEBS Lett, 2007. **581**(15): p. 2829-38.
153. Jeong, K.C., et al., *Polar delivery of Legionella type IV secretion system substrates is essential for virulence.* Proc Natl Acad Sci U S A, 2017. **114**(30): p. 8077-8082.

154. Gaspar, A.H. and M.P. Machner, *VipD is a Rab5-activated phospholipase A1 that protects Legionella pneumophila from endosomal fusion.* Proc Natl Acad Sci U S A, 2014. **111**(12): p. 4560-5.
155. Schlumberger, M.C., et al., *Real-time imaging of type III secretion: Salmonella SipA injection into host cells.* Proc Natl Acad Sci U S A, 2005. **102**(35): p. 12548-53.
156. Ghosh, S. and T.J. O'Connor, *Beyond Paralogs: The Multiple Layers of Redundancy in Bacterial Pathogenesis.* Front Cell Infect Microbiol, 2017. **7**: p. 467.
157. Luo, Z.Q. and R.R. Isberg, *Multiple substrates of the Legionella pneumophila Dot/Icm system identified by interbacterial protein transfer.* Proc Natl Acad Sci U S A, 2004. **101**(3): p. 841-6.
158. O'Connor, T.J., et al., *Minimization of the Legionella pneumophila genome reveals chromosomal regions involved in host range expansion.* Proc Natl Acad Sci U S A, 2011. **108**(36): p. 14733-40.
159. Kotewicz, K.M., et al., *A Single Legionella Effector Catalyzes a Multistep Ubiquitination Pathway to Rearrange Tubular Endoplasmic Reticulum for Replication.* Cell Host Microbe, 2017. **21**(2): p. 169-181.
160. Qiu, J., et al., *Ubiquitination independent of E1 and E2 enzymes by bacterial effectors.* Nature, 2016. **533**(7601): p. 120-4.
161. VanRheenen, S.M., et al., *Members of a Legionella pneumophila family of proteins with ExoU (phospholipase A) active sites are translocated to target cells.* Infect Immun, 2006. **74**(6): p. 3597-606.
162. Xu, L., et al., *Inhibition of host vacuolar H+-ATPase activity by a Legionella pneumophila effector.* PLoS Pathog, 2010. **6**(3): p. e1000822.
163. Allgood, S.C., et al., *Legionella Effector AnkX Disrupts Host Cell Endocytic Recycling in a Phosphocholination-Dependent Manner.* Front Cell Infect Microbiol, 2017. **7**: p. 397.
164. Machner, M.P. and R.R. Isberg, *Targeting of host Rab GTPase function by the intravacuolar pathogen Legionella pneumophila.* Dev Cell, 2006. **11**(1): p. 47-56.
165. Pan, X., et al., *Ankyrin repeat proteins comprise a diverse family of bacterial type IV effectors.* Science, 2008. **320**(5883): p. 1651-4.
166. Banga, S., et al., *Legionella pneumophila inhibits macrophage apoptosis by targeting pro-death members of the Bcl2 protein family.* Proc Natl Acad Sci U S A, 2007. **104**(12): p. 5121-6.
167. Weber, S.S., C. Ragaz, and H. Hilbi, *The inositol polyphosphate 5-phosphatase OCRL1 restricts intracellular growth of Legionella, localizes to the replicative vacuole and binds to the bacterial effector LpnE.* Cell Microbiol, 2009. **11**(3): p. 442-60.
168. Ge, J., et al., *A Legionella type IV effector activates the NF-kappaB pathway by phosphorylating the IkappaB family of inhibitors.* Proc Natl Acad Sci U S A, 2009. **106**(33): p. 13725-30.

169. Cazalet, C., et al., *Evidence in the Legionella pneumophila genome for exploitation of host cell functions and high genome plasticity.* Nat Genet, 2004. **36**(11): p. 1165-73.
170. de Felipe, K.S., et al., *Evidence for acquisition of Legionella type IV secretion substrates via interdomain horizontal gene transfer.* J Bacteriol, 2005. **187**(22): p. 7716-26.
171. Franco, I.S., H.A. Shuman, and X. Charpentier, *The perplexing functions and surprising origins of Legionella pneumophila type IV secretion effectors.* Cell Microbiol, 2009. **11**(10): p. 1435-43.
172. Gomez-Valero, L., et al., *Comparative and functional genomics of legionella identified eukaryotic like proteins as key players in host-pathogen interactions.* Front Microbiol, 2011. **2**: p. 208.
173. Gomez-Valero, L., et al., *Extensive recombination events and horizontal gene transfer shaped the Legionella pneumophila genomes.* BMC Genomics, 2011. **12**: p. 536.
174. Ninio, S., J. Celli, and C.R. Roy, *A Legionella pneumophila effector protein encoded in a region of genomic plasticity binds to Dot/Icm-modified vacuoles.* PLoS Pathog, 2009. **5**(1): p. e1000278.
175. Stone, B.J. and Y. Abu Kwaik, *Natural competency for DNA uptake by Legionella pneumophila and its association with expression of type IV pili.* J.Bacteriol., 1999. **181**: p. 1395-1402.
176. Lurie-Weinberger, M.N., et al., *The origins of eukaryotic-like proteins in Legionella pneumophila.* Int J Med Microbiol, 2010. **300**(7): p. 470-81.
177. Price, C.T., et al., *Exploitation of conserved eukaryotic host cell farnesylation machinery by an F-box effector of Legionella pneumophila.* J Exp Med, 2010. **207**(8): p. 1713-26.
178. Al-Quadan, T. and Y.A. Kwaik, *Molecular Characterization of Exploitation of the Polyubiquitination and Farnesylation Machineries of Dictyostelium Discoideum by the AnkB F-Box Effector of Legionella Pneumophila.* Front Microbiol, 2011. **2**: p. 23.
179. Price, C., et al., *Host FIH-Mediated Asparaginyl Hydroxylation of Translocated Legionella pneumophila Effectors.* Front Cell Infect Microbiol, 2017. **7**: p. 54.
180. Price, C.T. and Y. Abu Kwaik, *Exploitation of Host Polyubiquitination Machinery through Molecular Mimicry by Eukaryotic-Like Bacterial F-Box Effectors.* Front Microbiol, 2010. **1**: p. 122.
181. Li, J., A. Mahajan, and M.D. Tsai, *Ankyrin repeat: a unique motif mediating protein-protein interactions.* Biochemistry, 2006. **45**(51): p. 15168-78.
182. Mosavi, L.K., et al., *The ankyrin repeat as molecular architecture for protein recognition.* Protein Sci, 2004. **13**(6): p. 1435-48.
183. Wright, L.P. and M.R. Philips, *Thematic review series: lipid posttranslational modifications. CAAX modification and membrane targeting of Ras.* J Lipid Res, 2006. **47**(5): p. 883-91.
184. Price, C.T., et al., *Host proteasomal degradation generates amino acids essential for intracellular bacterial growth.* Science, 2011. **334**(6062): p. 1553-7.

185. Price, C. and Y. Abu Kwaik, *Amoebae and Mammals Deliver Protein-Rich Atkins Diet Meals to Legionella.* Microbe, 2012. **7**: p. 506-513.
186. Habyarimana, F., et al., *Molecular characterization of the Dot/Icm-translocated AnkH and AnkJ eukaryotic-like effectors of Legionella pneumophila.* Infect Immun, 2010. **78**(3): p. 1123-34.
187. Habyarimana, F., et al., *Role for the Ankyrin eukaryotic-like genes of Legionella pneumophila in parasitism of protozoan hosts and human macrophages.* Environ Microbiol, 2008. **10**(6): p. 1460-74.
188. Von Dwingelo, J., et al., *Interaction of the Ankyrin H Core Effector of Legionella with the Host LARP7 Component of the 7SK snRNP Complex.* mBio, 2019. **10**(4): p. e01942-19.
189. Abu Khweek, A. and A.O. Amer, *Factors Mediating Environmental Biofilm Formation by Legionella pneumophila.* Front Cell Infect Microbiol, 2018. **8**: p. 38.
190. Fields, B.S., *The molecular ecology of legionellae.* Trends.Microbiol, 1996. **4**: p. 286-290.
191. Horwitz, M.A. and S.C. Silverstein, *Legionnaires' disease bacterium (Legionella pneumophila) multiples intracellularly in human monocytes.* J.Clin.Invest., 1980. **66**: p. 441-450.
192. Best, A.M. and Y. Abu Kwaik, *Evasion of phagotrophic predation by protist hosts and innate immunity of metazoan hosts by Legionella pneumophila.* Cell Microbiol, 2019. **21**(1): p. e12971.
193. Boamah, D.K., et al., *From Many Hosts, One Accidental Pathogen: The Diverse Protozoan Hosts of Legionella.* Frontiers in Cellular and Infection Microbiology, 2017. **7**(477).
194. Wu, H.-Y., et al., *Legionella qingyii sp. nov., isolated from water samples in China.* International Journal of Systematic and Evolutionary Microbiology, 2019. **69**(7): p. 2017-2022.
195. Gomez-Valero, L., et al., *Comparative analyses of Legionella species identifies genetic features of strains causing Legionnaires' disease.* Genome Biol, 2014. **15**(11): p. 505.
196. Isberg, R.R., T.J. O'Connor, and M. Heidtman, *The Legionella pneumophila replication vacuole: making a cosy niche inside host cells.* Nat Rev Microbiol, 2009. **7**: p. 13-24.
197. Luo, Z.Q., *Legionella secreted effectors and innate immune responses.* Cell Microbiol, 2011.
198. Molmeret, M., et al., *Amoebae as training grounds for intracellular bacterial pathogens.* Appl Environ Microbiol, 2005. **71**(1): p. 20-8.
199. Oliva, G., T. Sahr, and C. Buchrieser, *The Life Cycle of L. pneumophila: Cellular Differentiation Is Linked to Virulence and Metabolism.* Frontiers in Cellular and Infection Microbiology, 2018. **8**(3).
200. Shin, S. and C.R. Roy, *Host cell processes that influence the intracellular survival of Legionella pneumophila.* Cell Microbiol, 2008. **10**(6): p. 1209-20.
201. Swart, A.L., et al., *Acanthamoeba and Dictyostelium as Cellular Models for Legionella Infection.* Frontiers in Cellular and Infection Microbiology, 2018. **8**(61).

202. Coers, J., C. Monahan, and C.R. Roy, *Modulation of phagosome biogenesis by Legionella pneumophila creates an organelle permissive for intracellular growth.* Nature Cell Biol, 1999. **1**(7): p. 451-453.
203. Kagan, J.C. and C.R. Roy, *Legionella phagosomes intercept vesicular traffic from endoplasmic reticulum exit sites.* Nat Cell Biol, 2002. **4**(12): p. 945-54.
204. Steiner, B., et al., *ER remodeling by the large GTPase atlastin promotes vacuolar growth of Legionella pneumophila.* EMBO Rep, 2017.
205. Ensminger, A.W., *Legionella pneumophila, armed to the hilt: justifying the largest arsenal of effectors in the bacterial world.* Current Opinion in Microbiology, 2016. **29**(Supplement C): p. 74-80.
206. Schroeder, G.N., *The Toolbox for Uncovering the Functions of Legionella Dot/Icm Type IVb Secretion System Effectors: Current State and Future Directions.* Frontiers in Cellular and Infection Microbiology, 2018. **7**(528).
207. Manske, C. and H. Hilbi, *Metabolism of the vacuolar pathogen Legionella and implications for virulence.* Frontiers in Cellular and Infection Microbiology, 2014. **4**.
208. Qiu, J. and Z.-Q. Luo, *Legionella and Coxiella effectors: strength in diversity and activity.* Nat Rev Micro, 2017. **15**(10): p. 591-605.
209. Qiu, J. and Z.-Q. Luo, *Hijacking of the Host Ubiquitin Network by Legionella pneumophila.* Frontiers in Cellular and Infection Microbiology, 2017. **7**(487).
210. Best, A. and Y. Abu Kwaik, *Evolution of the Arsenal of Legionella pneumophila Effectors To Modulate Protist Hosts.* MBio, 2018. **9**(5).
211. Ivanov, S.S., et al., *Lipidation by the host prenyltransferase machinery facilitates membrane localization of Legionella pneumophila effector proteins.* J Biol Chem, 2010. **285**(45): p. 34686-98.
212. Kubori, T., A. Hyakutake, and H. Nagai, *Legionella translocates an E3 ubiquitin ligase that has multiple U-boxes with distinct functions.* Mol Microbiol, 2008. **67**(6): p. 1307-19.
213. Zamboni, D.S., et al., *The Birc1e cytosolic pattern-recognition receptor contributes to the detection and control of Legionella pneumophila infection.* Nat Immunol, 2006. **7**(3): p. 318-25.
214. Binz, H.K., et al., *Crystal structure of a consensus-designed ankyrin repeat protein: implications for stability.* Proteins, 2006. **65**(2): p. 280-4.
215. Bork, P., *Hundreds of ankyrin-like repeats in functionally diverse proteins: mobile modules that cross phyla horizontally?* Proteins, 1993. **17**(4): p. 363-74.
216. Kohl, A., et al., *Designed to be stable: crystal structure of a consensus ankyrin repeat protein.* Proc Natl Acad Sci U S A, 2003. **100**(4): p. 1700-5.
217. Mosavi, L.K., D.L. Minor, Jr., and Z.Y. Peng, *Consensus-derived structural determinants of the ankyrin repeat motif.* Proc Natl Acad Sci U S A, 2002. **99**(25): p. 16029-34.
218. Yu, H., et al., *Molecular dynamics study of the stabilities of consensus designed ankyrin repeat proteins.* Proteins, 2006. **65**(2): p. 285-95.
219. Sedgwick, S.G. and S.J. Smerdon, *The ankyrin repeat: a diversity of interactions on a common structural framework.* Trends Biochem Sci, 1999. **24**(8): p. 311-6.

220. Islam, Z., et al., *New paradigm in ankyrin repeats: Beyond protein-protein interaction module.* Int J Biol Macromol, 2018. **109**: p. 1164-1173.
221. Al-Khodor, S., et al., *Functional diversity of ankyrin repeats in microbial proteins.* Trends Microbiol, 2010. **18**(3): p. 132-9.
222. Belyi, Y., et al., *Lgt: a family of cytotoxic glucosyltransferases produced by Legionella pneumophila.* J Bacteriol, 2008. **190**(8): p. 3026-35.
223. Shen, X., et al., *Targeting eEF1A by a Legionella pneumophila effector leads to inhibition of protein synthesis and induction of host stress response.* Cell Microbiol, 2009. **11**(6): p. 911-26.
224. Li, T., et al., *SET-domain bacterial effectors target heterochromatin protein 1 to activate host rDNA transcription.* EMBO Rep, 2013. **14**(8): p. 733-40.
225. Rolando, M., et al., *Legionella pneumophila effector RomA uniquely modifies host chromatin to repress gene expression and promote intracellular bacterial replication.* Cell Host Microbe, 2013. **13**(4): p. 395-405.
226. Lee, P.C. and M.P. Machner, *The Legionella Effector Kinase LegK7 Hijacks the Host Hippo Pathway to Promote Infection.* Cell Host Microbe, 2018. **24**(3): p. 429-438.e6.
227. Price, D.H., *P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II.* Mol Cell Biol, 2000. **20**(8): p. 2629-34.
228. Zhou, Q., T. Li, and D.H. Price, *RNA polymerase II elongation control.* Annu Rev Biochem, 2012. **81**: p. 119-43.
229. Uchikawa, E., et al., *Structural insight into the mechanism of stabilization of the 7SK small nuclear RNA by LARP7.* Nucleic Acids Res, 2015. **43**(6): p. 3373-88.
230. Barboric, M., et al., *7SK snRNP/P-TEFb couples transcription elongation with alternative splicing and is essential for vertebrate development.* Proc Natl Acad Sci U S A, 2009. **106**(19): p. 7798-803.
231. He, N., et al., *A La-related protein modulates 7SK snRNP integrity to suppress P-TEFb-dependent transcriptional elongation and tumorigenesis.* Mol Cell, 2008. **29**(5): p. 588-99.
232. Muniz, L., S. Egloff, and T. Kiss, *RNA elements directing in vivo assembly of the 7SK/MePCE/Larp7 transcriptional regulatory snRNP.* Nucleic Acids Res, 2013. **41**(8): p. 4686-98.
233. Xue, Y., et al., *A capping-independent function of MePCE in stabilizing 7SK snRNA and facilitating the assembly of 7SK snRNP.* Nucleic Acids Res, 2010. **38**(2): p. 360-9.
234. Guo, J. and D.H. Price, *RNA polymerase II transcription elongation control.* Chem Rev, 2013. **113**(11): p. 8583-603.
235. Jeronimo, C., et al., *Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme.* Mol Cell, 2007. **27**(2): p. 262-74.
236. Michels, A.A., et al., *Binding of the 7SK snRNA turns the HEXIM1 protein into a P-TEFb (CDK9/cyclin T) inhibitor.* Embo j, 2004. **23**(13): p. 2608-19.

237. Yik, J.H., et al., *Inhibition of P-TEFb (CDK9/Cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA.* Mol Cell, 2003. **12**(4): p. 971-82.
238. Nguyen, V.T., et al., *7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes.* Nature, 2001. **414**(6861): p. 322-5.
239. Yang, Z., et al., *The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription.* Nature, 2001. **414**(6861): p. 317-22.
240. Perumal, K., et al., *Purification, characterization, and cloning of the cDNA of human signal recognition particle RNA 3'-adenylating enzyme.* J Biol Chem, 2001. **276**(24): p. 21791-6.
241. Sinha, K.M., et al., *Adenylation of small RNAs in human cells. Development of a cell-free system for accurate adenylation on the 3'-end of human signal recognition particle RNA.* J Biol Chem, 1998. **273**(12): p. 6853-9.
242. Peterlin, B.M. and D.H. Price, *Controlling the elongation phase of transcription with P-TEFb.* Mol Cell, 2006. **23**(3): p. 297-305.
243. Krueger, B.J., et al., *LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated.* Nucleic Acids Res, 2008. **36**(7): p. 2219-29.
244. Holm, L. and P. Rosenstrom, *Dali server: conservation mapping in 3D.* Nucleic Acids Res, 2010. **38**(Web Server issue): p. W545-9.
245. Abendroth, U., et al., *Identification of new protein-coding genes with a potential role in the virulence of the plant pathogen Xanthomonas euvesicatoria.* BMC Genomics, 2017. **18**(1): p. 625.
246. Chosed, R., et al., *Structural analysis of Xanthomonas XopD provides insights into substrate specificity of ubiquitin-like protein proteases.* J Biol Chem, 2007. **282**(9): p. 6773-82.
247. Pruneda, J.N., et al., *The Molecular Basis for Ubiquitin and Ubiquitin-like Specificities in Bacterial Effector Proteases.* Mol Cell, 2016. **63**(2): p. 261-276.
248. Yang, A., S. Pantoom, and Y.W. Wu, *Elucidation of the anti-autophagy mechanism of the Legionella effector RavZ using semisynthetic LC3 proteins.* Elife, 2017. **6**.
249. Horenkamp, F.A., et al., *The Legionella Anti-autophagy Effector RavZ Targets the Autophagosome via PI3P- and Curvature-Sensing Motifs.* Dev Cell, 2015. **34**(5): p. 569-76.
250. Finn, R.D., et al., *InterPro in 2017-beyond protein family and domain annotations.* Nucleic Acids Res, 2017. **45**(D1): p. D190-d199.
251. Rawlings, N.D., A.J. Barrett, and R. Finn, *Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors.* Nucleic Acids Res, 2016. **44**(D1): p. D343-50.
252. Rawlings, N.D., et al., *The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database.* Nucleic Acids Res, 2018. **46**(D1): p. D624-d632.
253. Stone, B.J. and Y. Abu Kwaik, *Expression of multiple pili by Legionella pneumophila: Identification and characterization of a type IV pilin gene and its*

role in adherence to mammalian and protozoan cells Infect.Immun., 1998. **66**: p. 1768-1775.
254. Bruckert, W.M. and Y. Abu Kwaik, *Lysine11-Linked Polyubiquitination of the AnkB F-Box Effector of Legionella pneumophila.* Infect Immun, 2015. **84**(1): p. 99-107.
255. Wilkins, M.R., et al., *Protein identification and analysis tools in the ExPASy server.* Methods Mol Biol, 1999. **112**: p. 531-52.
256. Minor, W., et al., *HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes.* Acta Crystallogr D Biol Crystallogr, 2006. **62**(Pt 8): p. 859-66.
257. Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.
258. Chen, V.B., et al., *MolProbity: all-atom structure validation for macromolecular crystallography.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 1): p. 12-21.
259. Qiu, J. and Z.Q. Luo, *Legionella and Coxiella effectors: strength in diversity and activity.* Nat Rev Microbiol, 2017. **15**(10): p. 591-605.
260. Du Toit, A., *The effector repertoire of Legionella.* Nat Rev Microbiol, 2019. **17**(3): p. 126.
261. Altschul, S.F., et al., *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Res, 1997. **25**(17): p. 3389-402.
262. Lando, D., et al., *Asparagine hydroxylation of the HIF transactivation domain a hypoxic switch.* Science, 2002. **295**(5556): p. 858-61.
263. Mahon, P.C., K. Hirota, and G.L. Semenza, *FIH-1: a novel protein that interacts with HIF-1α and VHL to mediate repression of HIF-1 transcriptional activity.* Genes & Development, 2001. **15**(20): p. 2675-2686.
264. Hewitson, K.S., et al., *Hypoxia-inducible factor (HIF) asparagine hydroxylase is identical to factor inhibiting HIF (FIH) and is related to the cupin structural family.* J Biol Chem, 2002. **277**(29): p. 26351-5.
265. Yang, M., et al., *Asparagine and aspartate hydroxylation of the cytoskeletal ankyrin family is catalyzed by factor-inhibiting hypoxia-inducible factor.* J Biol Chem, 2011. **286**(9): p. 7648-60.
266. Coleman, M.L., et al., *Asparaginyl hydroxylation of the Notch ankyrin repeat domain by factor inhibiting hypoxia-inducible factor.* J Biol Chem, 2007. **282**(33): p. 24027-38.

CURRICULUM VITAE

Juanita E. Von Dwingelo, M.S.
233 Haas Ln
Sellersburg, IN 47172
jevond01@louisville.edu

EDUCATION

08/2005-05/2010	B.S. in Biology, Murray State University, Murray, KY
08/2010-05/2012	M.S. in Biology, Murray State University, Murray, KY Thesis Title: Abundance of *Methanosaeta concillii*-like species in the sediment of Ledbetter Creek Embayment of Kentucky Lake
08/2012-12/2014 Louisville, KY	M.S. in Microbiology and Immunology, University of Louisville,
08/2012-Present Louisville, KY	Ph.D. in Microbiology and Immunology, University of Louisville, Dissertation Title: "Role of the AnkH effector of *Legionella pneumophila* in intracellular proliferation" Principal Investigator: Yousef Abu Kwaik Anticipated graduation August 2019

ACADEMIC APPOINTMENTS

08/2010-05/2012	Graduate Teaching Assistant, Department of Biology, Murray State University, Murray, KY
08/2012-Present	Graduate Research Assistant, Department of Microbiology and Immunology, University of Louisville, Louisville, KY

OTHER POSITIONS AND EMPLOYMENT

06/2003-08/2006	Cashier and Shift Leader, Dairy Queen, Marion, KY
03/2007-12/2012	Server, Patti's 1880s Settlement Grand Rivers, KY

PROFESSIONAL MEMBERSHIPS AND ACTIVITIES

2008-2012	Member of Tri Beta, Biological Honor Society

2010-Present	Member of the Kentucky Academy of Science
2014-Present	Member of the American Society of Microbiology

HONORS AND AWARDS

11/2011	Kentucky Academy of Science Graduate Research Competition, Microbiology, First Place
08/2012-Present	Integrated Programs in Biomedical Sciences Fellow, University of Louisville, Louisville, KY

COMMITTEE ASSIGNMENTS AND ADMINISTRATIVE SERVICES

05/2010-05/2011	Treasurer of Tri Beta, Murray State University, Murray, KY
05/2009-05/2010	President of Tri Beta, Murray State University, Murray, KY
05/2013-05/2014	Treasurer for Microbiology and Immunology Student Organization (MISO), University of Louisville, Louisville, KY
01/2015-05/2016	President and Co-founder of Science Policy and Outreach Group, University of Louisville, Louisville, KY

EDUCATIONAL ACTIVITIES

LECTURES/PRIMARY INSTRUCTION

08/2010-05/2012	Laboratory Instructor for Freshman Biology, Department of Biology, Murray State University, Murray, KY
08/2010-05/2012	Laboratory Instructor, Microbiology, Department of Biology, Murray State University, Murray, KY
2014-2015	Graduate Teaching Academy, University of Louisville, Louisville, KY

OUTREACH

01/2015-Present	Kentucky Science Center Volunteer
03/2015	Louisville Regional Science and Engineering Fair Mentor, Louisville, KY
03/2015	Career Fair volunteer, Stuart Middle School, Louisville, KY
03/2015	Louisville Regional Science and Engineering Fair Judge, Louisville, KY
10/2015	PULSE day at the Kentucky Science Center, Volunteer leader and speaker, Louisville, KY
12/2015 - 03/2016	Student Mentoring for Louisville Science Fair Students, Louisville, KY
03/2016	Louisville Regional Science and Engineering Fair Mentor, Louisville, KY

12/2016 – 03/2017	Student Mentoring for Louisville Science Fair Students, Louisville, KY
03/2017	Louisville Regional Science and Engineering Fair Judge, Louisville, KY
03/2018	Louisville Regional Science and Engineering Fair Judge, Louisville, KY

ABSTRACTS AND PRESENTATIONS

ORAL PRESENTATIONS

Local/Regional

1. **Von Dwingelo JE** (2011) Abundance of *Methanosaeta concillii*-like species in the sediment of Ledbetter Creek Embayment of Kentucky Lake. Kentucky Academy of Science Annual Meeting Seminar, Murray State University, Murray, KY

2. **Von Dwingelo JE** (2014) Role of the AnkH Effector of *Legionella Pneumophila* in Intracellular Proliferation. Department of Microbiology and Immunology Seminar, University of Louisville, Louisville, KY

3. **Von Dwingelo JE** (2015) Role of the AnkH Effector of *Legionella Pneumophila* in Intracellular Proliferation. Department of Microbiology and Immunology Seminar, University of Louisville, Louisville, KY

4. **Von Dwingelo JE** (2016) Role of the AnkH Effector of *Legionella Pneumophila* in Intracellular Proliferation. Department of Microbiology and Immunology Seminar, University of Louisville, Louisville, KY

5. **Von Dwingelo JE** (2017) Role of the AnkH Effector of *Legionella Pneumophila* in Intracellular Proliferation. Department of Microbiology and Immunology Seminar, University of Louisville, Louisville, KY

POSTERS

Local/Regional

1. **Von Dwingelo JE,** Price CTD, Jones S, Chung I, Cygler M, Abu Kwaik Y (2017) Role of the AnkH Effector of *Legionella Pneumophila* in Intracellular Proliferation. Research! Louisville. University of Louisville, Louisville, KY.

ABSTRACTS

Local/Regional

1. **Von Dwingelo JE**, Johnston T (2011) Abundance of *Methanosaeta concillii*-like species in the sediment of Ledbetter Creek Embayment of Kentucky Lake. Kentucky Academy of Science Annual Meeting. Murray State University, Murray, KY.

2. **Von Dwingelo JE,** Price CTD, Jones S, Chung I, Cygler M, Abu Kwaik Y (2017) Role of the AnkH Effector of *Legionella Pneumophila* in Intracellular Proliferation. Research! Louisville. University of Louisville, Louisville, KY.

PUBLICATIONS

PEER-REVIEWED

1. #Richards AM, #**Von Dwingelo JE**, Price CTD, Abu Kwaik Y. Cellular Microbiology and Molecular Ecology of Legionella-amoeba Interaction. Virulence 2013;4(4):301-14. doi: 10.4161/viru.24290

   # denotes co-first authorship

2. Price CTD, Richards AM, **Von Dwingelo JE**, Samara HA, Abu Kwaik Y. Amoeba host-Legionella Synchronization of Amino Acid Auxotrophy and its Role in Bacterial Adaptation and Pathogenic Evolution. Environ Microbiol. 2014;16(2):350-8. doi: 10.1111/1462-2920.12290

3. Price CTD, Merchant M, Jones S, Best AM, **Von Dwingelo JE**, Lawrenz MB, Alam N, Schueler-Furman O, Abu Kwaik Y. Host FIH-Mediated Asparaginyl Hydroxylation of Translocated Legionella pneumophila Effectors. Front Cell Infect Microbiol. 2017;7:54. doi: 10.3389/fcimb.2017.00054

PEER-REVIEWED- In Press

1. **Von Dwingelo JE**, Chung I, Price CTD, Li L, Jones S, Cygler M, Abu Kwaik Y. Nuclear Trafficking of Core Effector of *Legionella pneumophila* Required for Intracellular Replication. mBio, 2019 (In press)

TRAININGS

Biosafety Training, Radiation Safety, Formaldehyde Training, Recombinant DNA Guidelines, HIPPA (Health Insurance Portability and Accountability Act), Blood-Borne Pathogen Training, Biohazard and Laboratory Safety Training