



**HAL**  
open science

# New insights about LegK4 from *Legionella pneumophila*

Quentin Blache

► **To cite this version:**

Quentin Blache. New insights about LegK4 from *Legionella pneumophila*. Molecular biology. Université de Lyon, 2022. English. NNT : 2022LYSE1078 . tel-04789815

**HAL Id: tel-04789815**

**<https://theses.hal.science/tel-04789815v1>**

Submitted on 19 Nov 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N°d'ordre NNT : 2022LYSE1078

**THESE de DOCTORAT DE L'UNIVERSITE DE LYON**  
opérée au sein de  
**l'Université Claude Bernard Lyon 1**

**Ecole Doctorale N° accréditation**  
**Evolution, Ecosystèmes, Microbiologie, Modélisation ED341**

**Spécialité de doctorat** : Infectiologie fondamentale et Biologie  
moléculaire

Soutenue publiquement le 22/06/2022, par :  
**Quentin Blache**

---

**New insights about LegK4 from**  
***Legionella pneumophila***

---

Devant le jury composé de :

Bruel Christophe	Professeur	Université de Lyon	Président du jury
Molmeret Maëlle	Professeure	Université de Toulon	Rapporteuse
Samba Louaka Ascel	Maître de Conférences	Université de Poitiers	Rapporteur
Poussereau Nathalie	Maître de Conférences	Université de Lyon	Directrice de thèse



N°d'ordre NNT : 2022LYSE1078

**THESE de DOCTORAT DE L'UNIVERSITE DE LYON**  
opérée au sein de  
**l'Université Claude Bernard Lyon 1**

**Ecole Doctorale N° accréditation**  
**Evolution, Ecosystèmes, Microbiologie, Modélisation ED341**

**Spécialité de doctorat** : Infectiologie fondamentale et Biologie  
moléculaire

Soutenue publiquement le 22/06/2022, par :

**Quentin Blache**

INSERM U1111, CNRS, UM5308, ENS, CIRI

---

**New insights about LegK4 from**  
***Legionella pneumophila***

---

Devant le jury composé de :

Bruel Christophe	Professeur	Université de Lyon	Président du jury
Molmeret Maëlle	Professeure	Université de Toulon	Rapporteuse
Samba Louaka Ascel	Maître de Conférences	Université de Poitiers	Rapporteur
Poussereau Nathalie	Maître de Conférences	Université de Lyon	Directrice de thèse

Copyright ©  
Quentin BLACHE  
2022

## ACKNOWLEDGMENTS

Je tiens premièrement à remercier les membres du jury pour avoir accepté la lourde tâche de juger mes travaux de thèse.

Un immense merci à Nathalie Poussereau pour avoir accepté de m'aider, d'avoir lu et relu cette thèse. Merci pour votre aide inestimable, vos conseils pertinents ainsi que votre bienveillance. Tout cela n'aurait jamais été possible sans vous.

À Bénédicte qui a su toujours trouver les mots pour réconforter lors des différentes crises. Tu es une grande dame (ça fait un peu éloge funèbre...) et un modèle à suivre. J'admire avec quelle facilité tu gères ton double emploi d'enseignante/chercheuse et de maman. Tu es un exemple à suivre pour tous !

Je tiens à remercier toutes ma famille qui malgré mes doutes, mes craintes à toujours cherché à me pousser à finir cette thèse.

Papa, tu es un pilier sur lequel je peux compter et n'oublies pas que tu peux toujours compter sur moi. I love you to the moon and back.

Elodie, ma marâtre préférée, la distance rend notre relation plus compliquée mais merci à toi d'avoir été un modèle en termes de gentillesse et de bienveillance. Je te

souhaite le meilleur.

Neo, tu mérites d'apparaître ici. Merci d'avoir été là depuis le début de cette thèse, du passage du concours à la fin de la thèse. Tu as vu que la chenille est finalement devenue un papillon VOILÀÀÀÀÀÀÀÀ Bon, Paris et les dimanches brunch/anime n'attendent que toi, donc BOUGE TOI.

Patrick, première personne que j'ai rencontré en arrivant à Paris et quelle rencontre ! Je pense qu'on tombera d'accord sur le fait que cette rencontre devait se faire. Finalement, en quelques mois grâce à ton soutien, tes conseils et ton écoute j'ai grandi. Je pense que tout le monde devrait avoir un Patrick dans sa vie. Sache que cette thèse n'aurait jamais été fini sans toi et nos séances (intensive) de coworking #HUBSY. Je souhaite, de tout cœur, que notre amitié grandisse encore et encore. Quand le jour viendra pour toi de faire ta thèse, saches que tu pourras compter sur moi.

Un immense merci à Amélie, par votre travail, votre patience, ensemble nous avons réussi à dépasser les objectifs que nous nous étions fixés.

Un grand merci à Davide pour ta présence (le ramen), avoir trouver une solution concernant la perte totale de ma bibliographie et pour m'avoir appris à écrire sous LaTeX.

José, mon gym buddy préféré, merci d'être là tous les matins avec moi. Désolé d'être (trop) autoritaire et relou avec toi concernant les temps de repos (tu sais que j'ai raison).

Anthony, tu as bien évidemment ta place ici. Tu n'imagines vraiment pas l'aide et

le soutien que m'a apporté. Tu as dû lire des pages et des pages de texte qui ont dû te sembler être du chinois au début mais maintenant tu es au niveau en bio. Un immense merci à toi qui a été là à des moments vraiment critiques des larmes à la joie de pouvoir déposer ce manuscrit.

Pendant cette thèse, j'ai finalement perdu beaucoup de personne, des amitiés ou des personnes qui comptaient pour moi, je tenais malgré tout à vous remercier.

Enfin, je remercie sincèrement toutes les personnes qui m'ont fait douter de moi, de mes capacités. Vous avez été les moteurs qui m'ont poussé à me dépasser et à chercher le meilleur en moi.





# Contents

<b>ACKNOWLEDGMENTS</b> . . . . .	ii
<b>TABLE OF CONTENTS</b> . . . . .	vi
<b>ABSTRACT IN FRENCH</b> . . . . .	xi
<b>ABSTRACT</b> . . . . .	xiii
<b>ACRONYMS</b> . . . . .	xxiv
<b>List of Tables</b>	xxiv
<b>List of Figures</b>	xxv
<b>I Background</b>	1
<b>1 <i>Legionella pneumophila</i></b>	2
<b>1.1 Legionnaires' disease history and symptoms</b> . . . . .	2
<b>1.2 The causative agent of Legionnaires' disease</b> . . . . .	3
<b>1.3 Human infection, disease diagnosis and treatment</b> . . . . .	5
<b>1.4 <i>Legionella</i> life cycle in amoeba and macrophages</b> . . . . .	9
<b>1.4.1 Adhesion to eukaryotic cell</b> . . . . .	9
<b>1.4.2 Entry of <i>Legionella</i></b> . . . . .	11
<b>1.4.3 Intracellular lifestyle and phase transition</b> . . . . .	12
<b>1.5 Genome and <i>Legionella</i> strains</b> . . . . .	13

<b>2 Type IV Secretion System Dot/Icm and Effectors</b>	<b>16</b>
2.1 Structure of Dot/Icm T4SS	17
2.1.1 Core-transmembrane subcomplex	18
2.1.2 Coupling protein subcomplex	19
2.1.3 Cytoplasmic components	21
2.1.4 Proteins with unknown function	21
2.2 The 300 Dot/Icm effectors	22
2.2.1 Endosomal-lysosomal pathway escape	23
2.2.2 Nutrition and energy	35
2.2.3 Vesicular trafficking modulation	38
2.2.4 Autophagy	43
2.2.5 Apoptosis, cell cycle and death	44
2.2.6 Host gene expression control	50
2.2.7 Metaeffectors	54
<b>3 The endoplasmic reticulum and <i>L. pneumophila</i></b>	<b>58</b>
3.1 The endoplasmic reticulum	58
3.2 HSP70 family chaperones	60
3.2.1 Hsc70	62
3.2.2 Hsp70	63
3.2.3 Grp75	63
3.2.4 BiP	64
3.3 Unfolded Protein Response	65
3.3.1 UPR signaling pathways	65
3.3.2 <i>Legionella</i> and UPR	69

<b>4</b>	<b>New Target of <i>Legionella</i>: The Nucleus</b>	<b>72</b>
4.1	Levels of epigenetic control	72
4.1.1	DNA methylation	73
4.1.2	Non-coding RNA	75
4.1.3	Histone modifications	75
4.1.4	Histone H3	76
4.2	The new era of nucleomodulins	79
4.2.1	Protein nuclear import	79
4.2.2	Nucleomodulins	85
<b>II</b>	<b>New insights about LegK4 Paris from <i>L. pneumophila</i></b>	<b>89</b>
1	Localization of LegK4	92
1.1	LegK4 Paris localized in the nucleus of transfected HeLa cells	92
1.2	LegK4 Philadelphia localized into the cytosol of transfected HeLa cells	94
1.3	LegK4 Philadelphia is wrongly annotated in databases	96
1.4	Determination of the nuclear localization signal (NLS) of LegK4 Paris	100
1.5	Migration of LegK4 Paris to the nucleus	104
1.6	Translocation of LegK4 Paris into infected cells	106
1.7	Localization of LegK4 Paris during the infection of human cells	107
2	LegK4 interactant and substrate	113
2.1	Yeast two-hybrid screening	113
2.2	Validation of the interactants	115
2.3	<i>In vitro</i> validation of LegK4 substrates	118
2.4	Validation of LegK4 substrates in transfected cells	123

3	Role of LegK4 Paris	128
3.1	LegK4 does not have an impact on <i>Legionella</i> intracellular repli-	
	cation	128
3.2	Impact of LegK4 on human cells	132
<b>III Discussion and Conclusion</b>		<b>147</b>
<b>Annexes</b>		<b>163</b>
<b>References</b>		<b>207</b>



## ABSTRACT

*Legionella pneumophila* est la bactérie responsable de la légionellose, une pneumonie grave chez l'Homme. C'est un problème de santé publique en raison d'un taux de létalité d'environ 10%, même lorsque le patient est pris en charge à temps avec le bon traitement antibiotique. Les souches pathogènes de *Legionella* émergent de l'environnement après leur multiplication intracellulaire dans leurs hôtes environnementaux, comme les amibes. Les légionelles peuvent être disséminées par de fines particules d'eau en suspension appelées aérosols, qui sont formés par des technologies humaines telles que les fontaines ou les tours aérorefrigérantes. Lorsqu'ils sont inhalés, ces aérosols permettent aux bactéries d'atteindre les voies respiratoires basses, où des cellules immunitaires comme des macrophages alvéolaires sont présentes. Pour survivre alors à la phagocytose, un système de sécrétion Dot/Icm de type 4 (SST4) et la sécrétion d'environ 300 protéines, appelés effecteurs, sont absolument nécessaires.

Cette thèse, menée au sein de l'équipe LegioPath (INSERM U1111-CNRS UM5308), a permis de contribuer à une meilleure compréhension d'un effecteur de *Legionella* appelé LegK4, en combinant différentes approches expérimentales. Dans un premier temps, nous avons identifié sa localisation par microscopie confocale via l'expression ectopique de LegK4 associé à la GFP. Nous nous sommes ensuite focalisés en deuxième partie sur l'identification des interactants et des substrats de LegK4 par co-immunoprécipitation, création d'une lignée cellulaire stable exprimant LegK4, western blot et spectrométrie de masse. Enfin, la troisième partie concerne le rôle de LegK4 ainsi que son impact sur

les cellules infectées. Nous avons cherché à identifier un défaut de réplication en mettant en concurrence des souches de *Legionella* fluorescentes possédant ou non le gène *legK4*. Nous avons ensuite vérifié si l'activité de phosphorylation de LegK4 pouvait avoir un impact sur le profil d'expression génique de l'hôte.

L'introduction est composée de quatre chapitres qui définissent le contexte de ce travail. Le premier chapitre donne des renseignements importants sur la légionellose, ainsi que le modèle de notre étude *Legionella pneumophila*. Le deuxième chapitre est une description des principaux facteurs de virulence de *L. pneumophila*, et plus particulièrement, le rôle des effecteurs de système de sécrétion de type IV. Un troisième chapitre est dédié à LegK4, ayant été auparavant caractérisé comme impactant la voie de résolution du stress du réticulum (UPR), en interagissant avec les protéines chaperonne de la famille HSP70. Enfin, le dernier chapitre concerne le noyau, le contrôle épigénétique ainsi que les nucléomodulines chez *L. pneumophila*, étant donné que ce travail a démontré que LegK4 pourrait être nucléaire et moduler la transcription du gène hôte.

## ABSTRACT

*Legionella pneumophila* is the bacteria responsible for legionellosis, a severe pneumonia in humans. It is a public health issue due to a 10% mortality rate, even with the proper antibiotic therapy. Pathogenic strains of *Legionella* emerge from the environment after intracellular multiplication in amoebae. Bacteria can be disseminated by small water particles called aerosols formed by human technologies, such as ornamental fountains. When inhaled, these aerosols allow bacteria to reach the lower respiratory tract, where alveolar macrophages are present. To survive phagocytosis, a functional Dot/Icm type 4 secretion system (T4SS) and the secretion of approximately 300 proteins, called effectors, are absolutely required.

The objective of this thesis, conducted in the LegioPath team (INSERM U1111-CNRS UM5308) is to contribute to a better understanding of a *Legionella* effector called LegK4 by combining different experimental approaches. In the first step, we aimed mainly at identifying its localization using ectopic expression of GFP-tagged LegK4 and confocal microscopy. The second part focuses on the identification of LegK4 interactants and substrates using a co-immunoprecipitation assay, creation of a stable cell-line expressing LegK4, western blot and mass spectrometry. Finally, the third part concerns the LegK4 role and its impact on the infected cells. We aimed to find a different phenotype using competition between fluorescent *Legionella* with or without *legK4* gene. Then, we checked if the phosphorylation activity of LegK4 could impact the host gene transcription.



The introduction is composed of four chapters that set the context of this work. The first chapter gives important information about *Legionella* and legionellosis. The second chapter is a review of the major virulence factors of *L. pneumophila*, and more particularly, the role of type IV secretion system effectors. Because LegK4 was previously characterized as interacting with the endoplasmic reticulum chaperone from the HSP70 family and impacting the unfolded protein response (UPR), a third chapter is dedicated to these points. Finally, the last chapter concerns the nucleus, epigenetic control and nucleomodulin because this work demonstrated that LegK4 could be nuclear and modulate host gene transcription.

# List of Acronyms

<i>asd</i>	aspartate $\beta$ semialdehyde dehydrogenase
AD	Activation domain
ADP	Adenosine diphosphate
ATF4	Activating transcription factor 4
ATF6 $\alpha$	Activating transcription factor 6 $\alpha$
ATP	Adenosine triphosphate
ARF	ADP-ribosylation factor
AROS	Active regulator of SIRT1
ASK1	Apoptosis signal regulating kinase 1
AYE	ACES buffered yeast extract
BiP	Binding immunoglobulin protein
Ceg	Coregulated with Effector Genes
CGI	CpG island
CpG	Cytosine-guanine dinucleotide
CHOP	C/EBP homologous Protein
CNRL	Centre national de reference pour legionelles
CoA	Acetyl coenzyme A
Co-IP	Co-immunoprecipitation

CLCK1	Cdc2-like kinase
CUL4	Cullin-4
CYE	Charcoal yeast extract agar
DAG	Diacylglycerol
DAPI	4',6-diamidino-2-phenylindole, dihydrochloride
DAF	Decay acceleration factor
DBD	DNA binding domain (DBD)
DFC	Dense fibrillar component
DNA	Deoxyribonucleic acid
DDR	DNA damage repair
DMEM	Dulbecco's modified eagle medium
<i>dot</i>	Defective in organelle trafficking/intracellular replication
DTT	1,4-dithiothreitol, threo-1,4-Dimercapto-2,3-butanediol
EDTA	Ethylenediaminetetraacetic acid
EFA1	Elongation factor 1-Alpha-1
EIF2 $\alpha$	Eukaryotic initiation factor 2 $\alpha$
EGF	Epidermal growth factor
eGFP	emerald green fluorescent protein
ELP	Eukaryotic-like proteins
EM	Eukaryotic motif
EPEC	Enteropathogenic <i>E. coli</i>
ER	Endoplasmic reticulum
ERAD	Endoplasmic reticulum-Associated Degradation
ERdjs	Endoplasmic reticulum-Localized DnaJ family members
ERK	Extracellular signal-Regulated Kinase
ERQC	Endoplasmic reticulum Quality Control

ERS	Endoplasmic reticulum Stress
ETR	Effector triggered response
FACS	Fluorescence activated cell sorting
FC	Fibrillar center
FIB1	Fibrillarin
<i>flaA</i>	flagellin
FRET	Fluorescence resonance energy transfer
GANC	Glucosidase $\alpha$
GAP	GTPase-activating protein
GC	Guanine-Cytosine
GC	Granular component
GEF	Guanine nucleotide exchange factors
GDP	Guanosine-diphosphate
GDI	Guanosine dissociation inhibitor
GDF	Guanosine dissociation factor
GPCR	G protein-coupled receptor
GST	Glutathione S-transferase
GTP	Guanosine-triphosphate
GMP	Guanosine-monophosphate
GNAT	GCN5-related N-Acetyltransferases
GobX	Golgi localized U-box E3 ligase
GRV	Groundnut rosette virus
HA	Hemagglutinin
HAT	Histone acetyltransferase
HCl	Hydrochloric acid
HDAC	Histone deacetylase

Hep-2	Human epithelial cell type 2
HGT	Horizontal gene transfer
HHEX	Hematopoietically expressed homeobox
HIV	Human immunodeficiency virus
HPTM	Histone post-translational modification
HR	Homologous Recombination
<i>icm</i>	Intracellular multiplication
IKK	I $\kappa$ B kinase
IMPA	Inositol monophosphatase
IPTG	isopropyl $\beta$ -D-1-thiogalactopyranoside
IRE1 $\alpha$	Inositol-requiring protein
IRES	Internal ribosome entry site
JNK	c-Jun N-terminal kinase
KAP	Karyopherin
kDa	Kilo Dalton
LB	Luria-Bertani
Lcl	<i>Legionella</i> collagen-like protein
LCV	<i>Legionella</i> -containing vacuole
Lec	Lethal effect on yeast growth
LegK	<i>Legionella</i> eukaryotic-like gene Kinase
LegS	<i>Legionella</i> S1P-Lyase
Lem	<i>Legionella</i> Effector identified by Machine learning
Lep	<i>Legionella</i> effector protein
Lgt	<i>Legionella</i> glucosyltransferase
LLO	Listeriolysin
Lna	<i>Legionella</i> NF- $\kappa$ B activator

LncP	Legionella nucleotide carrier protein
LncRNA	long non-coding RNA
LpdA	Lipoamide dehydrogenase
Lpg	<i>L. pneumophila</i> Philadelphia-1 gene
LpnE	<i>L. pneumophila</i> entry
LPS	Lipopolysaccharide
LRR	Leucine-Rich repeat
Lse	<i>L. pneumophila</i> SNARE effector
Ltp	<i>L. pneumophila</i> Translocated Protein
LTR	Long terminal repeats
Lub	<i>L. pneumophila</i> U-box protein
Lup	<i>L. pneumophila</i> ubiquitin-specific protease
Lvg	<i>Legionella</i> virulence gene
LYNUS	Lysosomal Nutrient Sensing
MAM	Mitochondria-associated ER membrane
MAPK	Mitogen-activated protein kinase
MavC	More regions allowing vacuole colocalization
Mb	Megabase
MBP	Methyl-binding protein
<i>mip</i>	Macrophage infectivity potentiator
mRNA	Messenger ribonucleic acid
miRNA	micro ribonucleic acid
MLST	Multilocus sequence typing
MOI	Multiplicity of infection
<i>momp</i>	Major outer membrane protein
mTORC1	Mammalian target of rapamycin

MTS	Mitochondrial Targeting Sequence
NBS	Nucleotide-Binding Site
ncRNA	non-coding RNA
<i>neuA</i>	N-acetylneuraminate cytidylyl transferase
NBD	Nucleotide-binding domain
NCBI	National center for biotechnology information database
NF- $\kappa$ B	Nuclear factor-kappa $\beta$
NHEJ	Non-homologous end joining
NoDS	Nucleolar detention sequence
NoLS	Nucleolar localization signal
NLR	NOD-Like Receptor
NLS	Nuclear localization signal
NPC	Nuclear pore complex
NRK Cells	Normal rat kidney epithelial cells
OCRL1	Oculocere-Brorenal syndrome of lowe 1
PA	Phosphatidic acid
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PE	Phosphatidylethanolamine
PERK	Protein kinase RNA-like endoplasmic reticulum kinase
PFKFB2	6-Phosphofructo-2-Kinase/Fructose-2,6-Biphosphatase 2
pH	Potential of hydrogen
PIK4	Phosphatidylinositol-4-kinase
<i>pilE</i>	pilin
PMSF	Phenylmethylsulfonyl fluoride
<i>proA</i>	zinc metalloprotease

PtdIns	Phosphatidylinositol
PtdIns(3)P	Phosphatidylinositol-3-Phosphate
PtdIns(4)P	Phosphatidylinositol-4-Phosphate
PtdIns(3,4) $P_2$	Phosphatidylinositol-3,4-diphosphate
PtdIns(3,5) $P_2$	Phosphatidylinositol-3,5-diphosphate
PtdIns(4,5) $P_2$	Phosphatidylinositol-4,5-diphosphate
PtdIns(3,4,5) $P_3$	Phosphatidylinositol-3,4,5-triphosphate
PTM	Post-Translational Modification
PYG	Peptone yeast glucose
qRT-PCR	Quantitative reverse transcription polymerase chain reaction
Rab	RAS (rat sarcoma) analog in the brain
Ral	Recruitment of ARF to the <i>Legionella</i> phagosome
RanBP1	RAN binding protein 1
Rav	Region allowing vacuole colocalization
RBBP4	Retinoblastoma binding Protein 4
RCC1	Regulator of chromosome condensation 1
RCE1	RAS converting CAAX endopeptidase
RILP	RAB-Interacting lysosomal protein
RIPA	Radioimmunoprecipitation assay buffer
RNA	Ribonucleic acid
RPMI	Roswell Park memorial institute
rRNA	ribosomic ribonucleic acid
RtxA	Repeat in toxin
Sac1	Suppressor of actin
SAM	S-Adenosine-methionine
SBT	Sequence-based typing



SCF	Skp, Cullin, F-box containing complex
SDS	Sodium dodecyl sulfate
SET	Su(var)3-9, enhancer-of-zeste and trithorax
SEM	scanning electron microscope
sfGFP	Susperfolder green fluorescent protein
Sg	Serogroup
Sid	Substrate of Icm/Dot transporter)
SLC25	Solute carrier family 25
SNARE	Soluble N-ethylmaleimide-sensitive factor attachment protein receptor)
SP-STP	<i>Streptococcus pyogenes</i> serine/threonine phosphatase
ST	Sequence type
STX	Syntaxin
SBD	Substrate-binding domain
SdcA	SidC paralogue
SdeA	Sde substrate of the Dot/Icm system
SdhA	Sdh succinate dehydrogenase
SDS	Sodium dodecyl Sulfate
SDS-PAGE	Sodium Dddecyl sulfate polyacrylamide gel electrophoresis
Sid	Substrate of Icm/Dot transporter
SusF	Suppressor of SidI
T2SS	Type 2 secretion system
T4CP	T4SS coupling protein
T4SS	Type 4 secretion system
TBC1d5	Tre-2/Bub2/Cdc16 domain family member 5
TBS	Tris buffered saline
TF	Transcription factor

TFF1	Trefoil factor 1
TGN	Trans-Golgi network
TOX4	TOX high mobility group box family member 4
TRAF2	Tumor necrosis factor receptor-associated factor 2
Tris	Trishy-droxymethylaminomethanex
TRL	Toll-Like Receptor
tRNA	Transfert ribonucleic Acid
TRIM21	Tripartite motif-containing protein
UPR	Unfolded protein response
V-ATPase	Vacuolar ATPases
Vip	VPS inhibitor Protein
VSV	Vesicular stomatitis virus
XBP1	X-box Binding Protein
Xist	X inactive specific transcript
Wip	WASP interacting protein
YPD	Yeast extract, peptone, Dextrose
5mC	5-Methyl-Cytosin

# List of Tables

1	List of characterized <i>Legionella pneumophila</i> effectors	163
2	List of human HSP70	170
3	List of characterized NoLS	172
4	<i>In cellula</i> LegK4 substrates	174
5	Modulation of UPR-related genes	175
6	Modulation of Glycolysis-related genes	178
7	Modulation of apoptosis-related genes	180
8	List of strains	185
9	List of plasmids	188
10	List of antibodies	189
11	List of primers	190

# List of Figures

1	<i>Legionella</i> and protozoa relationship	4
2	<i>Legionella</i> 's lifestyle in the environment (This study)	5
3	Infection by <i>Legionella</i> due to contaminated man-made devices (This study)	6
4	Evolution of reported legionellosis cases in France from 1988 to 2020 (from 'Santé Publique France')	8
5	<i>Legionella pneumophila</i> infectious cycle and LCV biogenesis (This study)	10
6	Architectural model and <i>Legionella</i> Dot/Icm T4BSS protein composition (This study)	18
7	LCV phosphatidylinositol composition modulation (This study)	29
8	Host small GTPase recruitment and modification by effectors (This study)	32
9	Impact of <i>Legionella</i> 's effectors on vesicle migration and LCV mobility (This study)	42
10	Apoptotic pathways (This study)	46
11	<i>Legionella</i> effector having an impact on autophagy or cell death (This study)	49
12	<i>Legionella</i> effectors involved in host gene expression (This study)	51
13	HSP70s folding cycle (This study)	61

14	Unfolded protein response (UPR) signaling pathway (This study)	67
15	Nuclear import mechanism and recycling of essential transport components (This study)	84
16	LegK4 localization depending on <i>Legionella</i> strain	93
17	Identity level of LegK4 between the different strains	95
18	LegK4 Philadelphia is wrongly annotated in database	98
19	LegK4 Paris NLS prediction made with NLStradamus	101
20	LegK4 Paris possesses an atypical NLS in its Cap domain	102
21	eGFPc and eGFP-LegK4 localization in yeast cells	105
22	Determination of LegK4 localization during macrophages infection	109
23	H3 and BiP interact with LegK4 in HeLa cells	117
24	H3 and BiP are host proteins targeted by LegK4	121
25	Intracellular growth of <i>L. pneumophila</i> in U937 macrophages	129
26	Intracellular growth of fluorescent <i>L. pneumophila</i> in U937 macrophages	131
27	Creation of a stable cell lineage	134
28	HA- <i>LegK4</i> <sub>D195N</sub> induction with doxycycline	137
29	Relationship between LegK4 and BiP	144
30	Comparison between LegK4 and MavR	156
31	Intracellular growth of <i>L. pneumophila</i> in U937 macrophages	157
32	Creation of scar free deletion in <i>Legionella</i> genome	195

Copyright ©  
Quentin BLACHE  
2022

# Part I

## Background

# Chapter 1

## *Legionella pneumophila*

### 1.1 Legionnaires' disease history and symptoms

The first outbreak of Legionnaires' disease occurred in July 1976 during the American Legion 58<sup>th</sup> annual convention, Philadelphia, Pennsylvania. Of the 182 infected patients, 29 died from a severe form of pneumonia, even though under antibiotic therapy [1]. A few months later, the outbreak agent was isolated and identified: it was caused by a bacterium that had never been described. The analysis led to naming this brand-new bacterial genus *Legionella* and the new species, *Legionella pneumophila* [2, 3].

After being exposed to the bacteria, this illness usually develops quickly within two to five days, but in some cases, illness can take longer, around two weeks. Legionnaires' disease frequently begins with headache, muscles aches and high fever, over 40°C. After two or three days, the patient rapidly develops a strong cough, with or without mucus and blood, chest pain, shortness of breath, gastrointestinal symptoms such as diarrhea and nausea, and confusion. Even though this disease primarily affects the lungs, it can, on occasion, cause infection in other parts of the body.



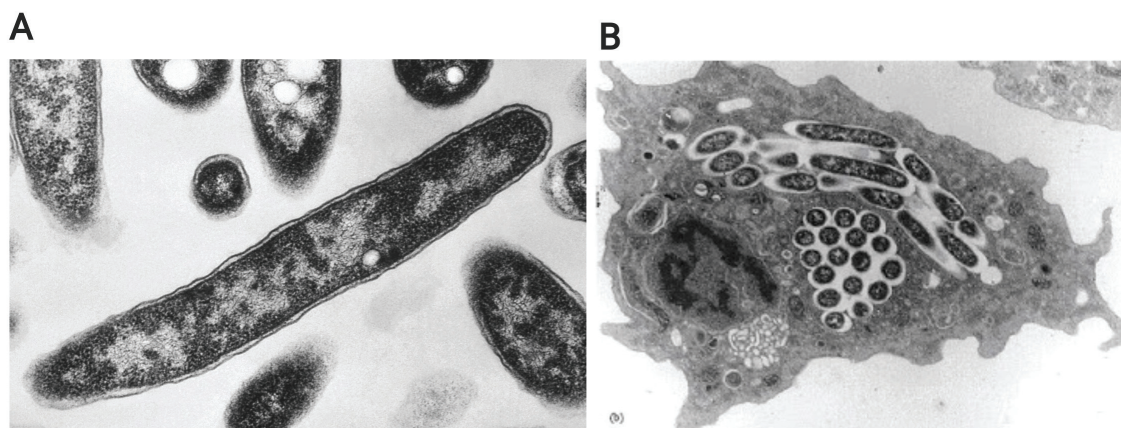
A mild form of *Legionella*-related illness known as ‘Pontiac Fever’ was first identified in 1968 in Pontiac, Michigan. Pontiac fever is considered a non-lethal disease described by flu-like symptoms that heal after a few days without any sort of treatment [4]. That is why the disease is rarely diagnosed even in case of an outbreak, like, for example, the 1976’s outbreak in Philadelphia when public health authorities were able to prove that the same bacteria caused both diseases [5].

## 1.2 The causative agent of Legionnaires’ disease

To date, 65 *Legionella* species have been described [3], but interestingly, 90% of legionellosis are due to species *L. pneumophila*. Amongst them, there are 16 serogroups (Sg) based on the LPS (lipopolysaccharide) structure. For an unknown reason, among all *Legionella* Sg, the Sg-1 is mainly found associated with legionellosis cases [6].

Belonging to *Legionellales* order, *L. pneumophila* is an aerobic, Gram-negative coccobacillus phylogenetically close to the *Coxiella* and *Rickettsiella* genera [7, 8].

The *L. pneumophila* size ranges between 0.3 and 0.4  $\mu\text{m}$  width and between 2 and 4  $\mu\text{m}$  in length [9, 2] (Figure 1 A). In particular conditions such as axenic culture, bacteria can be found in a filamentous state with a length close to 50  $\mu\text{m}$  [10]. Unable to produce endospores [2], the bacterium possesses one or several flagella for motility and fimbriae that seem to be not expressed during the infection [11].



**Figure 1: *Legionella* and protozoa relationship**

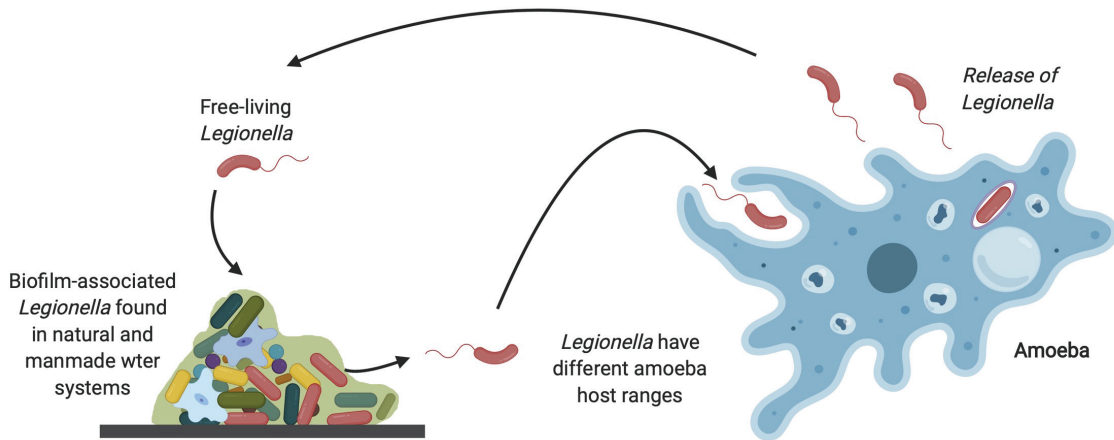
A. *Legionella* under transmission electron microscopy (TEM) [5].

B. *Legionella* inside *Acanthamoeba* spp under SEM [12].

In the environment, *L. pneumophila* is found in freshwater ecosystems, present in both planktonic and biofilm form (Figure 2).

*L. pneumophila* is known to be mostly found associated with protozoa, such as amoeba or ciliate (Figure 1 B). Protozoa provide Legionella with a suitable environment for survival and replication, playing a crucial role in the lifestyle of the latter.

Present in biofilm, several protozoan species have been found associated with *L. pneumophila* [13]. Protozoa, such as amoeba, often graze on bacteria present in biofilm to feed. *L. pneumophila* exploits this phenomenon to replicate, although it is capable of growing off the debris from dead amoebae. Interestingly, by an unknown mechanism, the amoeba *Willaertia magna* C2c maky are able to phagocytose and eliminate *L. pneumophila*. In addition to replication, protozoa provide protection from environmental stressors, such as temperature or biocides used in hydric systems [14]. *L. pneumophila* uses this last as a barrier to protect from hostile environments such as temperature and antibiotics and as a replication niche. Putative coevolution between *Legionella* and the latter has led to the acquisition of various effective defense strategies to bypass the pathway of phagosomal maturation [15].



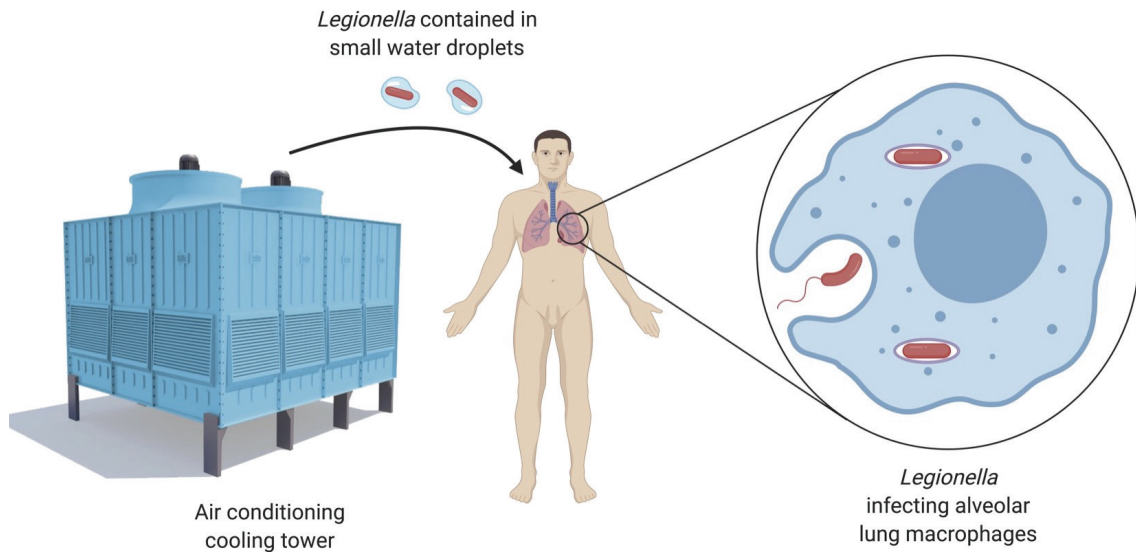
**Figure 2: *Legionella*'s lifestyle in the environment (This study)**

*Legionella* can be found in a planktonic form or associated in a biofilm with another bacteria or protozoa. After phagocytosis by an amoeba, *Legionella* is able to survive and replicate into its host. When the nutrients are insufficient, *Legionella* will be released in the environment via a fine genetic regulation and can infect another host. Modified from [12] and [5].

These strategies allow *Legionella* to create a permissive niche, the *Legionella* containing vacuole (LCV), in its environmental host or an accidental host, human macrophages, allowing their propagation and survival.

### 1.3 Human infection, disease diagnosis and treatment

*Legionella* is also able to thrive through association with biofilms and amoebae in manufactured environments such as hydric systems that have temperatures ranging from 25°C to 45°C [16]. *Legionella*'s ability to form biofilms and replicate in amoebae presents a challenge for tackling and destroying it [17].



**Figure 3: Infection by *Legionella* due to contaminated man-made devices (This study)**

*Legionella* can also be found in man-made hydric systems such as cooling towers where aerosols are created, enabling the infection. The inhalation of aerosols contaminated by *Legionella* can travel to the lower respiratory tract, where many immune cells are present [18]. In the lungs, *Legionella* is phagocytosed by macrophages but can survive and replicate similarly to how it replicates in amoebae [18]. Importantly, humans constitute a dead-end route for *Legionella*, because even infected casualties cannot expel bacteria.

To prevent *Legionella pneumophila* contamination, the hot water distribution systems must be maintained at 60°C, and cold water distribution networks should be kept below 25°C [19]. Moreover, the addition of copper-silver ions and monochloramine have also proven to be efficient measures [20]. The greatest challenge comes from airborne water aerosols formed by devices such as air conditioning systems and cooling towers (Figure 3) [21]. The inhalation of aerosols contaminated by *Legionella* allows bacteria to travel to the lower respiratory tract, where many immune cells, particularly macrophages, are present [18, 17]. Macrophages, like amoeba, are phagocytic cells; as is the case with protozoa, *Legionella* can infect and replicate inside of these cells, eventu-

ally lysing them (Figure 1). Anecdotally, one putative case of inter-human transmission was described in a study from Portugal [22], although aerosol inhalation seems to be the only mechanism of human infection.

Legionnaires' disease is usually diagnosed via chest X-ray [23] and the presence in the patient's urine of *Legionella* specific antigens (LPS), which are detectable from 1-2 days after the onset of the first clinical signs [24]. This test is particularly sensitive to *L. pneumophila* Serogroup-1 (Sg-1), which is predominant in clinical cases [25]. A serogroup (Sg) represents bacteria with the same LPS structure. To date, there is no explanation for the predominance of *L. pneumophila* Sg-1 in clinical cases but not in the environment, although it may result from higher infectivity or more efficient intracellular growth [26, 27]. In France, this serogroup is responsible for 95% of legionellosis cases but represents only 28% of environmental isolates [28].

In Australia and New Zealand, *L. pneumophila* and *L. longbeachae* have similar prevalences and represent between third and half of the community-acquired cases of Legionnaires' disease [7].

Depending on its the severity, Legionnaires' disease is treated with antibiotics by a macrolide monotherapy or a combination of fluoroquinolones and macrolides [29]. However, treatment failures, defined by the absence of clinical improvement and by the persistence of *L. pneumophila* despite appropriate antibiotic therapy, are regularly reported, particularly in immunosuppressed patients [30]. *In vitro*, *Legionella* is sensitive to these two antibiotics, although some strains of *L. pneumophila* Sg-1 isolated from patients exhibit a resistance to fluoroquinolones [30]. Currently, no explanation for treatment failures has been reported, but a patient's genetic background or the presence of bacterial resistance to antibiotics may be involved [31, 32].

The average mortality rate is approximately 10%, even with early and appropriate antibiotic treatment. Studies showed that certain factors such as sex, age, smoking habits, or immunodeficiency also impact on patients: being a man, older than 50 years, or a smoker significantly increases the chance of being infected and developing legionellosis [25, 29].

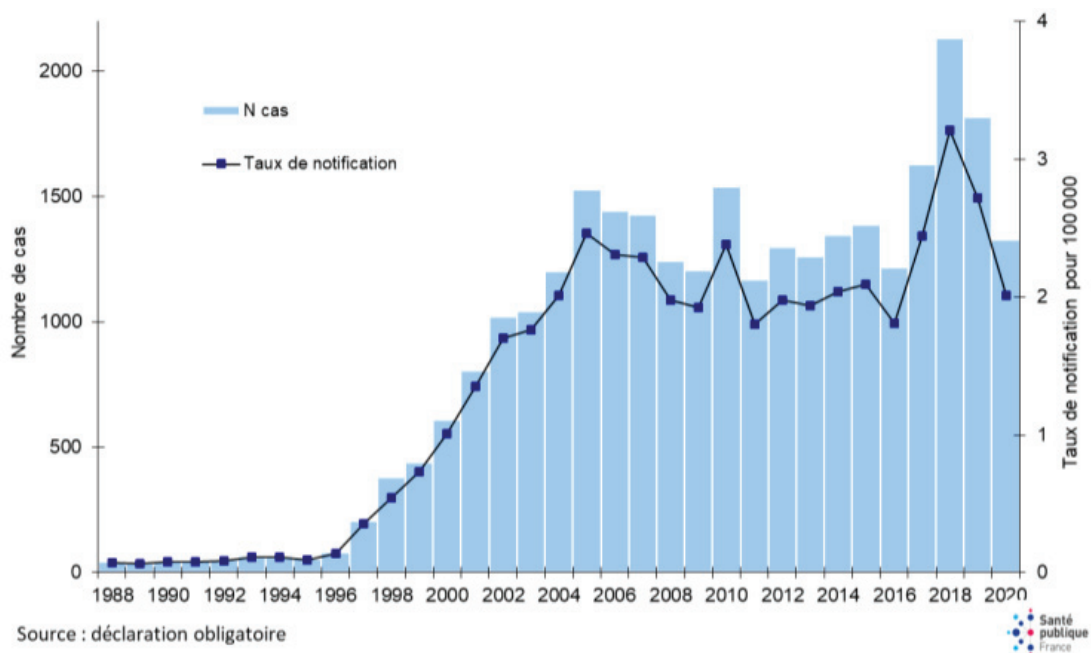


Figure 4: Evolution of reported legionellosis cases in France from 1988 to 2020 (from ‘Santé Publique France’)

[33]

In France, legionellosis has been a reportable disease since 1987, and approximately 1,600 cases are recorded every year (Figure 4). *Le Centre National de Référence pour Legionella (CNRL)* supervises the typing and detection of all environmental samples and samples collected during an outbreak.

The Institut de Veille Sanitaire is responsible for the surveillance of all legionellosis cases. This public institution publishes the results in the ‘*Bulletin Epidémiologique*

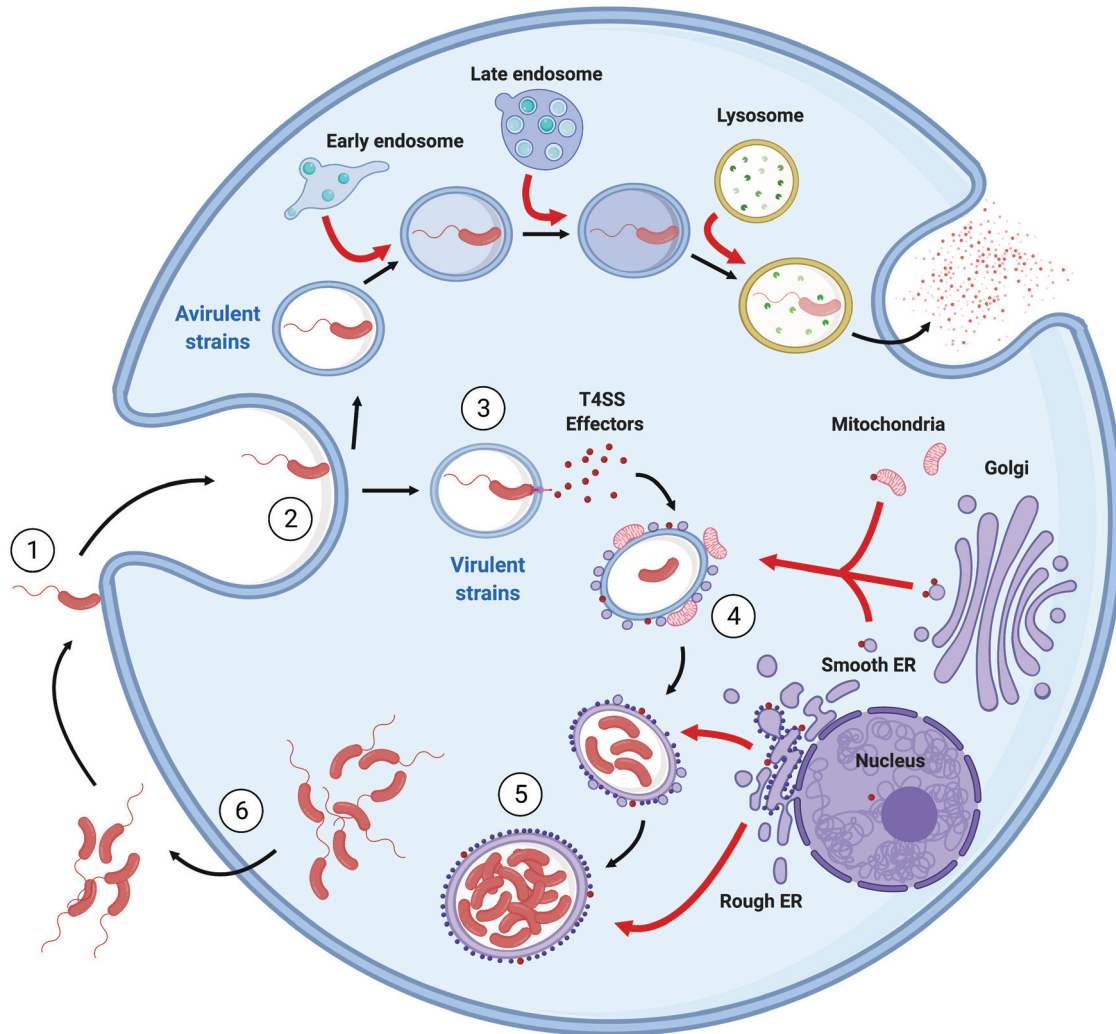
*Hebdomadaire*' [34] every year.

## 1.4 *Legionella* life cycle in amoeba and macrophages

During *Legionella* human lung infection, the primary infection reservoir could be macrophages and neutrophils until all these cells are lysed [35]. Then, *Legionella* can infect alveolar epithelial cells that represent the secondary infection reservoir, allowing bacteria to escape the immune system [36]. The intracellular life cycle of *L. pneumophila* is highly similar both in its environmental host amoebae and its accidental host human macrophages (Figure 5). This is due to the many cellular pathways that are conserved in both amoebae and macrophages [35].

### 1.4.1 Adhesion to eukaryotic cell

*L. pneumophila* flagella allow it to move in the environment until it meets a host to infect [37] (Figure 5). Complement component C1 and C3 receptors were first described to be involved in phagocytosis of *L. pneumophila* in primary monocytes [38].



**Figure 5: *Legionella pneumophila* infectious cycle and LCV biogenesis (This study)**

① *L. pneumophila* finds a eukaryotic host and via several proteins attached to it. ② Bacteria are internalized within the host cell and reside in a phagosome. ③ Virulent strains inject more than 300 effector proteins via their T4SS to control LCV biogenesis and host defenses. ④ Many organelles, such as mitochondria and smooth ER-derived vesicles, are recruited around LCV to escape endocytosis degradation, creating a replicative niche. ⑤ Later on, *L. pneumophila* recruits ribosomes on the LCV surface and replicates itself efficiently within this ER-like vacuole. ⑥ When nutrients dwindle, *L. pneumophila* expressed its virulence factors, such as flagella, to exit the LCV and so on its host.

*L. pneumophila* avirulent strains, such as those deficient for Dot/Icm T4SS, lose the ability to recruit organelles and are rapidly cleared via the classical endosomal pathway.



The bacterial protein Momp, localized at the *Legionella* surface, binds complement Component C1 and C3 on the eukaryotic cell, facilitating the entry of bacteria [39]. Via its Type 2 secretion system (T2SS), *Legionella* secretes Lcl (*Legionella* collagen-like protein) at its surface that interacts with complement component C1q receptor. This interaction promotes the entry of bacteria into monocyte-derived macrophages (U937) and pneumocytes (A549) [40]. Repeats in toxin (RtxA) and EnhC bacterial proteins are believed to be involved in bacterial adhesion to laryngeal epithelial cells (Hep-2, human epithelial cell type 2) and monocyte-derived macrophages (THP-1) [41].

LpnE (*L. pneumophila* entry) is another crucial protein for adhesion of bacteria to monocyte-derived macrophages (THP-1) and pneumocytes (A549) [42]. Other bacterial proteins whose cell targets are not identified are involved in the adhesion of *L. pneumophila*. Interestingly, some of these proteins are not present into non-pneumophila species and could explain the difference in virulence observed between *Legionella* species [43, 44, 45].

Finally, type IV pili have a role in the adhesion of bacteria to epithelial cells (HeLa, Henrietta Lacks), monocyte-derived macrophages (U937), and amoeba *Acanthamoeba polyphaga* [43].

### 1.4.2 Entry of *Legionella*

Once bacteria are attached to a eukaryotic cell, they will be uptaken inside the host inside a phagosome. *L. pneumophila*'s uptake in phagocytic cells such as mouse macrophages and *Dictyostelium* amoebae seems to take place mainly through the macropinocytosis [46] (Figure 5). Macropinocytosis is known to be used as a classical uptake mechanism for many pathogens, such as *Mycobacteria*, *Salmonella* [47] and HIV-1 [48] in non-phagocytic and phagocytic cells. Moreover, experiments using drugs

inhibiting endocytosis suggested the entry of *Legionella* could be possibly mediated by clathrin [49, 50].

### 1.4.3 Intracellular lifestyle and phase transition

*L. pneumophila* adopts a biphasic life cycle: when the bacterium is free in the environment and infects a new host to form its replicative vacuole, it is in the transmissive form (Figure 5). Next, in the replicative form, a phase transition allows the bacteria to multiply actively inside the vacuole [51, 52]. While little is known about the mechanisms involved in the transition from the transmissive to the replicative forms, a notable change is the tolerance of bacteria to acidification. Indeed, at early stage of infection, bacteria inhibit the acidification of LCV and its fusion with lysosomes. However, when the bacteria begin to replicate, the LCV acidifies and acquires lysosomal proteases [51, 52]. Interaction with the endocytic degradation pathway allows bacteria to obtain the nutrients and energy necessary for bacterial replication [53]. The replicative phase also corresponds to the stimulated expression of genes involved in metabolic, energy, and cell-division pathways [52]. Deficiencies of nutrients such as amino acids or fatty acids can serve as signals for bacteria to regain the characteristics of the transmissive form [51, 52] such as flagella and T4SS. Quorum sensing, which allows bacteria to adapt their physiology to the density of the bacterial population, is also involved in this phase transition and in the expression of some Dot/Icm substrates [54, 55]. //

## 1.5 Genome and *Legionella* strains

Since the first outbreak in 1976, many other *Legionella* species have been discovered, and to date, 65 *Legionella* species have been characterized. Another species that also provokes legionellosis, prevalent in the Australian and New Zealand soil, is *Legionella longbeachae*.

Genomic sequences available from a vast number of *Legionella* species showed a broad diversity in terms of genome size and virulence-associated protein genes [56]. Sg1 Philadelphia-1, Paris, and Lens strains were the first to be sequenced. The Philadelphia strain genome was totally sequenced in 2004, 28 years after the first outbreak [57]. The Paris strain genome was wrongly named “Paris” because it was considered to be endemic to the Paris region according to a ten-year study [58]. Between 1987 and 1997, this strain was responsible for a third of legionellosis cases. This strain was later found in another French region and then globally [59, 60]. The Lens strain was identified after the outbreak from November 2003 to January 2004 in Lens city, France. The investigation brought to light that contaminated water droplets could be transported in the atmosphere over a long distance (over 12 kilometers). This had never been reported before [61].

Comparison of the Paris and Lens strain genomes revealed a high degree of genetic variability: 14% of genes from Paris strain were not found in Lens strain, and 10% of Lens strain genes were found exclusively in this strain [62].

Finally, a comparative genomic analysis using several *L. pneumophila* Sg-1s was performed to identify genetic diversity. These strains were the clinical strain Corby, UK [63, 64]; Alcoy strain, Spain [65]; the 130b strain, USA [44]; Lorraine strain, France [66, 67]; and the French environmental strain HL 0604 1035 [68]. Each of the sequenced

strains has a circular chromosome with a size between 3.3 and 3.6 million base pairs, representing about 3,000 genes. Each strain also has a GC (guanine-cytosine) percent of 38% [68]. Each of these strains presents up to 11% specific genes, confirming the genetic diversity within the *L. pneumophila* sg-1 [69].



## Chapter 2

# Type IV Secretion System Dot/Icm and Effectors

Type IV secretion systems (T4SS) plays a crucial role in the pathogenesis of many Gram-negative bacteria. The system looks like a needle allowing the injection of proteins, called effectors, and nucleic acids inside the host. Two different types of T4SSs can be distinguished: a T4ASS that is important for conjugative plasmid transfer and a T4BSS that transfers both proteins and nucleic acids [70]. T4BSS is more complex as it is made up of 27 components versus 12 for T4ASS.

In *Legionella*, T4BSS is also called Dot/Icm T4SS and was discovered by two different teams at the same time. Mutants of *dot* genes (defective in organelle trafficking/intracellular replication) found by the first team are enabled to bypass the phagosomal maturation pathway [71], and *icm* genes (Intracellular multiplication) mutants identified by the other team are unable to replicate inside macrophages [72].

The  $\Delta dotA$  defective mutant, a component of the secretion system whose function is still undescribed, is unable to replicate in phagocytic cells and is usually used as an avir-

ulence control. Experiments using mutants from a mini-Tn10::kan bank and TUNEL assays have shown that the Dot/Icm T4SS secretion system is essential for induction of macrophage apoptosis [73]. Nine Dot/Icm secretion system mutants were isolated and were severely affected in apoptosis induction. In these mutants, an introduction of wild-type gene on a plasmid recovered the wild-type phenotype.

In *Legionella*, T4BSS, also called “Dot/Icm T4SS,” is localized at the bacterium pole, and this particular localization is necessary for *Legionella* virulence and so for LCV biogenesis. Dot/Icm T4SS allows the secretion of more than 300 effectors in the host cytoplasm, which is crucial for bacterial replication in the host cell [74]. Dot/Icm T4SS is the main virulence factor for *Legionella*, and deletion of one of the T4SS-related genes often led to a complete avirulence.

## 2.1 Structure of Dot/Icm T4SS

In 2017, a team described the first *in situ* structure of *L. pneumophila* Dot/Icm T4SS [75]. Dot/Icm T4SS is encoded by 27 *dot/icm* genes localized in two pathogenic islands. The 27 proteins forming the Dot/Icm T4SS apparatus include several inner and outer membrane components that span the LCV membrane [75].

Two different protein complexes exist in *Legionella*’s Dot/Icm T4SS that are crucial for effector translocation: the core transmembrane complex composed of DotC, DotD, DotH, DotF and DotG; and the coupling protein subcomplex: DotL, DotM, DotN, IcmSW and IcmS-LvgA (Figure 6). Many other proteins such as DotA have an essential role in effector translocation, although their function is not yet clear.

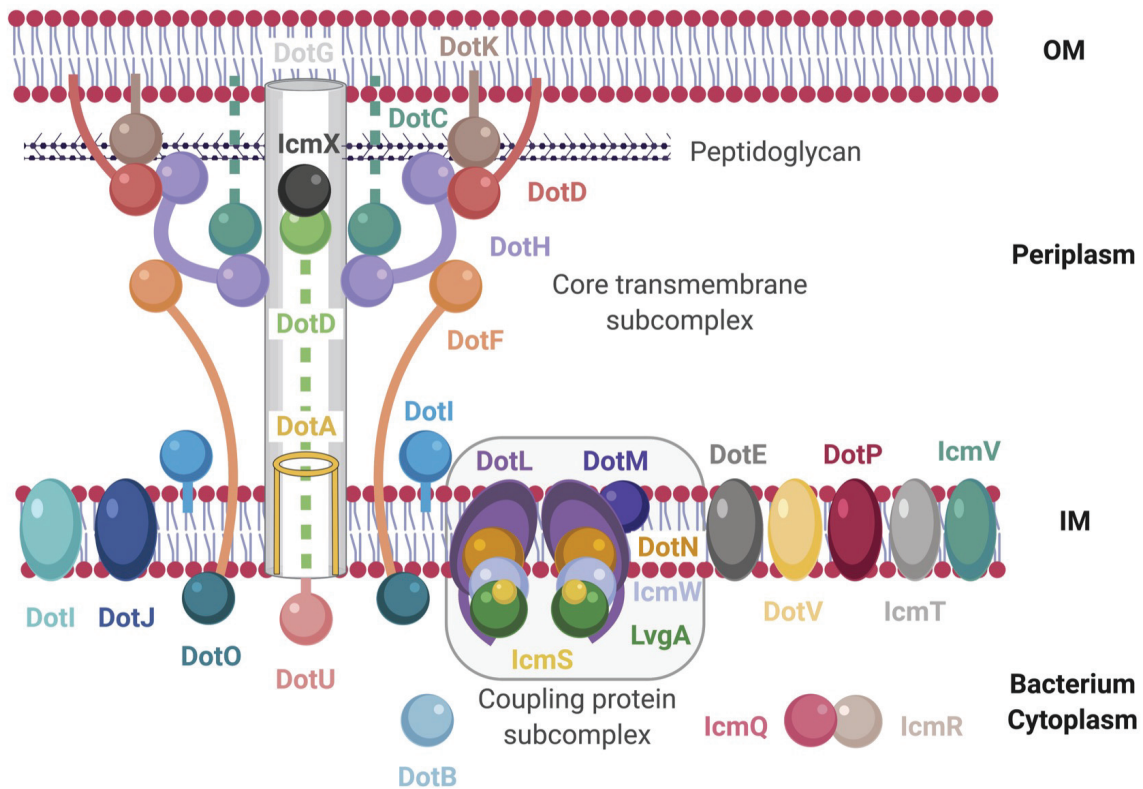


Figure 6: Architectural model and *Legionella* Dot/Icm T4BSS protein composition (This study)

### 2.1.1 Core-transmembrane subcomplex

Dot/Icm T4SS system genomic and biochemical analysis led to the discovery of a major subassembly called the core-transmembrane subcomplex (Figure 6), which is composed of five proteins: DotC, DotD (IcmF), DotF (IcmG), DotG (IcmE), and DotH (IcmK) [75].

Polar targeting of this subcomplex is mediated by two proteins DotU and IcmF (UF), alongside two other Dot/Icm proteins that can pole localize by themselves. Both were described as being crucial for the core-transmembrane subcomplex initial step assembly. The UF proteins are homologs of TssL and TssM, two components of the Type VI



secretion system (T6SS) membrane complex [75]. DotF, DotG and DotH are localized to the bacterial pole due to UF [76]. Even though the DotC role is poorly characterized, it does interact with DotH and can form a periplasmic subcomplex made of DotC/DotD/DotH [75]. DotD is a lipoprotein attached to the outer membrane and has been proposed to be *Xanthomonas citri* VirB7 homolog; both share an N0 secretin domain [77]. This domain is known to be present in the Type II secretion system (T2SS) and functions as an outer membrane channel enabling the passage of folded proteins from the periplasm into the extracellular medium [78]. DotD localization is not well characterized because there is no Dot/Icm T4SS particle formation in a  $\Delta dotD$  mutant [79]. It has been reported that DotF interacts with several Dot/Icm substrates, but these interactions, which were identified using a yeast two-hybrid screen, have to be confirmed [80].

DotF was finally suggested to interact with and regulate the activity of DotG in the inner membrane [81, 82]. Because it is a VirB10 homolog from *Agrobacterium tumefaciens*, DotF is present in both T4ASS and Dot/Icm T4SS and presents similarly in their C-termini [83]. VirB10 transduces energy from the inner membrane to the outer membrane via TonB action [84]. A homolog of VirB9 is DotH forming a ring just below the outer membrane with DotC, even though it is not a lipoprotein [79].

### 2.1.2 Coupling protein subcomplex

In addition to the core-transmembrane complex, another subcomplex exists in Dot/Icm T4SS, the coupling protein subcomplex (Figure 6), composed of DotL (IcmO), DotM (IcmP), DotN (IcmJ), IcmS, IcmW and LvgA [85, 75].

Coupling proteins are considered crucial elements in Dot/Icm T4SS because they are thought to play two essential roles in substrate secretion. They constitute an inner membrane receptor that links substrates to the secretion machinery and provides energy for substrate translocation from the ATP hydrolysis [86]. Interestingly, DotL contains a nucleotide-binding motif and shows sequence similarity to the conjugal transfer protein TrbC [87] and ATPases named T4SS coupling proteins (T4CPs), such as VirD4 from *A. tumefaciens* [88]. DotL has been proposed to function as the T4CP for Dot/Icm secretion system and may likely perform a critical role. It has been proposed that T4CPs act like molecular pumps secreting effectors via ATP hydrolysis [89]. This kind of ATPase forms a hexameric ring like a channel where substrates come through during translocation. The opening and closing of this channel are dependent on ATP fixation and release [90].

In the bacterium inner membrane, DotL interacts with DotM, and DotN stabilizes DotL [91]. The DotL/DotM/DotN complex then interacts with IcmS/IcmW to better stabilize DotL. IcmS and IcmW form a chaperone complex necessary for the translocation of specific effectors into the host cell and function as a Type III secretion chaperone [92]. The IcmSW complex interacts with a wide variety of Dot/Icm substrates, such as SdeA, SidA, SidB, SidC, SidD, SidE, SidG and SidH, and is required for their export via the T4BSS [93]. *IcmS* and *icmW* defective mutants can replicate to some extent, but eventually, their LCVs will fuse with lysosomes [94]. This indicates that the effector chaperoned by IcmS/IcmW may be necessary for avoiding lysosomal fusion. IcmS was also characterized as interacting with LvgA to form the IcmS/LvgA complex, and it can also interact with a various range of effectors like lcmSW [95].

### 2.1.3 Cytoplasmic components

IcmS, IcmW, LvgA, IcmQ and IcmR are T4SS Dot/Icm components are well-characterized proteins that are localized in the cytoplasm [96]. They can interact with each other and act as a chaperone complex where IcmR is the chaperone and IcmQ its substrate [97, 98]. In oligomeric form, IcmQ can insert into lipid membranes forming pores up to 26Å [97]. IcmR/IcmQ interaction inhibits the formation of IcmQ oligomeric form and prevents it from being found associated with the membrane [97, 99]. DotB and DotO are ATPases essential to effector translocation and belong to the same classes as VirB/D4 systems [80].

### 2.1.4 Proteins with unknown function

Among the Dot/Icm T4SS components, many proteins do not yet have a characterized function, such as IcmX. IcmX is the only soluble periplasmic protein of the Dot/Icm T4SS and has been found well conserved throughout the *Legionellaceae* family, in (*L. pneumophila*, *L. micdadei*, *L. bozemanii*, *L. gratiana*) and in *Coxiella* [94, 100]. During *L. pneumophila* culture, a fragment of IcmX (165-466) is secreted into culture supernatant depending on T4BSS. However, the secretion of this fragment into the eukaryotic cytoplasm has not been detected [100].  $\Delta icmX$  mutants were found unable to replicate in neither human (U937) or A/J-derived mice macrophages even though T4BSS was still well-assembled [100]. IcmX seems to play an essential role in the regulation of *L. pneumophila* phagosome trafficking, a crucial step for replicative niche establishment [101].

## 2.2 The 300 Dot/Icm effectors

T4SS Dot/Icm translocates a massive number of over 300 bacterial proteins, called effectors, into the host cell cytosol. Unfortunately, only around 60 have a described function during the infection (Annexe Table [1](#)). Effectors are crucial for many steps of *Legionella* infection, and their expression is regulated at the transcriptional level: Dot/Icm effectors are produced before contact with the host cell.

Effectors from *Legionella* have been identified using different techniques: *in silico* searches for genes encoding eukaryotic-like domains [\[45\]](#), altered growth viability [\[102\]](#), the presence of a C-terminal secretion signal [\[103\]](#), interaction with the Dot/Icm components [\[104, 92\]](#), yeast genetic assays [\[105, 106\]](#) and a machine-learning approach [\[107\]](#).

Comparative analysis of five *Legionella*'s species (*L. pneumophila*, *L. longbeachae*, *L. micdadei*, *L. hackeliae* and *L. fallonii*) have highlighted the presence of more than 400 proteins with a putative eukaryotic origin. Phylogenetic reconstruction clearly showed that among these proteins, 40 had been acquired through horizontal gene transfer from eukaryotes [\[108, 56\]](#). Interestingly, *L. longbeachae* and *L. fallonii* possessed around two times more proteins with EMs than any other genomes. This could be easily explained by their larger genome size (4.1 Mb for *L. longbeachae* versus 3.5 Mb for *L. pneumophila* Paris strain) [\[108\]](#). Among the 300 effectors secreted by *L. pneumophila*, several do share functional redundancy. The deletion of a single gene effector or a group does not usually have an impact on the fate of an LCV. This suggests that some effectors target the same cellular pathway [\[109, 110\]](#). Due to this redundancy, classical genetic approaches have not been very successful, and to date, only around fifty effectors are functionally characterized.

Overall, T4SS effectors are involved at each step of *L. pneumophila*'s infection and target the host cellular pathway to hijack the cell, creating a permissive replicative niche. For example, they allow bacteria to escape from phagosomal degradation, rearrange the actin cytoskeleton, or even change the host gene expression to its advantage.

The effector translocation from bacterial cytoplasm to host cytoplasm requires the recognition of a secretion signal located at the C-terminal region [111, 112]. Although this secretion signal is very different between effectors, it is characterized by a 35-amino-acid, negatively charged Glu-rich signal at the C-terminal important for the recognition by Dot/Icm system [113]. A recent study showed that the timely delivery of Dot/Icm substrates was c-di-GMP-signaling-dependent and not dependent on the chaperones picking up [96]. Below, we will provide an overview of *L. pneumophila* effector roles during infection.

### 2.2.1 Endosomal-lysosomal pathway escape

The main trait of the *L. pneumophila* infection cycle is to avoid the degradation by the endosomal pathway and create an ER-derived replicative niche.

The endosomal-lysosomal pathway comprises a set of intracellular membranous compartments that dynamically interact with each other, including early and late endosomes and lysosomes. After phagocytosis, bacteria are contained in a phagosome that will mature following interaction in a specific order with endocytic pathway organelles. The phagosome maturation consists of the rearrangement of its membrane composition and a drastic change in its content, becoming more acidic and oxidative. At the end of phagosome maturation, all the optimal conditions are present to degrade internalized

bacteria.

First, the phagosome will fuse with early endosomes and so will expose specific patterns of the early endosome at its surface. The fusion with early phagosome causes enrichment of phosphatidylinositol-3-phosphate (PtdIns(3)P), Rab5 and EEA1 (early endosomal antigen 1) proteins. EEA1 is anchored on the phagosome membrane via its interaction with PtdIns(3)P, allowing it to interact with Rab5. The change in membrane composition seems to be involved in the attachment with other vesicles, such as multivesicular body or late endosome. The early phagosome presents a weak hydrolytic activity and a pH between 6.1 and 6.5 driven by the acquisition of proton pump sets called vacuolar ATPases (V-ATPases). V-ATPases are protein complexes that catalyze ATP hydrolysis using released energy to import hydrogen ions ( $H^+$ ) through the phagosome membrane. After fusion with late endosomes, the phagosome will present LAMP (lysosomal-associated membrane proteins) and Rab7 on its surface. These are late-endosome-specific proteins. Rab7 is known to control trafficking between phagosome and lysosomes as it can recruit specific proteins on the phagosome membrane, such as RILP (RAB-interacting lysosomal protein). RILP interacts with dynein or dynactin molecular motors that are associated with microtubules promoting migration of phagosome to lysosomes.

Finally, the phagosome and lysosomes fuse, creating a phagolysosome. At this step, the phagolysosome becomes more acid (pH 4.5) by recruiting other proton pumps. This leads to the inhibition of bacterial growth and an increase of hydrolytic enzymes.

By modifying LCV composition with ER markers, such as PtdIns(4)P, and recruiting small GTPases, *L. pneumophila* inhibits recruitment of early endosomes and spatially moves the LCV away from the endocytic degradation pathway. In addition, *L. pneumophila* prevents the fusion of an LCV with endosomes and lysosomes, while at the same time limiting its acidification [114].

*Legionella* can also interfere with phagosome acidification. Indeed, SidK, via its N-terminal domain, interacts specifically with VatA, a subunit of the v-ATPase proton pump, leading to its inactivation [114]. During the early stages of infection, the LCV's acidification inhibition prevents the degradation of bacteria. Interestingly, it has been shown that SidK's expression in macrophages prevents degradation of non-pathogenic *Escherichia coli* by inhibiting phagosome acidification [115].

### 2.2.1.1 LCV lipidic composition modulation

In eukaryotic cells, cellular compartments are mainly characterized by their lipid composition, notably, phosphoinositides [116]. Phosphoinositides are lipids derived from phosphatidylinositol (PtdIns) that can be phosphorylated/dephosphorylated via cellular kinases/phosphatases.

Early stages of LCV formation consist of modulation of lipidic membrane composition and recruitment of particular proteins on its surface. Within 1 min of bacterial uptake, the LCV is composed by *PtdIns*(4,5) $P_2$  (phosphatidylinositol-4,5-bisphosphate) and *PtdIns*(3,4,5) $P_3$  (phosphatidylinositol-3,4,5-triphosphate) that are a set of specific patterns of internalization by phagocytosis/macropinocytosis and PtdIns(4)P (phosphatidylinositol-4-phosphate) derived from the plasma membrane [117, 116]. During phagosome maturation, *PtdIns*(3,4,5) $P_3$  and PtdIns(4)P are normally modified to PtdIns(3)P (phosphatidylinositol-3-phosphate), a crucial element for the recruitment of EEA1 or SNX1 that together promote endosomal fusion [117, 116].

Dot/Icm secretion is essential to avoid PtdIns(3)P accumulation and so on endocytic pathway activation, as an  $\Delta icmT$  defective mutant cannot avoid [118]. Instead of Pt-

dIns(3)P, the LCV will present PtdIns(4)P, a lipid mainly found in the Golgi apparatus and the ER [116]. PtdIns(4)P accumulation allows the LCV to recruit and fuse with ER vesicles to escape endocytic degradation, increasing the LCV's size [119] (Figure 7). Several effectors contribute to the LCV lipidic composition, such as LepB. This effector is a bifunctional protein possessing PtdIns 4 kinase (PI4K) activity at the N-terminal region [120] and RabGAP activity at the C-terminal [121].

3D structure analysis and biochemical investigation led to identifying that LepB N-terminal domain (LepB\_NTD) possesses a catalytic motif similar to PI4K [120]. This domain converts PtdIns(3)P from the LCV into  $PtdIns(3,4)P_2$ , and it was found to be functional in eukaryotic cells. LepB\_NTD's ectopic expression in HeLa cells was reported to cause total disruption of Golgi-related structures [120].

SidF is one of the first effectors discovered to modify the phagosome phosphoinositides composition (Figure 7). It possesses PtdIns polyphosphate 3-phosphatase activity, hydrolyzing  $PtdIns(3,4)P_2$  and  $PtdIns(3,4,5)P_3$ . It is anchored on the LCV via transmembrane domains. SidF phosphatase activity causes PtdIns(4)P accumulation on the LCV, necessary for endocytic pathway evasion. It also facilitates PtdIns(4)P-binding effectors anchoring to the LCV [122]. With a similar mechanism, SidP uses its PtdIns polyphosphate 3-phosphatase activity to decrease vacuolar PtdIns(3)P accumulation [123].

A cysteine phytase, LppA, was demonstrated to have PtdIns phosphatase activity. Analysis of the LppA catalytic site (HCRGGKGRT) showed a high similarity with human PtdIns 3-phosphatase PTEN (phosphatase and tensin homolog) (HCR/KAGKGRT) and PtdIns phosphatase MptpB (HCFAGKDRT) from *Mycobacteria*. LppA was described as dephosphorylating *in vitro*  $PtdIns(3,4)P_2$  and  $PtdIns(4,5)P_2$  very efficiently and  $PtdIns(3,4,5)P_3$  and  $PtdIns(3,5)P_2$  with less efficiency [124] [47] (Figure 7). Dur-



ing *Dictyostelium discoideum* infection, LppA is translocated into infected cells but seems not to play a major role in PtdIns(4)P accumulation. Within 1 h or 2 h postinfection and compared to a wild-type strain, an *lppA* defective mutant accumulates the same amount of LCV PtdIns(4)P, suggesting other effectors participate in PtdIns(4)P accumulation [125]. In addition to its phosphatase activity, LppA possesses a phytase activity, degrading phytate into inositol and orthophosphates. Phytate can chelate micronutrients such as iron [124], which is an essential nutrient [126, 127]. With phytate secretion, *Legionella* could probably release and uptake the host's iron [47]. Together, the synergistic actions of LepB, SidF, SidP and LppA led to LCV PtdIns(4)P enrichment (Figure 7).

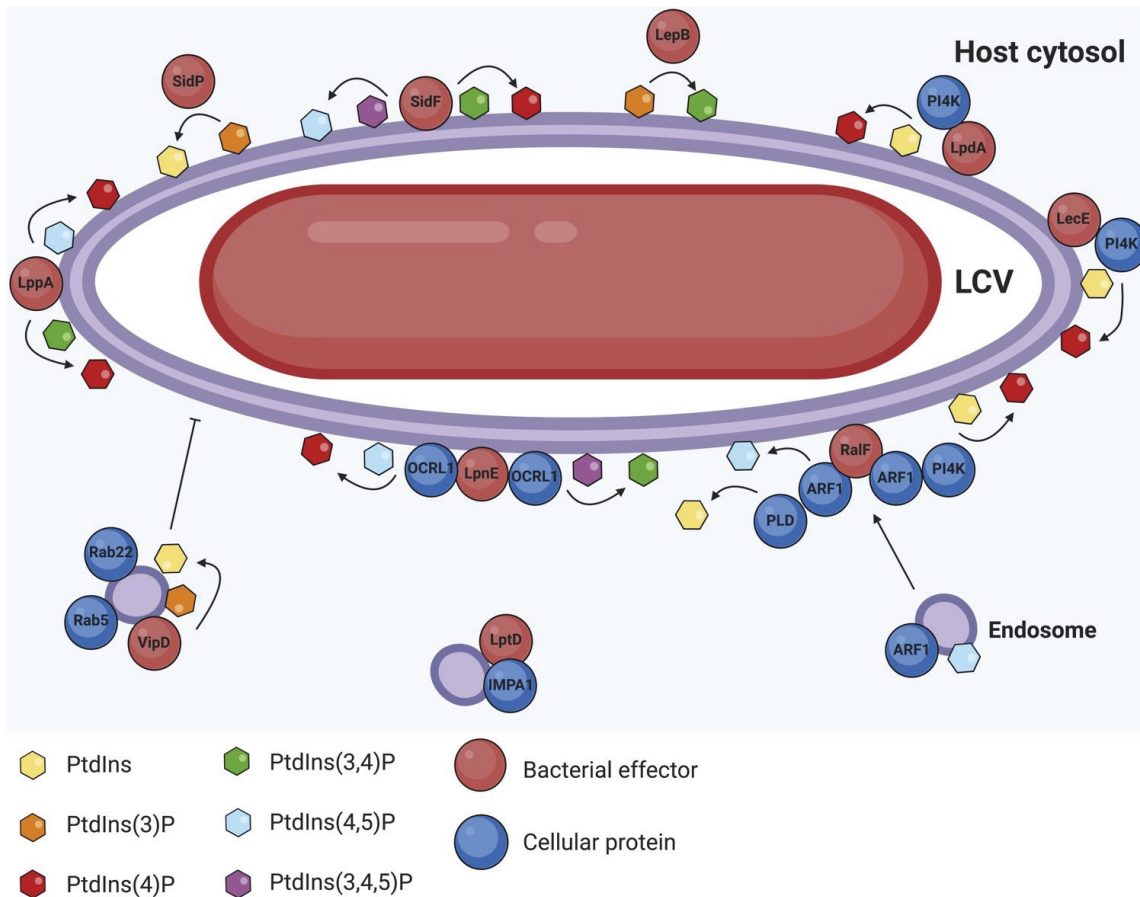
Some of the *L. pneumophila* effectors are unable to directly modify phosphoinositides but need to recruit host enzymes to modify LCV lipidic composition.

RalF is localized and anchored to the LCV via its C-terminal capping domain. This effector has a GEF activity enabling recruitment and activation of ARF1 host trafficking small GTPase [128] (Figure 7). ARF1, like other members of the ARF family, regulates *PtdIns(4, 5)P<sub>2</sub>* synthesis in Golgi apparatuses [129, 130]. *PtdIns(4, 5)P<sub>2</sub>* is required for many cell signaling aspects, such as endocytosis/exocytosis and reorganizing cytoskeleton. It can also act as a substrate for generating lipid-derived second messengers. ARF1 is also known to recruit different coat proteins to form transport vesicles leading to control of protein transport between organelles [131]. As an example, ARF1 can recruit coatomers, a soluble macromolecular complex. It is involved in membrane trafficking through the Golgi apparatus and in COPI (Coat protein I) vesicle formation, mediating protein transport from the Golgi apparatus to the ER [132, 133]. ARF family members have been shown to activate phospholipase D (PLD) that degrades phospholipids such as phosphatidylcholine (PC) or PtdIns [129]. They can also act as cellular process sec-

and messengers, such as vesicle trafficking. ARF1 can recruit PI4KIII $\beta$  and may lead to LCV PtdIns(4)P accumulation [134, 135].

Present on the LCV membrane, the combined action of LpdA and LecE activity could contribute to LCV PtdIns(4)P enrichment, converting PC to DAG (diacylglycerol) [136, 137] (Figure 7). After secretion, LpdA will be post-translationally modified using host palmitoylation activity to anchor it on the LCV surface. LpdA induces the Golgi apparatus disruption and seems to contribute to *L. pneumophila*'s virulence in permissive mice. Via its phospholipase activity, LpdA hydrolyzes PC in phosphatidic acid (PA), then LecE probably via a host phosphatase that converts PA into DAG. Accumulation of higher LCV DAG levels allows the recruitment of PKC and PKD on the LCV surface, leading to PtdIns(4)P accumulation via LCV PI4KIII  $\beta$  recruitment [136, 138].

LpnE can recruit a host PtdIns-metabolizing enzyme such as OCRL1 (oculocerebrorenal syndrome of Lowe 1), which is a PtdIns 5-phosphatase [139, 42, 140] (Figure 7). Localized on the LCV membrane, OCRL1 is implicated in several cellular processes, such as phagocytosis, endocytosis and endosome–trans-Golgi apparatus retrotrafficking [140, 141, 142]. They both convert  $PtdIns(4, 5)P_2$  and  $PtdIns(3, 4, 5)P_3$  into PtdIns(4)P and  $PtdIns(3, 4)P_2$ , respectively. This will lead to PtdIns(4)P increase, making it available for SidC or SidM anchoring.



**Figure 7: LCV phosphatidylinositol composition modulation (This study)**

These lipidic modifications of the LCV are essential to help bacteria avoid the host endosomal degradation pathway. *Legionella* secretes effectors that can either modify PtdIns composition via their activity or recruit cellular proteins to do so.

VipD (VPS inhibitor protein D) is a phospholipase whose activation is dependent on Rab5 or Rab22 endosomal GTPases binding [143] (Figure 7). This catalytic activity does participate in decreasing endosome's PtdIns(3)P level and leads to recruitment inhibition of crucial proteins for membrane fusion such as EEA1 [143, 144]. VipD disrupts the endocytic degradation pathway, allowing the LCV to avoid fusion with early endosomes [145].

LtpD (*Legionella* translocated protein) binds PtdIns(3)P present on early endosomes, late endosomes, lysosomes and also on the LCV surface [146] (Figure 7). LtpD in-

teracts with eukaryotic IMPA1 (inositol monophosphatase), a phosphatase producing myo-inositol, a precursor of PtdIns and, therefore, has an impact on endosomal trafficking disorganization [146].

### 2.2.1.2 Host small GTPases recruitment and modification

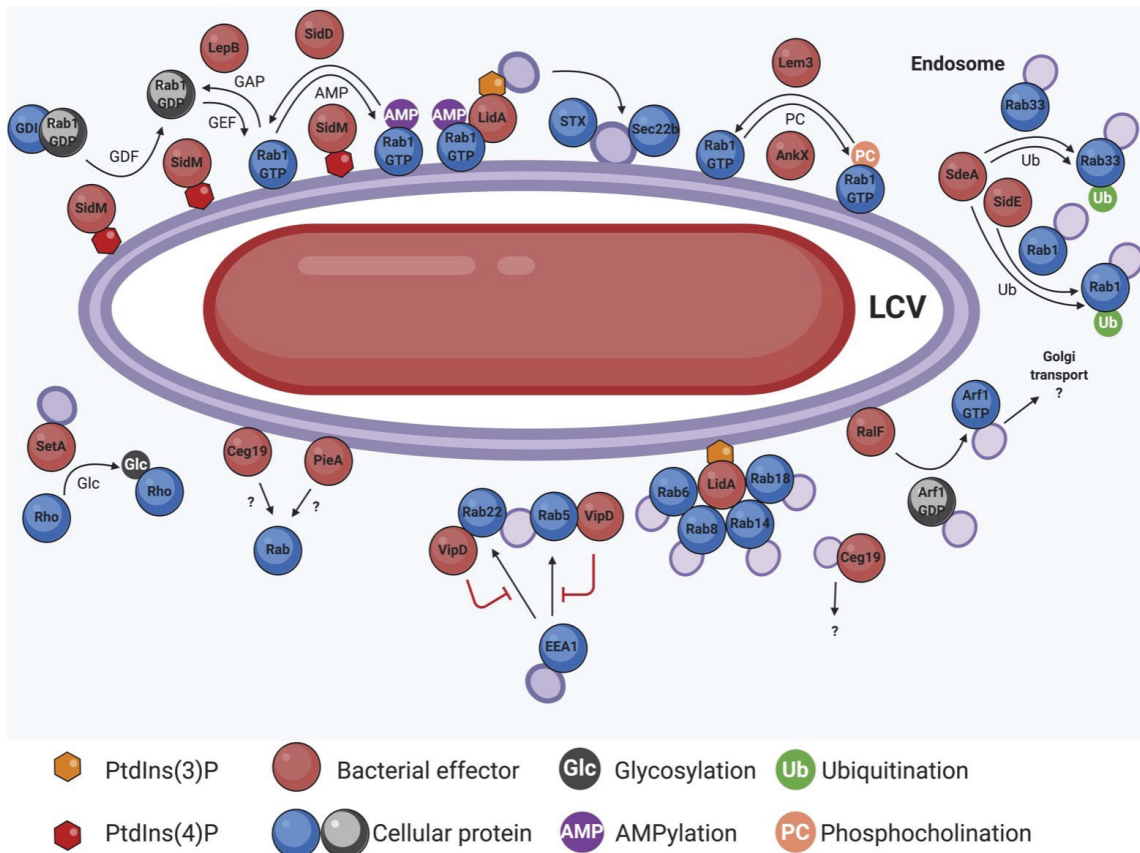
GTPases are molecular switches organized mainly in two protein families: Rab (RAS (rat sarcoma) analog in the brain) and ARF (ADP (adenosine diphosphate)-ribosylation factor) GTPases. The activity of these proteins is modulated by GDI (guanosine dissociation inhibitor), GDF (GDI dissociation factor), GEF (guanine nucleotide exchange factor), and GAP (GTPase-activating protein) activities [147]. Using effectors targeting host GTPases, *Legionella* can modulate at its benefit the host vesicle trafficking dynamic to its own benefit (Figure 8).

- **Rab1 GTPase**

Rab1 is known to be a crucial protein for *Legionella* or other intracellular pathogens during infection. In addition to interacting with ARF1 on the LCV, *Legionella* effectors can recruit other host small GTPases. This interaction allows fusion between the LCV and ER vesicles to expand its vacuole and escape to endocytic degradation (Figure 8). SidM/DrrA was the first bacterial effector found to recruit and activate host cellular Rab1 GTPase, known to be involved in trafficking regulation between the ER and Golgi apparatuses [148]. SidM is anchored on the LCV via its interaction with PtdIns(4)P, where SidM can recruit and activate PtdIns(4)P via its GEF activity [148, 149]. Due to its adenylyltransferase activity, SidM can also activate Rab1 by modifying it with AMP. Rab1 AMPylation blocks GAP action and inhibits GDI, leading to a permanent

activation status [150, 151, 152].

LidA (lowered viability in the presence of dotA) potentiates the SidM/DrrA effect. LidA is anchored to the LCV due to its PtdIns(3)P and PtdIns(4)P interaction domain [153, 148]. It can stabilize Rab1 in the activated/AMPylylated form and may participate in its LCV sequestration [154]. Moreover, LidA could also act as an adaptor interacting with Rab1 or the LCV [155]. By an unknown mechanism, LidA can also interact with other small GTPases such as Rab6, Rab8, Rab14 and Rab18 found in Golgi apparatuses and also on the LCV surface [156, 157, 158, 149]. Rab6 and Rab8 can interact with cellular OCRL1 and may participate in PtdIns(4)P accumulation on the LCV surface [159].



**Figure 8: Host small GTPase recruitment and modification by effectors (This study)**

Effectors can modify host GTPase from the Rab family or others to finely inhibit the endosomal degradation pathway for its own benefit. Effectors modify host GTPases by glycosylation, AMPylation, ubiquitination or phosphocholination, allowing bacteria to control the LCV's future.

When ER vesicles are closed to the LCV, SidM boosts membrane fusion due to host SNARE (soluble N-ethylmaleimide-sensitive factor attachment protein receptor) [160]. This could be explained by the presence of STX (syntaxin) proteins derived from the plasma membrane on the LCV and Sec22b found on ER vesicles [161].

A third effector, AnkX, targets Rab1 GTPase catalyzing another post-translational modification (PTM), called phosphocholination, without SidM action [162]. In infected cells, Rab1 can be found in both AMPylated or phosphocholinated form, and the

LidA effector could probably participate in the stabilization of both modified states [155, 152]. Furthermore, AnkX can also modify Rab35, which is involved in the plasma membrane and endosome vesicles trafficking [157]. Whereas Rab1 phosphocholination moderately decreases SidM GEF activity, phosphocholinated Rab35 inhibits its activation by the connecdenn (denn (differentially expressed in neoplastic versus normal cells) domain-containing protein) cellular protein [163]. Even though the precise role of Rab35 phosphocholination is not clearly characterized, this modification on Rab1 is known to participate in the LCV, avoiding fusion with late endosome by inhibiting microtubule vesicles trafficking [162].

Rab1 AMPylation or phosphocholination does not have an impact on cellular GAP activity but, interestingly, inhibits LepB activity [150, 163]. Importantly, Rab1 can be de-AMPyated by SidD and de-phosphocholinated by Lem3, two other effectors [164, 165]. Then, the LepB GAP activity can also inhibit other small GTPases such as Rab3, Rab8, Rab13 and Rab35 at the LCV surface, suggesting a broad activity spectrum [166] and, thus, contributing to a fine temporal activation of Rab1. Rab35, like Rab1, is regulated by SidM AMPylation, which will activate it, de-AMPyated by SidD and finally inactivated by LepB [166, 167]. At the same time, Sac1 (suppressor of actin) is a host phosphatase that decreases PtdIns(4)P concentration on the LCV surface [168]. This will lead to effectors like SidM, which is involved in early LCV maturation, to be released from the LCV. Together, those interactions participate in temporal modulation of the *L. pneumophila* effectors.

Many other effectors, such as SidE family members, are known to modify Rab1 GTPase and modulate its activity, suggesting a fine temporal control of Dot/Icm effector secretion [169].

- Other GTPases

Due to its C-terminal domain, VipD binds to small GTPases present on early endosomes such as Rab5 and Rab22. These interactions would prevent Rab5 and Rab22 from binding to substrates such as EEA1, Rabaptin-5 or Rabenosyn-5, blocking endosomal trafficking and lysosomal degradation [170, 143] (Figure 8).

Lpg0393 (*L. pneumophila* Philadelphia-1 gene) is an effector that contains a Vsp0-like domain that is very similar to the human Rabex-5 catalytic core, suggesting that it can also activate Rab5, Rab21 and Rab22 by its GEF activity [171]. Even though its role is not yet determined, Rab5 and Rab22 activation by Lpg0393 could participate in VipD endosome targeting and inhibit early endosome recruitment on the LCV surface [171].

SetA (subversion of eukaryotic traffic) binds to PtdIns(3)P and can be found on early endosomes, late endosomes, and lysosomes alongside the LCV surface [172, 106]. It is a multidomain protein that possesses glucosyltransferase activity at its N-terminal and is anchored at early endosomes via its C-terminal PtdIns(3)P-binding domain. Although its role is not yet identified, its N-terminal domain presents a homology to *Clostridium difficile*'s AB toxins [173]. These enzymes are known to inactivate small GTPases from the Rho family through glycosylation and could participate in the endocytic degradation pathway's disorganization [173].

Ceg19 can be found on late endosomes surfaces and would play a role in transport between Golgi apparatuses, late endosomes, and lysosomes [106]. Ceg19 is also recruited on the LCV surface and interacts with several Rab GTPases involved in the ER, Golgi apparatuses and endosomes trafficking. Even though these Rab GTPases may locate on the LCV where they have various activities, we cannot rule out that this effector may also disrupt vesicular trafficking [106].



### 2.2.2 Nutrition and energy

Within its replicative niche, *L. pneumophila* must obtain nutrients to multiply efficiently. Some Dot/Icm effectors specifically target and hijack host pathways, such as the proteasome, to compete with the cell for nutrient acquisition.

Protein ubiquitination is considered as a eukaryote-specific post-translational modification critical for regulation of numerous cellular processes, such as cell homeostasis [174], immune response [175], cell cycle progression [176], DNA repair [177] and vesicular trafficking [178]. Ubiquitination is mediated by the sequential action of three enzymes: E1 ubiquitin-activating, E2 ubiquitin-conjugating and E3 ubiquitin-protein ligase. At the start of the ubiquitin cascade, the E1 enzyme activates ubiquitin using ATP- $Mg^{2+}$ , leading to its binding to an E1 cysteine residue in the thioester bond. Then, activated ubiquitin in the E1-ubiquitin complex is transferred to an active site of the E2 ubiquitin-carrier protein on a cysteine residue. The last step is mediated by ubiquitin-protein ligase or E3 enzyme, consisting of the transfer of ubiquitin from the E2-ubiquitin complex to the substrate protein lysine residue. As many host cellular pathways, the ubiquitin machinery is hijacked by *L. pneumophila* to monitor the activity of its effectors, cellular proteins, or obtain nutrients.

At the beginning of the infection, upon the first contact between host and *Legionella*, host-mediated farnesylation and ubiquitination machinery are recruited on the plasma membrane. This recruitment is dependent on Dot/Icm system and exclusively localized at the bacterial attachment site [179].

*Legionella* can hijack host farnesylation machinery and, in particular, RCE1 (RAS converting CAAX endopeptidase) to the anchored effector on the LCV, such as AnkB [180]. There, it can recruit and interact with the SCF complex that ubiquitinates many proteins on the LCV's lysine 48 residue [179]. AnkB can be poly-ubiquitinated on the

lysine 11 residue by host TRIM21 (tripartite motif-containing protein) via its E3 ubiquitin ligase activity. Surprisingly, this modification does not affect its stability, but its role is not yet characterized [180]. AnkB is colocalized with host ParvB, suggesting their interaction and decreasing ParvB ubiquitination level. This protein is known to be involved in actin cytoskeleton dynamic and cell survival [181] [182] [183].

Mainly known to interfere with autophagy, RavZ was shown to be able to interfere with ubiquitin recruitment of *Salmonella*-containing vacuoles [184]. This observation suggests a putative role in ubiquitin deconjugation and probably a similar role in *Legionella*.

LegU1 interacts with host SKP1, leading to host BAT3 chaperone's ubiquitination, and could increase cell resistance to ER stress [182]. Lpg2160, AnkB/LegAU13 and LicA can also interact with SKP1, regulating other substrate functions through ubiquitination [182] [183].

SdeA/LaiA, SdeB, SdeC and SidE are members of the SidE family and are involved in ER- and LCV-associated small GTPases' ubiquitination, such as Rab1 and Rab33. However, their roles are not yet clear. Interestingly, SdeA ubiquitinates Rab33b, Rab1, Rab6a and Rab30 without the help of host E1 and E2 [185]. Moreover, SdeA also might add an ADP-ribosylated ubiquitin to these proteins, although this function has not yet been described [186, 187].

SidC and SdcA are localized to the LCV and can interact with PtdIns(4)P *in vitro* [81, 188]. They both can ubiquitinate Rab1 via their ubiquitin ligase activity. Even though their role is not yet precisely characterized, they may be involved in the regulation of Rab1 during recruitment of the LCV and ER vesicles [189, 190, 188]. Moreover, these effectors also possessed a de-ubiquitination domain essential to finely control the balance between ubiquitination and de-ubiquitination mediated by *Legionella* [191].

WipB is a serine/threonine phosphatase related to the eukaryotic phospho-protein phos-

phatase (PPP) family. It localizes to the lysosomal compartment to dephosphorylate lysosomal LAMTOR1 and two subunits of the v-ATPase [192]. These two proteins, whose activities are dependent on phosphorylation, are members of the lysosomal nutrient sensing (LYNUS) apparatus. They control the mammalian target of rapamycin (mTORC1) kinase complex [192]. During *L. pneumophila*'s infection, WipB targets lysosomal phosphoprotein to modulate host nutrient sensing for its own benefit.

Iron ( $Fe^{2+}$ ) is an essential metal for microbial growth, and intracellular pathogens have developed many sophisticated strategies to uptake and minimizing host cell iron storage. At the same time, hosts have also developed mechanisms to reduce iron acquisition by pathogens to limit infection [193].

MavN was first characterized as an iron transporter localized to LCVs, but a recent study demonstrated that MavN could also transport other divalent ions such as  $Mn^{2+}$ ,  $Co^{2+}$  or  $Zn^{2+}$  [194, 195]. Consistent with these observations, a  $\Delta mavN$  mutant exhibits a transcriptional iron-starvation signature and an intracellular growth defect that can be rescued only by the addition of iron excess. Interestingly, even though MavN was also described as transporting other metals, the addition of other metals in the culture medium did not rescue intracellular growth defects during infection. All these findings show that MavN plays a crucial role in iron uptake during *Legionella*'s infection.

LncP (*Legionella* nucleotide carrier protein) is a highly hydrophobic protein that can be anchored on the mitochondrial inner membrane [196]. This protein may be involved in ATP transport outside mitochondria to provide bacteria energy [196].

## 2.2.3 Vesicular trafficking modulation

### 2.2.3.1 Actin cytoskeleton

Actin exists in two different forms in eukaryotic cells: globular (G-actin) and filamentous form (F-actin), composed of polymerized G-actin. Arranged into filaments, the actin cytoskeleton plays a major role in various cellular processes such as mobility, phagocytosis and vesicular trafficking [197]. This is probably why actin is targeted by many intracellular pathogens such as *Listeria*, and *L. pneumophila* is not an exception to the rule; at least five Dot/Icm effectors target this pathway (Figure 9).

VipA was identified by yeast screening and shown to inhibit lysosomal trafficking [170]. Localized on early endosome surfaces, VipA interacts *in vitro* with actin and promotes actin nucleation [198]. VipA was proposed to impact endosomal trafficking due to its interaction with actin [198, 199]. Another effector, SidL (substrate of Icm dot transporter/coregulated with effector genes)/Ceg14, does the opposite, inhibiting spontaneous actin polymerization and, therefore, its arrangement [200].

LegK2 is translocated during infection on the LCV and phosphorylates host ARPC1B and ARP3 subunits from the ARP2/3 actin nucleator. Consequently, it inhibits actin polymerization, preventing LCV and late endosome/lysosome association [201]. LegK2 importantly contributes to bacterial endosomal degradation escape and LCV remodeling into a replicative niche by ER recruitment. WipA (IcmW-interacting protein A) is a tyrosine phosphatase inhibiting actin polymerization [202]. It can dephosphorylate many actin-associated proteins such as N-WASP and ARP2/3 [203]. WipA phosphatase activity leads to actin polymerization disturbance by reducing the F/G-actin ratio, which is a marker of actin polymerization status. Infection experiments using a defective  $\Delta wipA$  mutant showed in cells a significantly higher ratio than those in-

fectured with the wild-type strain. Because LegK2 and WipA both negatively affect actin polymerization, the authors investigated the potential synergistic effect of these two effectors, constructing a double knockout strain  $\Delta wipA/\Delta legK2$ . In macrophages, this double mutant has a significant intracellular defect compared to a single mutant but is complemented using only *wipA* [203]. Both LegK2 and VipA act to reduce actin polymerization and host trafficking.

Finally, RavK targets the host cytoskeleton and reduces actin filament abundance in mammalian cells [204]. RavK is a protease that cleaves actin between the T351 and F352 residues, preventing actin polymerization. *In cellula* experiments, overexpressing a non-cleavable form of actin (actin F352A) suppressed RavK-induced phenotype [204], suggesting RavK has a role in actin modulation.

### 2.2.3.2 Microtubules

Microtubules are highly dynamic structures made of  $\alpha$   $\beta$ -tubulin that are an essential component of the eukaryotic cell cytoskeleton, mitotic spindle, and flagella and serve as tracks for intracellular trafficking [205].

The effector LegG1, also called PieG, manipulates microtubules by promoting the activation of the small GTPase Ran (RAS-related nuclear protein) and is involved in many cellular processes such as nuclear transport [206] or microtubule nucleation [207]. LegG1 displays a homology sequence with host Ran GEF (guanine nucleotide-exchange factor) RCC1 (regulator of chromosome condensation 1) and contains a tetrapeptide CAAX motif at the C-terminal that will be prenylated, leading to its anchoring on the LCV [208, 209]. Thus, LegG1 might activate Ran by its GEF activity, leading to the recruitment on the LCV of RanBP1 (RAN binding protein 1), promoting microtubule

polymerization and stability [208, 210]. LCV motility on the microtubule network could be useful when removing phagosomes from the plasma membrane and moving closer to the ER.

Another effector RidL inhibits retrograde trafficking, thus playing a central role in endosomal trafficking. The retromer-mediated transport involved cargo protein delivery from endosomes back to the trans-Golgi network (TGN) and endosomal transmembrane proteins recycling to the cell surface [211]. This transport pathway is crucial for organelle identity, lipid homeostasis and many other cellular functions. During the infection, RidL interacts with PtdIns(3)P on the LCV and also with Vps29 [212], one of the components of retromer cargo recognition subcomplex Vps26-Vps35-Vps29 [213]. *In vivo*, RidL acts like a competitor for host Vps29 regulators: the Rab 7 GTPase-activating protein TBC1d5 (Tre-2/Bub2/Cdc16 domain family member 5) or VARP [214].

LtpM was found to block microtubule-related endosome movements. LtpM localized on Rab5- and Rab7-containing endosomes, where it reduces microtubule motility due to its glucosyltransferase activity [215]. During infection, LtpM was also found localized to the LCV and may have a role in an endosomal pathway or LCV movement. Moreover, the authors showed that LtpM overexpression in mammalian cells does not affect retromer-mediated retrograde transport but has a strong impact on endosome mobility related to LtpM's glucosyltransferase activity [215].

### 2.2.3.3 Other mechanisms

*L. pneumophila* globally disrupts the host trafficking pathway by the secretion of a large number of effectors by its Dot/Icm T4SS.

LegC2 (YlfB), LegC3 and LegC7 (YlfA) can disrupt the endolysosomal trafficking pathway in yeast and could participate to avoid the fusion between the LCV and lysosomes [216]. *In vitro*, LegC3 inhibits SNARE- and Rab GTPase-dependent membrane fusion in yeast [216]. This suggests that LegC3 could inhibit endosomal degradation by inhibiting LCV's fusion with lysosomes. LegC7/YlfA are effectors that disrupt the delivery of both biosynthetic and endocytic cargo to the yeast vacuole [109].

LpdA, already mentioned for its phospholipase activity, is also found on the plasma membrane, early endosomes and Golgi surface. Phosphatidic acid produced by this phospholipase disturbs Golgi apparatus integrity and could disrupt plasma membrane vesicular trafficking [137].

*Legionella* uses host farnesylation machinery to anchor LseA (*Legionella* SNARE effector) to the Golgi membrane. Even though its role is not well known, this effector would act as a SNARE protein and would, therefore, be involved in regulating membrane fusion at the Golgi apparatus level [217].

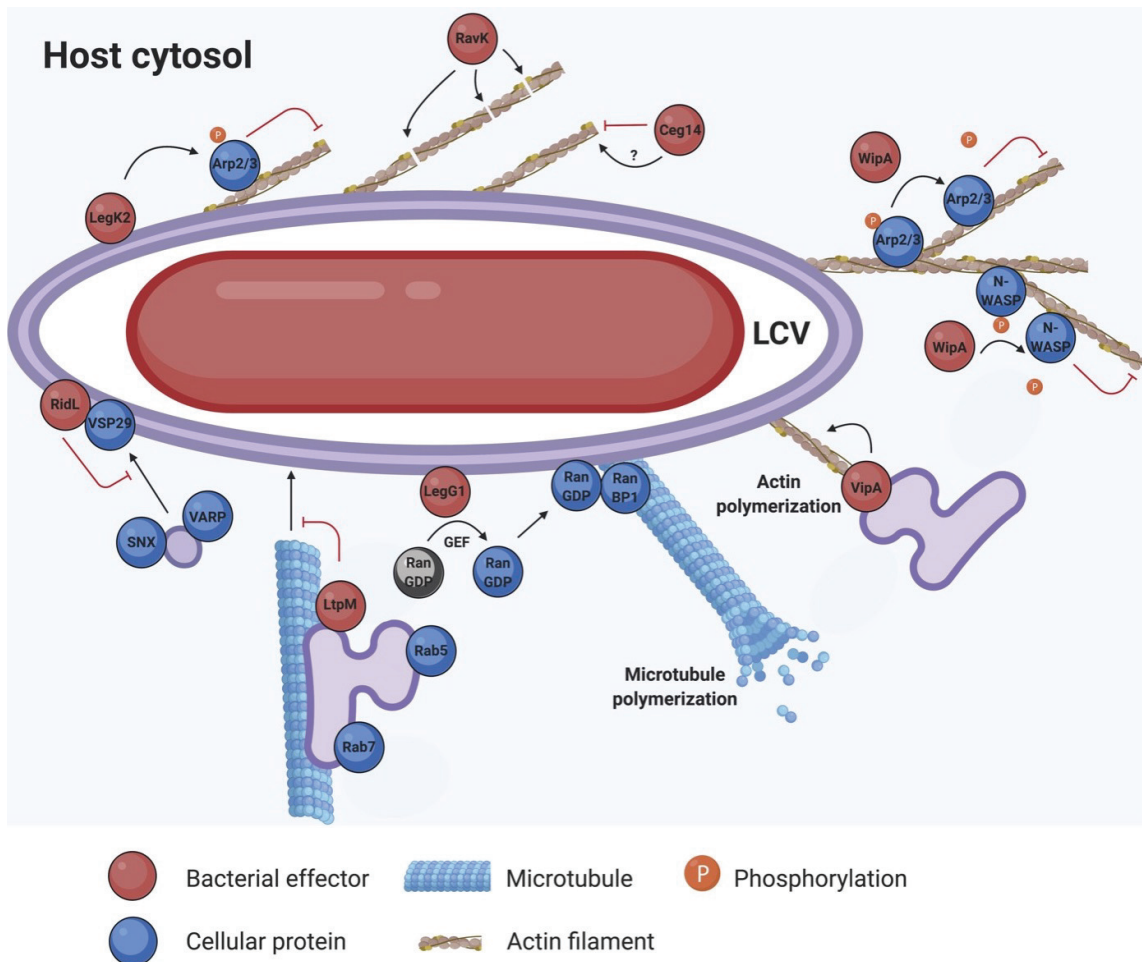


Figure 9: Impact of *Legionella*'s effectors on vesicle migration and LCV mobility (This study)

RavD binds to phosphatidylinositol-3-phosphate on the LCV and could be a part of the molecular mechanism that steers the LCV away from the endolysosomal maturation pathway [218].

Ceg9 disrupts trafficking occurring between the ER and Golgi apparatuses [219].

Based on sequence homology, VipF possesses a putative N-acetyltransferase [170, 220]. Moreover, it also has a strong similarity level with GCN5-related N-acetyltransferases (GNAT), enzymes that use acetyl coenzyme A (CoA) to transfer an acetyl group to many substrates, such as histones or antibiotics [221]. This modification is well known



to have a role in metabolism and chromatin remodeling and to be used by bacteria to inactivate antibiotics. Interestingly, chloramphenicol, an antibiotic produced by *Streptomyces venezuelae*, is a VipF substrate [220]. Even though it is not yet clarified, VipF could participate in the inactivation/activation of eukaryotic proteins or antibiotics.

## 2.2.4 Autophagy

Intracellular bacteria such as *Shigella* or *Listeria* target host cellular mechanisms such as autophagy and nutrition sensors to their own benefit. This allows bacteria to escape the host defenses and participating bacteria to uptake nutrients [222]. *Legionella* is generally thought to inhibit host cell autophagy, notably using three Dot/Icm effectors.

RavZ has been characterized as the first bacterial effector inhibiting host autophagy [223]. It specially binds to high-curvature membranes present in large numbers on PtdIns(3)P, such as phagophores and autophagosomes [224]. RavZ functions as a cysteine protease and irreversibly deconjugates host Atg8/LC3 proteins coupled to phosphatidylethanolamine (PE) on autophagosomal membranes [223, 225]. Host ATG4, which is also a cysteine protease, plays the same role as RavZ but in a reversible way. Interestingly, RavZ's lipid-binding site shows a similar folding to yeast phospholipid transfer proteins (Sec14 family) [226]. During infection, using a *ravZ* defective mutant, autophagosomes are still not recruited to the LCV, suggesting that other effectors hijack the autophagy pathway [223].

*L. pneumophila* Philadelphia-1's effector Lpg1137 has been experimentally characterized as a serine protease able to block autophagy and apoptosis [227]. It has been

shown that Lpg1137 presents a high identity with mitochondrial carrier proteins SLC25 (solute carrier family 25) and can cleave human STX17 (syntaxin 17) on mitochondria-associated ER membranes (MAM) and mitochondria [227, 228].

LegS2 (*Legionella pneumophila* S1P-Lyase)/LpSL is able to restrain autophagy by drastically reducing LC3, a marker of autophagy, in human macrophages [229].

## 2.2.5 Apoptosis, cell cycle and death

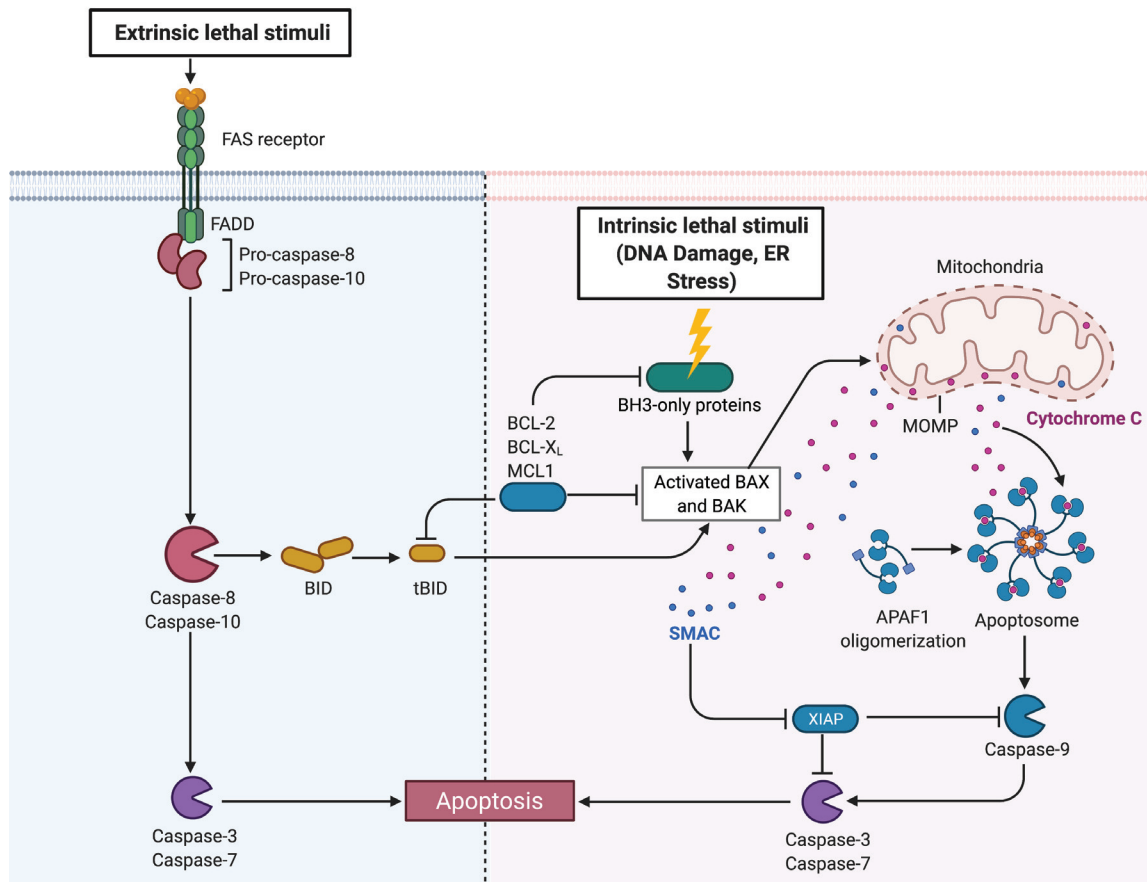
In addition to hijacking many vesicular traffic pathways, *L. pneumophila* finely regulates the survival of its host as well. Indeed, the bacterium first keeps its host alive to establish its replicative vacuole and then multiply. When replication is complete, the bacteria then trigger the death of their host to be released into the environment and thus infect new cells. Multiple Dot/Icm effectors have been characterized as regulating host cell death, although the overall mechanism remains poorly understood (Figure 11) [230, 231, 232].

### 2.2.5.1 Apoptosis

Apoptosis is one mechanisms of cell death used by eukaryotic cells. It is an active, programmed, and autonomous process of cellular dismantling without activation of inflammation [233]. This mechanism can be initiated by two distinct signaling pathways: the extrinsic pathway, triggered by the attachment of pro-apoptotic ligands to death receptors; and the intrinsic pathway, which is brought on by intracellular stress. Mitochondria play an important role in the intrinsic pathway, notably via the release of cytochrome c [234].

Apoptosis also involves specific proteases, called caspases, which can be divided into two groups: the so-called initiating caspases (caspases 2, 8, 9 and 10) activate the effector

caspases (caspases 3, 6 and 7) [233].



**Figure 10: Apoptotic pathways (This study)**

In the extrinsic apoptotic pathway, an extracellular ligand binds to death receptors such as FAS, which can activate initiator caspases (caspase 8 and caspase 10) through their dimerization, mediated by adaptor proteins such as FAS-associated death domain protein (FADD). Active caspase 8 and caspase 10 then cleave and activate the effectors caspase 3 and caspase 7, leading to apoptosis. The intrinsic pathway of apoptosis requires mitochondrial outer membrane permeabilization (MOMP). Cell stresses engage BCL-2 homology domain 3 (BH3)-only protein activation, leading to BAX and BAK activity that triggers MOMP. Anti-apoptotic proteins such as BCL-2 or MCL1 counteract this. Following MOMP, mitochondrial intermembrane space proteins, such as second mitochondria-derived activator of caspases (SMAC) and cytochrome c, are released into the cytosol. Cytochrome c interacts with apoptotic protease activating factor 1 (APAF1), triggering apoptosome assembly, which then activates caspase 9. Active caspase 9, in turn, activates caspase 3 and caspase 7, leading to apoptosis. Mitochondrial release of SMAC facilitates apoptosis by blocking the caspase inhibitor X-linked inhibitor of apoptosis protein (XIAP). Caspase 8 cleavage of the BH3-only protein BH3-interacting death domain agonist (BID) enables crosstalk between the extrinsic and intrinsic apoptotic pathways. (ER: endoplasmic reticulum; MCL1: myeloid cell leukaemia 1; tBID: truncated BID).

The latter cleave different substrates, resulting in cell fractionation. Cellular dismantling causes characteristic and specific morphological changes with nuclear and cytoplasmic condensation, as well as the formation of apoptotic bodies [235].

Although *L. pneumophila* triggers the activation of caspase 3 during the early stages of infection, apoptosis is only eventually induced in the late stages [236]. This delay in the activation of the cell death via apoptosis is due to the implementation of an anti-apoptotic balance by *Legionella* through various strategies (Figure 11). The SidF effector inactivates BCL-Rambo and BNIP3, two members of the pro-apoptotic BCL-2 family [237].

The interaction between SidF and these two proteins thus increases apoptotic resistance, although the deletion of the *sidF* gene has shown that it is necessary only late during infection [238]. Another way for *Legionella* to stop apoptosis is through indirect activation of the expression of anti-apoptotic genes via the NF- $\kappa$ B transcription factor [239].

Once the NF- $\kappa$ B pathway has been activated, IKK (I $\kappa$ B kinases) phosphorylate inhibitory proteins of the I $\kappa$ B family, leading to the release of the NF- $\kappa$ B factor and its translocation into the cell nucleus [233]. The LnaB bacterial effector strongly induces this cell pathway, but its mode of action has not yet been identified [240]. Unlike LnaB, the mechanism of activation of the NF- $\kappa$ B pathway by the Dot/Icm LegK1 substrate has been elucidated. LegK1 is a kinase that mimics host IKK and phosphorylates I $\kappa$ B, leading to the activation of the canonical NF- $\kappa$ B pathway [241, 240]. Interestingly, LegK1 can also induce a non-canonical NF- $\kappa$ B pathway by phosphorylating p100, leading to its cleavage into p52 [241]. Like many effectors, LegK1 is not essential for bacterial replication but activates the NF- $\kappa$ B pathway in macrophages or during *Acanthamoeba*

*castellanii* infection. The activation of host NF- $\kappa$ B signaling likely plays an essential role in the modulation of macrophage defenses, such as by preventing cytokine release by other cells.

The effectors Lgt1, Lgt2, Lgt3, SidI and SidL are also known to target this pathway by inhibiting the host's translation machinery, thereby decreasing I $\kappa$ B production [242, 243]. Notably, the inhibition of translation is due in part to the glycosylation of the elongation factor eEF1A [244]. It has been shown that another post-translational modification may alter the expression of immune genes: the methylation of histone H3 by RomA/ LegAS4 [245, 246]. In addition, the SdhA effector contributes to maintaining of the integrity of the LCV and is involved in the prevention of cell death [247]. In the absence of *sdhA*, bacterial phospholipase PlaA is no longer regulated, causes earlier release of bacteria into the cytoplasm of host cells, thus triggering immune responses of caspase types and cell death by pyroptosis [248, 249].

Finally, *L. pneumophila* stimulates apoptosis via various effectors, thus promoting its exit from the host cell. It seems that these bacteria proceed in two stages: first, they break the membrane of the LCV and enter the cytoplasm of the cell. They then continue to multiply before acquiring the virulence traits of the transmissive phase and being released into the external environment [250]. Electron microscopy observations reveal that the end of the infectious cycle is characterized by a condensation of the chromatin in the cells, indicating cell death by apoptosis [251].

A screening of *Legionella* effectors that activated the caspase 3 has led to the identification of Ceg18, Lpg0716, Lem12, LegS2, and VipD [230]. Most of the identified effectors are localized in the mitochondria, but only VipD has had its mechanism for activat-

ing caspase 3 finely detailed. VipD destabilizes the mitochondrial membrane with its phospholipase A activity, releasing cytochrome c, which activates the apoptotic pathway [252]. *L. pneumophila* thus finely modulates cell apoptosis, either inhibiting or activating it, following the stage of the infectious cycle (Figure 11).

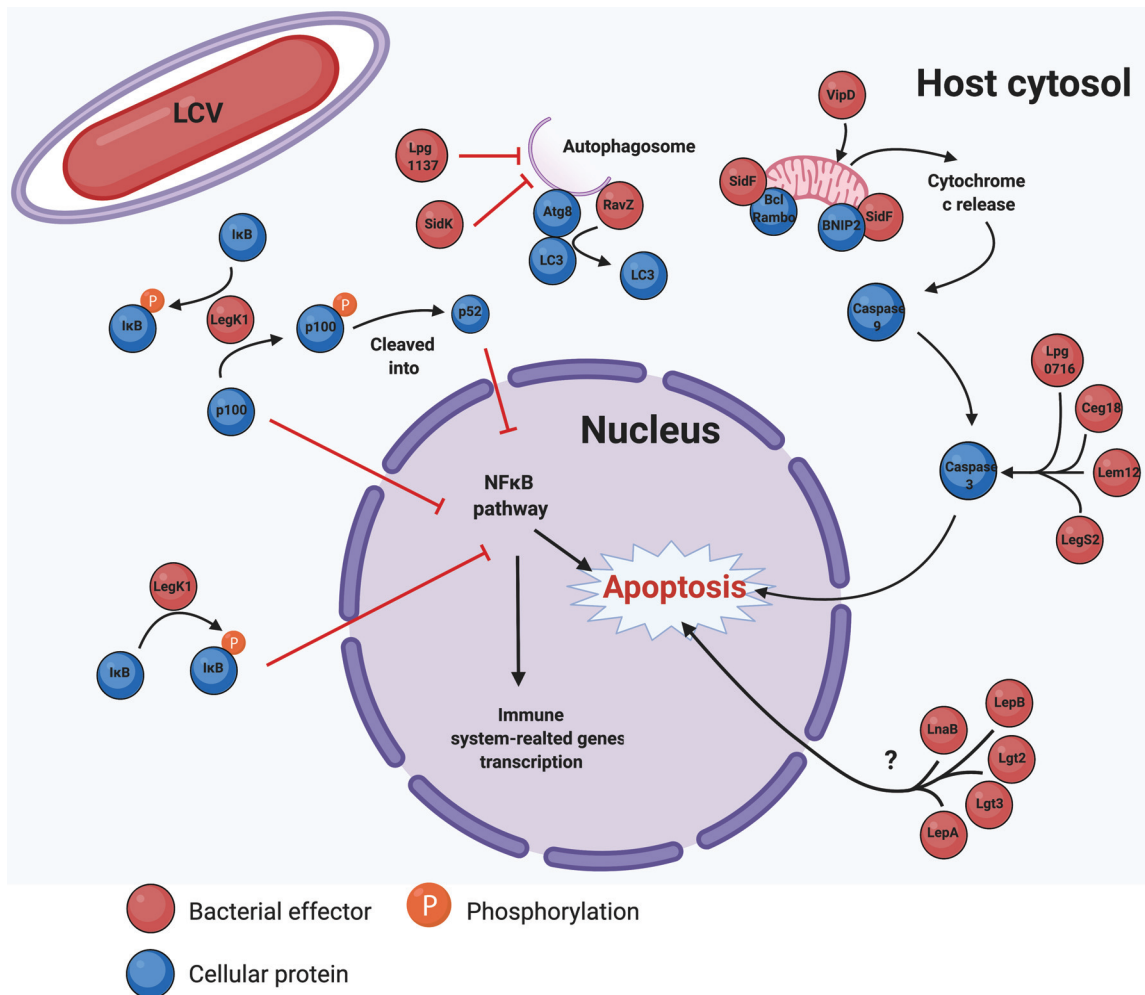


Figure 11: *Legionella* effector having an impact on autophagy or cell death (This study)

### 2.2.5.2 Cell cycle and bacterial egress

It has been previously demonstrated that the S phase of the cell cycle provides a hostile environment for bacterial replication, while cells at G1 and G2/M phases are

more permissive for *L. pneumophila* replication [253]. The translocation of five known translation inhibitors, Lgt1, Lgt2, Lgt3, SidI and SidL, during the G1 phase is required to induce cell cycle arrest [254].

It is also accepted that *L. pneumophila* could lyse its host cells using secreted proteins. It also has been proposed that the bacteria actively destroy the LCV membrane and plasma membrane through pore formation [118]. The discovery of mutants unable to lyse their host, being trapped in the cytoplasm after effective multiplication, strongly supports this hypothesis [255]. The cytolytic activity of *Legionella* is believed to be derived from various secreted proteins such as phospholipases [256]). As an example, phospholipase PlcC, hydrolyses a wide range of lipid membrane compounds which, during infection, could likely destabilize the different cell membranes [252]. However, this model of pore formation lysis is quite controversial since one study found that the release in cytoplasm was not dependent on this mechanism and that plasma membrane lysis would only involve apoptosis [250]. Finally, *L. pneumophila* could exit its protozoan hosts through a non-lytic process where the LCV membrane would fuse with the plasma membrane [257]. It has been proposed that LepA and LepB effectors be involved in this phenomenon, due to their similarity to SNARE membrane fusion proteins. However, the molecular mechanism put in place remains to be elucidated, particularly the role of LepB, already known to target the small GTPase Rab1 [164].

### 2.2.6 Host gene expression control

Host gene expression modulation allows *L. pneumophila* to establish proper conditions for its replication. *L. pneumophila*'s ability to reduce innate immunity gene



expression and also to induce proinflammatory gene response demonstrates its ability to manipulate the host depending on the intracellular life cycle step. Some effectors have been described as controlling host transcription or translation machinery via post-translational modification (PTM). Some effectors can migrate to the nucleus (Figure 12).

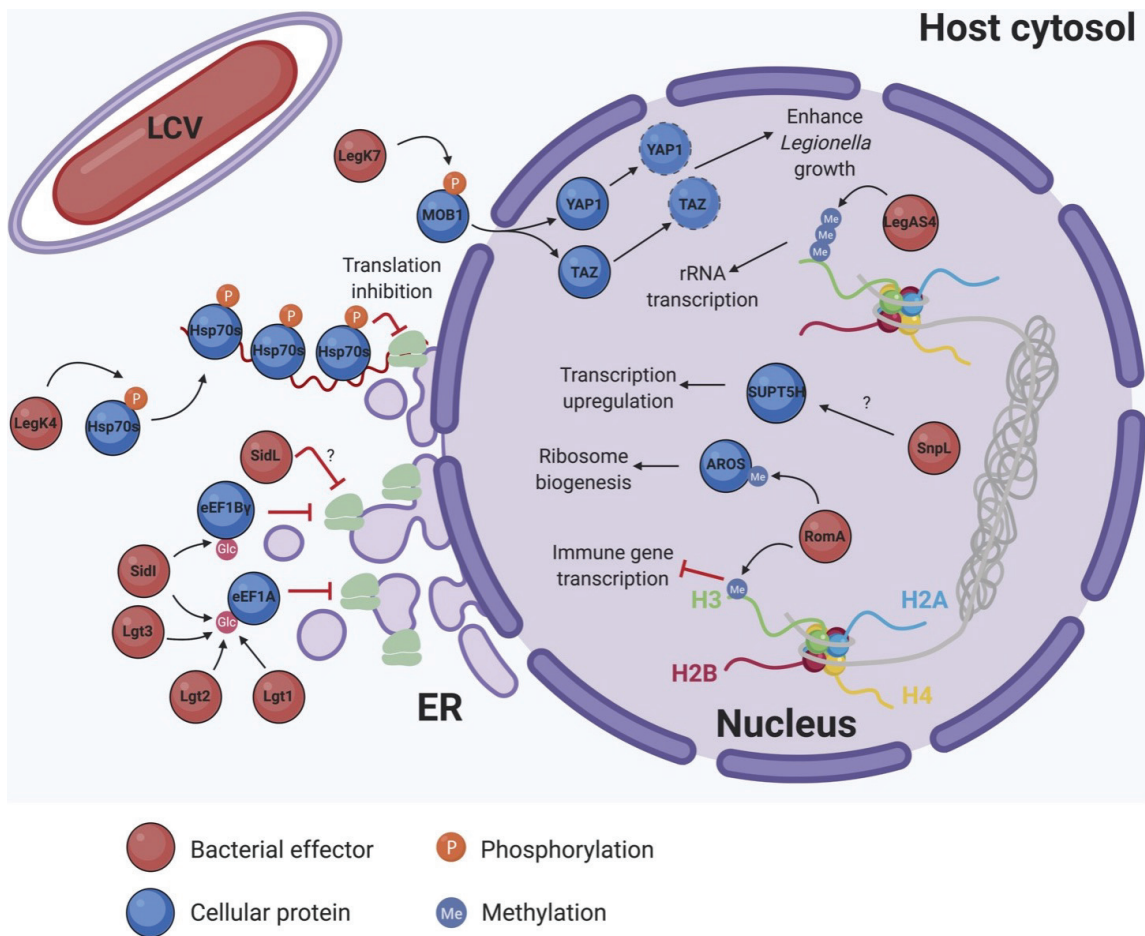


Figure 12: *Legionella* effectors involved in host gene expression (This study)

### 2.2.6.1 Transcription

To modulate the eukaryotic gene transcription, epigenetic controls by histone post-translational modification (HPTM), such as methylation, acetylation and phosphorylation, can be used by *L. pneumophila*. These post-translational modifications have an impact on chromatin supercoiling, allowing or preventing access to DNA for transcription initiation machinery. *L. pneumophila* secretes several Dot/Icm effectors that contribute to these modifications, such as HPTM, thus triggering epigenetic control of host gene expression, most likely to its benefit (Figure 12).

RomA is a methyltransferase present in the Paris strain that migrated to the host nucleus, where it trimethylates H3 histone on the lysine 14 residue. Methylation prevents subsequent acetylation on this residue. This leads to a repression of host gene expression, mainly innate immune-related genes [246]. During the infection, RomA can also methylate non-histone proteins that contain a G-K-X-(PA) sequence [258], such as AROS (Active regulator of SIRT1), whose participating in ribosome biogenesis helps to maintain protein synthesis at the maximal rate [259].

LegAS4, the RomA homolog in the Philadelphia-1 strain, is also a methyltransferase localized in nucleoli. There, it methylates histone H3 on the lysine 4 residue, leading to ribosomal RNA (rRNA) transcript increase [260]. For an unknown reason, *L. pneumophila* recruits ribosomes on its LCV, and LegAS4 could participate in producing more ribosomes [260]. Surprisingly, LegAS4 and RomA display more than 90% of identity and share a common SET domain, despite their different enzymatic activities and physiological effects [246, 261, 260].

LegK7 mimicks mammalian Hippo kinase MST1 activity, triggering a signaling cascade resulting in degradation of TAZ and YAP1, two transcriptional regulators. Transcriptomic analysis showed that LegK7-mediated phosphorylation alters the transcriptional

profile of macrophages. This contributes to rendering macrophages more permissive to *L. pneumophila*'s intracellular growth [262].

SnpL is an effector localized in the host cell nucleus during infection. Immunoprecipitation experiments followed by mass spectrometry led to the identification of the SnpL eukaryotic partner SUPT5H, a transcription elongation factor. SUPT5H is a part of a protein complex that regulates mRNA (messenger RNA) processing dependent on RNA polymerase activity. SnpL ectopic expression leads to a global host gene expression upregulation and cell death [263].

#### 2.2.6.2 Translation

During *L. pneumophila* infection, inhibition of host protein synthesis is essential to block the innate immune response. To date, six effectors have been described as inhibiting host translation machinery (Figure 12).

Lgt1, Lgt2 and Lgt3 are glycosyltransferases that can glycosylate host eEF1A (eukaryotic translation elongation factor) on the serine 53 residue [264, 172, 265]. Glycosylation on this residue, close to the GTP fixation site, leads to the inhibition of protein synthesis.

Lgt1 and Lgt3 were found in every sequenced strain; however, Lgt2 seems to be restricted to the Philadelphia-1 strain [264]. Lgt1, Lgt2 and Lgt3 are secreted at different infection times: while Lgt3 is secreted during the early step of *Acanthamoeba castellanii* infection, Lgt1 is secreted later [172]. The different secretion timing suggests that each effector has a role dependent on the infection step. Moreover, Lgt1 and Lgt2 are involved in inhibiting ER stress, which is also called the UPR. UPR is activated by the host cell after a stressful event [231].

As Lgt proteins, SidI especially targets eEF1A and eEF1B $\gamma$ , two proteins crucial for eukaryotic translation [243]. SidI expression leads to the inhibition of protein synthesis and cell death. Interestingly, mutations of only two specific residues suppressed SidI toxicity, but those mutations did not suppress the interaction with eEF1A and eEF1B $\gamma$  [243].

As previously described, SidL is an actin polymerization inhibitor, but it can also reduce protein translation by an unknown mechanism *in vitro* [266].

LegK4's 3D structure [267] and biochemistry analysis [268] showed that LegK4 is a serine/threonine kinase. A recent study using the Philadelphia strain showed that LegK4 inhibited global host translation through Hsp70 chaperone family phosphorylation [269]. *In vitro*, LegK4 phosphorylated-Hsp70 showed a reduced ATPase and refolding activity. In transfected cells, LegK4 seemed to impact host global protein synthesis through Hsc70 phosphorylation. The increase of the amount of Hsc70 associated with translating polysomes by LegK4's phosphorylation could explain the inhibition of translational machinery. Moreover, LegK4 ectopic expression in HEK-293T cells showed a significant decrease in the UPR sensor's BiP. The authors suggested that the reduced quantity of BiP leads to inhibition of UPR and cellular stress. This study uncovered Hsp70's role in protein synthesis and its link to the cellular translational machinery [269].

These six effectors strongly impact host global translation and could be related to a decrease of I $\kappa$ B (inhibitor of  $\kappa$ B) production that activates the effector triggered response (ETR) and the NF- $\kappa$ B pathway.

### 2.2.7 Metaeffectors

A metaeffector is a particular category of effector that regulates the activity of other effectors during the infection in a synergetic or antagonistic way. Regulation of other

effector activity or function by metaeffectors is a crucial component of the *L. pneumophila* virulence strategy and to date, a dozen metaeffectors have been discovered [270, 271].

LubX (*Legionella* U-box protein)/LegU2 and GobX (Golgi-localized U-box E3 ligase) both have an E3 ubiquitin ligase activity. LubX acts as a metaeffector regulating the function of other effectors, such as SidH [272]. By its E3 ligase activity, LubX will polyubiquitinate SidH, leading to its degradation by the proteasome [273]. LubX is crucial for the temporal regulation and fine-tunes control of SidH function during infection. LubX can also target cellular host proteins such as CLK1 (Cdc2-like kinase) involved in the RNA splicing and could impact target genes' expression [274].

GobX's role during infection is not yet clear; it possesses an E3 ubiquitin ligase activity mediated by a mammalian U-box domain [275]. It can mainly be localized to the Golgi apparatus by exploiting cellular S-palmitoylation activity, where it could probably ubiquitinate many cellular proteins [275].

LupA (*Legionella* ubiquitin-specific protease A) inactivates the LegC3 effector by removing ubiquitin, which suggests that LegC3 activity could depend on an endogenous E3-ligase [270].

SidP, previously described as a PtdIns(3)P phosphatase, can also inactivate the ATP hydrolysis activity of the MavQ effector whose function is unknown [270].

SidJ is known to be a metaeffector participating in the temporal effector regulation, eliminating SdeA or other SidE family members from the LCV [169, 276].

By its interaction and binding with RavJ, LegL1 blocks and inhibits the RavJ active site [270].

SusF (suppressor of SidI)/MesI (metaeffector of SidI) bind and regulate SidI-mediated translation inhibition by decreasing its GDP-dependent glycosyltransferase activity

[277, 278, 279]. SidI interacts with the host eEF1A and inhibits eukaryotic protein translation by an unknown mechanism, but SusF does not hinder SidI-eEF1A because these proteins bind to distinct regions of SidI [277].



# Chapter 3

## The endoplasmic reticulum and *L. pneumophila*

### 3.1 The endoplasmic reticulum

The ER is the larger organelle in the eukaryotic cell representing around 50% of the total cell membrane. This large and dynamic structure plays many roles, including calcium storage, lipid metabolism, protein synthesis, folding, and transport [280]. It represents the entry point into the secretory pathway where nascent proteins are released in a specialized environment for their folding and maturation. ER quality control (ERQC) is the ER surveillance system that ensures the handling of misfolded or misassembled proteins [281]. Right after their detection in the ER, these proteins are retrotranslocated back into the cytosol to be ubiquitinated and degraded by a proteasome using ERAD (ER-associated degradation) [282] or via ERQC-autophagy for ERAD-resistant proteins [283]. It is known that half of the proteins that pass through the ER failed and are eliminated via ERAD [284].

In mammals, ERAD pathways are highly complex, but in yeast, three ERAD mech-



anisms exist [285]. ERAD-L checks soluble proteins in the ER lumen, ERAD-M checks the transmembrane proteins, and the ERAD-C mechanism verifies cytosolic domains of membrane proteins. If a “defect” of folding is detected, proteins will be translocated to the cytosol, ubiquitinated, and eventually eliminated by the proteasome.

Chaperones present in the ER play a key role in protein folding and also ensure quality control. Several stresses, such as infection by *L. pneumophila*, misfolded protein build-up, increased calcium concentration, or heat shocks lead to a UPR or ERS physiological response. The accumulation of misfolded proteins will be detected by the ER-resident protein, BiP/Grp78. UPR will first slow down global translation through protein interactions and phosphorylation events. Then, UPR will activate the production of chaperones, such as foldase, to drastically reduce the quantity of unfolded proteins. Another mechanism leading to the same result is the identification and degradation of these proteins through the proteasome. If all the above steps failed, UPR would activate apoptosis. The ER is where the calcium concentration is higher in the cell. The typical  $Ca^{2+}$  cytosolic concentration is around 100 nM, while the concentration in the ER lumen is between 100  $\mu$ M–800  $\mu$ M, and the extracellular concentration is approximately 2 mM [286, 280].  $Ca^{2+}$  ions are stringently excluded from cytosol because they can bind water less efficiently than other divalent ions like  $Mg^{2+}$  and precipitate phosphate. Several calcium channels are used to concentrate  $Ca^{2+}$  in the ER. Calcium is a crucial signaling molecule that has an impact on diverse processes such as the association of proteins, organelles or nucleic acids.

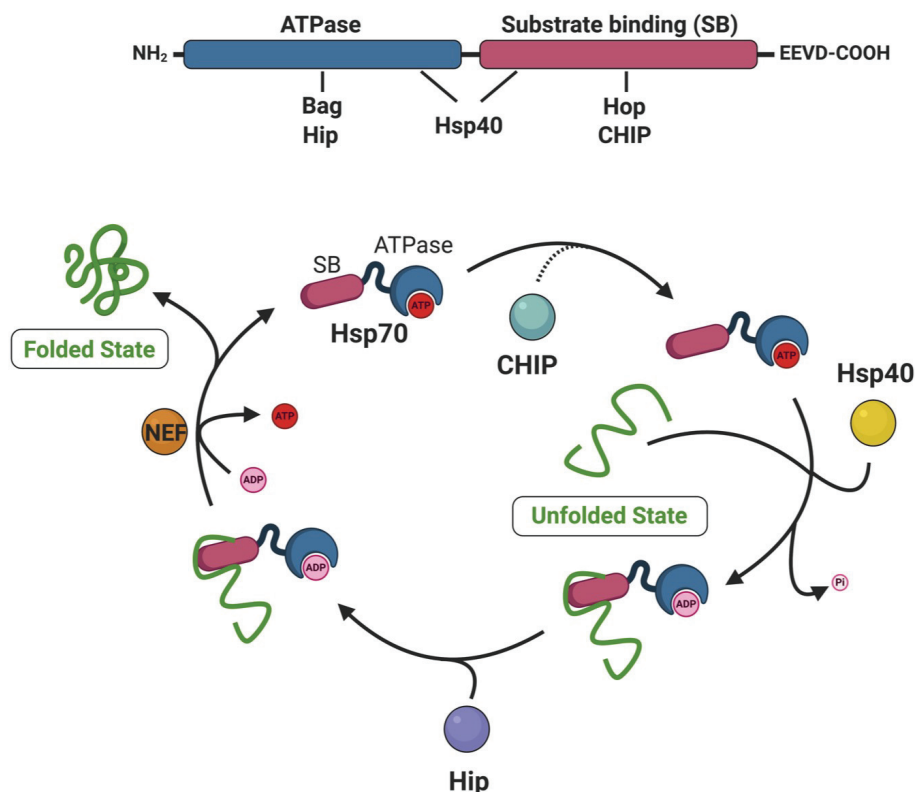
The ER is also the place of lipid synthesis, such as cholesterol biogenesis. The ER lumen is a highly oxidative place enabling the formation of disulfide bonds and other modifications. Numerous chaperones help to fold nascent proteins.

## 3.2 HSP70 family chaperones

When subjected to rapid change in their environment, organisms exhibit homeostatic-like responses. The ability to successfully adapt or acclimate to a new environment is critical to an organism's survival. The most studied response to sudden adverse environmental changes is the so-called heat shock or stress response. When confronted with relevant temperature increases, cells respond by quickly increasing the synthesis of a particular type of protein, the heat shock proteins (HSPs). Despite their designation as HSPs, studies showed almost all these proteins are, in fact, produced in unstressed cells (constitutive) and after metabolic insults (induced). HSPs are molecular chaperones that facilitate the early stages of folding and assembly of proteins, stabilize maturing polypeptides, and reduce incorrect folding or aggregation (Figure 13) [287]. They are traditionally divided into "families" based upon their molecular weight and sequence conservation [288].

The HSP70 family members or Hsc70s are constitutively expressed chaperones (Annexe Table 2). Present in organisms from archaeobacteria to mammals, proteins of the HSP70 family are among the most conserved proteins during evolution [289, 290]. Interestingly, HSP70 proteins from different organisms can share low identity but have the same functions.

As an example, Hsp72 and its prokaryotic analog DnaK share around 50% amino acid identity and possess the same activity. The ectopic expression of Hsp70 protein from *Drosophila* in human cells can protect those cells from various stresses. Interestingly, compared to most prokaryotes, all eukaryotes possess multiple genes encoding HSP70.



**Figure 13: HSP70s folding cycle (This study)**

Through cycles of substrates binding and release, HSP70s guarantee the proper folding of proteins via ATP binding and hydrolysis. HSP70s are made of two domains: an N-terminal ATPase domain and a C-terminal substrate-binding (SB) domain. Numerous cochaperones, such as Hsp40, are crucial components and help HSP70s in the folding process. HSP40 binds straight to HSP70s and stimulates its ATPase activity, stabilizing the HSP70s-substrate interaction. Other cochaperones such as Hip can modulate the HSP70s reaction cycle by binding to its ATPase domain, stabilizing the ADP-bound state and leading to better stabilization of HSP70 with the substrate. Nucleotide exchange factors (NEF), such as Bag-1, can also bind to the HSP70s ATPase domain, releasing bound ADP. HSP70 ATP in the bound state releases a properly folded substrate, completing the cycle. In the case of protein aggregation and misfolding, the cochaperone CHIP promotes ubiquitination leading to the proteasomal degradation of failed proteins.

In a cell, deregulation of chaperone balance or elevated levels have been linked with various cancers and diseases [291, 292]. Interestingly, the deletion of one HSP70 is compensated for by the high number of cytosolic HSP70 proteins, suggesting these

proteins present both overlapping and divergent functions. Every HSP70s possesses two essentially different domains, the nucleotide-binding domain (NBD) and the substrate-binding domain (SBD).

### 3.2.1 Hsc70

Hsc70 (heat shock cognate 71 kDa protein)/HSPA8 is a member of the heat shock protein 70 family (HSP70). Unlike some other chaperones, Hsc70 is a constitutively expressed chaperone protein involved in diverse cellular processes from protein folding to protein degradation [293]. With the help of other proteins, Hsc70 binds in an ATP-dependent manner to short hydrophobic stretches of nascent or unfolded polypeptides via its SBD. ATP hydrolysis leads its conformational change to induce protein folding and then release the substrate. An exchange of ADP for ATP with the help of nucleotide exchange factors allows Hsc70 to bind to new substrates [293]. Hsc70 also plays a role in substrate transport from the nucleus to the cytoplasm by shuttling across the nuclear membrane. Hsc70 helps in the export and recycling of import receptors in an ATP-dependent manner [294].

Hsc70 can recognize misfolded or denatured substrates during chaperone-mediated autophagy and targets them to the lysosome with the help of various chaperones. Then, substrates are translocated across the lysosomal membrane for degradation [295].

Finally, Hsc70 can also play a role in the disassembly of clathrin-coated vesicles with the help of auxilin during transport of membrane components through the cell [296].

### 3.2.2 Hsp70

In addition to its role as a molecular chaperone, Hsp70 protects cells from a broad range of apoptosis-inducing stresses [297]. Interestingly, Hsp70 inhibits several intrinsic apoptotic pathways, such as the apoptosome formation or the activation of stress-induced kinase ASK1, JNK, or P38 [298]. Likewise, Hsp70 may also boost the RNA activity of IRE1 $\alpha$ , enhancing IRE1 $\alpha$ /XBP1, showing a physical interaction between cytosolic chaperones and UPR [297].

Hsp70 is also a chaperokine and describes the dual role of an extracellular heat shock protein as both a chaperone and a cytokine [299]. Hsp70 can be found out of the cell, where it is called eHsp70 and may be the result of a passive (cell death) or active (exocytosis) release. Extracellular Hsp70 binds with high affinity to a broad range of receptors present on the plasma membrane of antigen-presenting cells, such as TLR2 (toll-like receptor 2), TLR4, CD40, or CD36 [300]. The binding between Hsp70 and the receptor elicits a rapid calcium ion flux, activating NF- $\kappa$ B, leading to an upregulation of the pro-inflammatory cytokine expression, such as TNF- $\alpha$ , interleukin-1 beta (IL-1 $\beta$ ) and IL-6 [301].

### 3.2.3 Grp75

The human mitochondrial Hsp70/Grp75, also called mortalin, is a constitutively expressed chaperone mainly present in the mitochondria [302]. In the mitochondrial matrix, Grp75 plays a crucial role in importing mitochondrial proteins through the mitochondrial membrane, oxidative stress response, regulation of mitochondrial membrane potential, energy generation, immune response, and protection against apoptosis [303, 304].

### 3.2.4 BiP

In mammalian cells, BiP is an abundant protein that exists in monomeric and oligomeric forms. However, its expression is overstimulated under stressful conditions, such as in the presence of an accumulation of unfolded polypeptides in the ER [305]. Like many members of the HSP70 family, BiP is an ATP-dependent chaperone and is assisted with its task by cochaperones like ERdj5 (ER-localized DnaJ family members), Grp94 from the HSP90 family or calnexin [306, 307]. In the ER lumen, BiP acts as a general chaperone thanks to its ability to recognize a wide variety of polypeptides [308]. BiP's substrates share no obvious sequence similarity, but they can efficiently discriminate properly folded proteins from the unfolded structure.

In addition to its chaperone activity, BiP acts as the major monitor and transducer of ERS [309]. In non-stressed cells, BiP binds and inhibits PERK, IRE1, and AFT6 from activating the UPR. In the case of ERS, BiP will bind preferentially to unfolded or misfolded proteins leading to UPR activation and biological changes to restore homeostasis.

BiP can be post-translationally modified by phosphorylation and by ADP ribosylation [310, 311]. All these modifications may be crucial in the regulation of the synthesis and polypeptide binding activity. Inhibition of protein synthesis mediated by cycloheximide or rapamycin induces BiP phosphorylation [312], but unfortunately, little is known about how these modifications affect the basal BiP's ATPase activity or its propensity to oligomerize. However, ERS causes an increase in the number of monomeric species and a decrease in the extent of BiP modification. Only monomeric and unmodified BiP molecules are found in interaction with unfolded or misfolded polypeptides [313]. Modification of BiP seems to provide a storage pool of BiP that can be recruited back to

the “active form” in the case of ERS.

### 3.3 Unfolded Protein Response

The unfolded protein response (UPR) is a master regulator in many inflammatory diseases and is an essential adaptive response. This mechanism provides protein translation and folding homeostasis, regulates immune responses to various stimuli such as pathogens, and eventually serves as a decision point of fight-or-die response [314]. Manipulation of ER chaperone activity to control the protein folding and avoid protein aggregation has been used by many pathogens, including *Legionella*, to prevent UPR activation.

To create its replicative niche, *Legionella* picks up ER membranes and mitochondria and these stressful events for the host should, in theory, activate the UPR, but it is not the case [231]. While UPR has a critical role in various molecular events, its purpose and how its dysregulation contributes to the *Legionella* pathogenesis remains unclear. However, a recent study suggested *L. pneumophila* could, thanks to LegK4, target BiP and downregulate UPR activity and ERS [269].

#### 3.3.1 UPR signaling pathways

The ER is in charge and ensures the release of adequately folded proteins, while misfolded proteins are degraded through ER-associated degradation (ERAD) or autophagy [280]. ER tasks are primary the biosynthesis, trafficking, and posttranslational modification of secreted and transmembrane proteins. Upon *Legionella* infection by ER recruitment on the LCV, its homeostasis can be disturbed, resulting in significant stress

that could lead to a series of adaptive mechanisms termed the UPR, which restores protein folding homeostasis [231].

In the ER lumen, three transmembrane proteins exist to detect stressful events: the protein kinase RNA-like endoplasmic reticulum kinase (PERK), the activating transcription factor 6 $\alpha$  (ATF6 $\alpha$ ), and the inositol-requiring protein 1 $\alpha$  (IRE1 $\alpha$ ). All these proteins physically interact with the chaperone immunoglobulin binding protein (BiP)/Grp78/HSPA5 and act as sensors of ERS. The UPR is triggered right after ERS is detected to restore protein folding homeostasis by monitoring protein translation, increasing folding capacity, or eventually activating ERAD [315]. In the worst-case scenario, apoptosis could be triggered to remove stressed cells and avoid the harmful effects of misfolded proteins (Figure 14).

### 3.3.1.1 UPR activation

It was originally believed that during ERS, the BiP dissociation from ER transmembrane sensors initiates the UPR (Figure 14). Nonetheless, BiP seems to be not the only chaperone starting the UPR for all ER transmembrane sensors, and unfolded proteins can bind straight to either IRE1 $\alpha$  or PERK [316]. The interaction between ER transmembrane sensors and peptides from unfolded protein luminal domains induces conformational changes initiating the UPR. Finally, a third model of UPR activation exists and proposes that both BiP dissociation and binding of unfolded proteins cause the UPR activation [317]. This model seems to be a more reliable and efficient way of handling a broad range of unfolded/misfolded proteins. The model suggests two different pathways for UPR exist: a translational pathway involving PERK and a transcriptional pathway that involves ATF6 and IRE1 $\alpha$ .



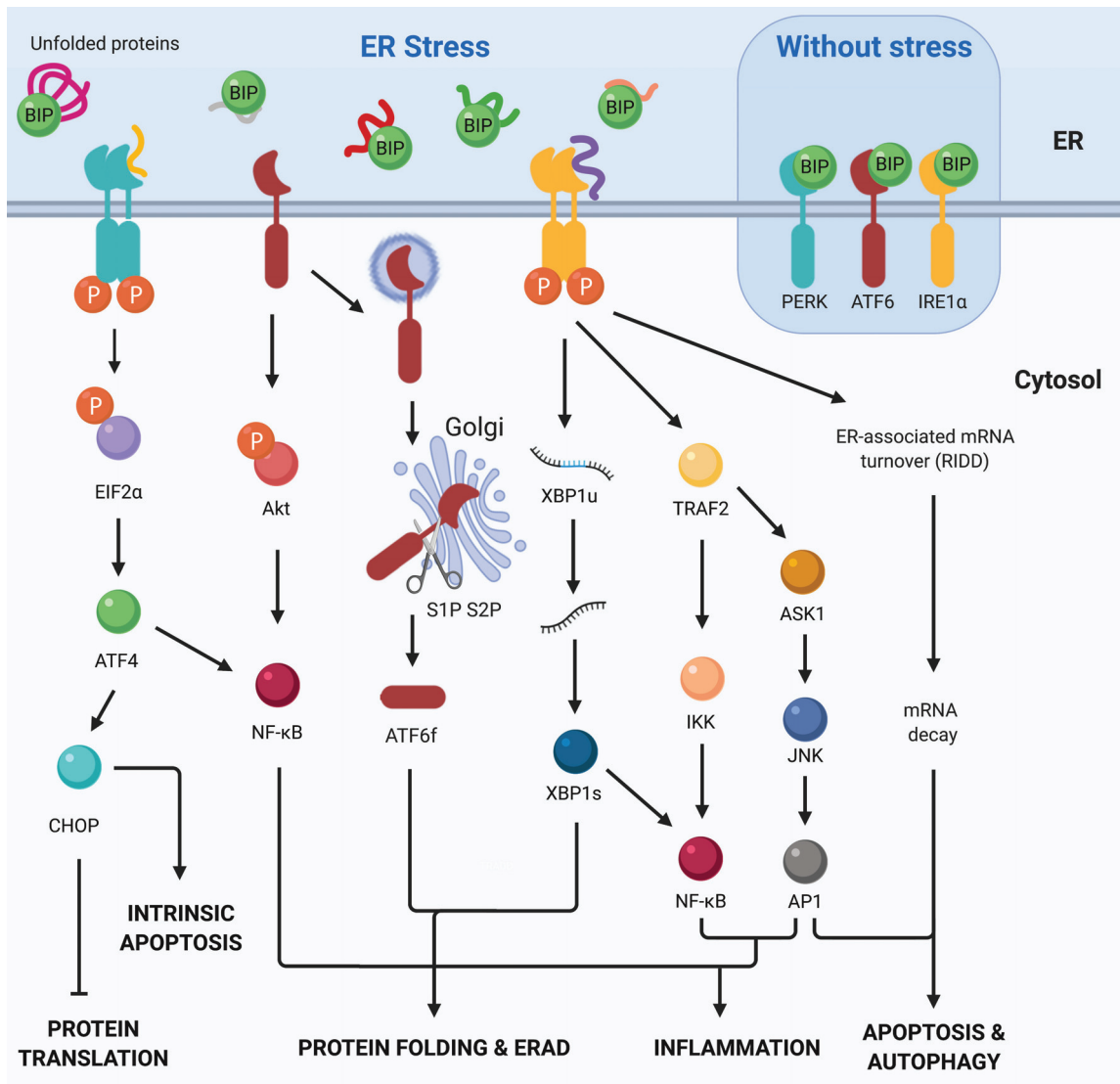


Figure 14: Unfolded protein response (UPR) signaling pathway (This study)

. Upon ER stress, such as a misfolded protein accumulation, ER-resident chaperon BiP dissociates from the ER membrane sensor to bind with misfolded or unfolded proteins present in the ER. BiP dissociation activates UPR to correct the protein folding or lead to cell death. Activation of the UPR is mainly due to three primary ER sensors: protein kinase RNA-like endoplasmic reticulum kinase (PERK), activating transcription factor 6 $\alpha$  (ATF6 $\alpha$ ), and inositol-requiring protein 1 $\alpha$  (IRE $\alpha$ ). The UPR could attenuate protein translation, increase the misfolded protein degradation (ER-associated degradation), increase protein folding, induce various inflammatory responses, and if the problem remains unresolved, the UPR could trigger autophagy and cell apoptosis with as an example the activation of the intrinsic apoptotic pathway by CHOP.

- **The translational pathway/PERK Signaling**

PERK is a type I ER transmembrane protein whose activation is dependent on the recognition of misfolded proteins inside the ER. PERK activation leads in its oligomerization and autophosphorylation, triggering phosphorylation of eukaryotic initiation factor  $2\alpha$  (eIF2 $\alpha$ ), leading to an attenuation of protein translation by controlling the protein load of the ER [318] (Figure 14). eIF2 $\alpha$  also activates eIF2 $\alpha$ -activating transcription factor 4 (ATF4) translation that, in turn, activates several UPR-related gene transcriptions, which include ERAD leading to apoptosis, autophagy and redox homeostasis [319].

- **The transcriptional pathway**

- **ATF6 signaling**

ATF6 is a type II ER transmembrane leucine zipper protein. Inside the ER, the accumulation of misfolded proteins leads to BiP dissociation from ATF6, allowing it to interact with the misfolded proteins [320]. Then, ATF6 translocates consequently into the Golgi apparatus and is cleaved by S1P and S2P proteases, leading to the release of its N-terminal domain. The latter can migrate to the nucleus to induce UPR-related genes transcription, which includes acute phase response (APR)-associated genes [321] (Figure 14).

- **IRE1 $\alpha$  signaling**

The third arm of the UPR pathway, IRE1, is a type I ER transmembrane protein that is activated by the dissociation of BiP from its luminal domain. Accumulation of misfolded proteins in the ER results in IRE1 $\alpha$  oligomerization, kinase domain autophosphorylation, and activation of the RNase domain (Figure 14). The latter induces

splicing of X-box binding protein 1 (XBP1) mRNA to express transcriptional factors and upregulates UPR-related genes to protein folding, protein secretion, and ERAD [320]. In the case of chronic persistence of the stress, activated IRE1 will interact with tumor necrosis factor receptor-associated factor 2 (TRAF2). The latter will lead to the activation of the apoptosis signal-regulating kinase 1 (ASK1)-c-Jun N-terminal kinase (JNK) signaling pathway, triggering inflammation or apoptosis [322].

### 3.3.2 *Legionella* and UPR

Bacterial pathogens are well-known for hijacking many cellular pathways to survive and replicate. Depending on the pathogen or under certain conditions, the UPR will be inhibited or activated to the benefit of the bacteria [323]. The UPR is initiated by the host cell to fight against ERS, which can be induced by the accumulation of abnormal proteins. Activation of the UPR can have many consequences, including the initiation of the innate immune response, apoptosis, proinflammatory programs, the inhibition of global protein synthesis, and upregulation of ERS proteins protein such as BiP and CHOP [324]. The consequences of activating the UPR include the inhibition of global protein synthesis, upregulation of the production of ER stress proteins (e.g., luminal chaperone BiP), and the initiation of apoptotic (e.g., through induction of CHOP) and proinflammatory (e.g., activation of NF- $\kappa$ B) programs.

Some pathogens may manipulate the UPR as a requirement for establishing an intracellular niche. Some might exploit the UPR to take advantage of its particular function (i.e., lipid biosynthesis, increasing of protein folding capacity and release of amino acids after ERAD). In contrast, others block specific pathways among UPR to prevent host defense activation, such as innate immunity or apoptosis.

During the infection, *L. pneumophila* creates a replicative vacuole using ER-like vesi-

cles that is a very stressful event for the host cell. To cope with ERS and initiation of the UPR, *Legionella* uses different proteins to inhibit UPR by targeting different UPR pathways.

Lgt1 and Lgt2, previously described as inhibiting global protein translation, are also two effectors inhibiting the UPR by targeting IRE1 [231]. A wild strain of *Legionella* reduces the amount of spliced XBP1 by IRE1, causing a lack of ER-related transcription factor, inhibiting the UPR.

Genomic deletion of the five proteins involved in host translation inhibition (Lgt1, Lgt2, Lgt3, SidI and SidL) or for the T4SS restored host-cell-splicing activity will affect UPR activity [242].

Moreover, the wild phenotype can be restored in these mutants by *lgt2* or *lgt3* plasmidic complementation.

In addition to phosphorylating Hsp70s family-related proteins and decreasing cytosolic ER sensor BiP, LegK4 was also described as inhibiting the UPR [269].



# Chapter 4

## New Target of *Legionella*: The Nucleus

The nucleus is the largest organelle inside eukaryotic cells and represents about a tenth of the entire cell volume. It contains genetic material and serves to maintain cellular integrity.

The nucleus is an essential organelle controlling many vital cellular processes, such as transcription, DNA replication, division, and growth. It stores all the genetic information of the cell and is responsible for gene expression and cell integrity. Having a central role in eukaryotic cells, the nucleus is the new target of bacterial protein, called nucleomodulin. These proteins usually change the host transcription gene primarily by epigenetic control.

### 4.1 Levels of epigenetic control

William Bateson defined genetics as the science of hereditary trait transmission in 1905 [325]. In comparison, the term “epigenetics” was first used in 1942 by the British

embryologist Conrad Hal Waddington to define what is literally “above genetics” [326]. After the discovery of the DNA double helix structure and genetic code, it was thought that all the information required for life, whatever species, was included in DNA sequence and, more precisely, in genes.

Later in the late 70s, it became clearer that genes alone could not explain everything. In 1990, epigenetics was characterized by Holliday as “the study of the mechanisms that impart temporal and spatial control on the activities of all those genes required for the development of a complex organism” [327]. The term “epigenetic” is now usually used to describe how other DNA sequences influence gene expression.

To date, three levels of epigenetic control that understood as being correlated to DNA level organization among chromosomes: DNA methylation and RNA and histones modification .

The human genome is composed of 23 chromosome pairs present in every cell. *In vivo*, the double helix DNA structure is finely organized, compacted in the nucleus and represents more than 10% of cell space. To fit in, DNA must be supercoiled. Supercoiled DNA is associated with essential proteins called histones, leading to chromatin formation.

### 4.1.1 DNA methylation

DNA methylation was the first epigenetic mark discovered to have an impact on gene expression. This modification is the reversible addition of a methyl ( $CH_3$ ) group to DNA. In the 70s, DNA methylation was proposed as a mark that could regulate gene expression and, in particular, the female X-chromosome inactivation [328]. Because it prevents translocations of repeated or transposable sequences, this modification contributes to genome stability and seems to play a role in development [329]. Interestingly,

it does not alter the DNA nucleotide sequence and cannot be detected by conventional sequencing techniques. However, bisulfite/bisulphite sequencing [330] or methylation-sensitive restriction endonucleases allows determine of DNA methylation patterns [331]. In vertebrates, DNA methylation usually occurs on cytosine, giving a 5-methyl-cytosine (5mC) within a cytosine-guanine dinucleotide (CpG). Methylation could occur elsewhere in plants and embryonic stem cells [332]. There are around 28 million CpG in the human genome, where over 60% are always methylated. These methylations mostly take place in gene regulatory sequences or in repeated sequences. Over 50% of genes possess dense CpG stretches called CpG island (CGI) in their promoter regions. DNA methylation's role in these regions is considered a stable epigenetic mark of suppressed gene expression. Interestingly, these modifications are maintained through mitosis and meiosis [333]. In 2008, Metivier [334] reported that promoters may be actively methylated and demethylated using as a model trefoil factor 1 (TFF1) from breast cancer cells. Many methyltransferases, such as MeCP2, SWI/SNF, DNMT1 and DNMT3a/b, are recruited to the TFF1 promoter, resulting in its methylation and, from there, to its inhibition [335, 336]. DNA methylation has two impacts: preventing DNA recognition from proteins such as transcription factors that are DNA-state sensitive or promoting methyl-binding proteins (MBP) that will inhibit gene transcription. This reversible modification, which has a substantial impact on gene expression, is finely controlled by cells. Indeed, it has been proposed there is a balance between DNA methylation from DNA methyltransferase and demethylation. Surprisingly, demethylation can be done passively during DNA replication or by spontaneous or enzymatic transformation of 5mC into uracil. The latter is a molecule not found in DNA that will be recognized by the repair enzymes system as a DNA sequence error, modifying uracil into cytosine. Experiments to disrupt methylation have shown that this epigenetic mark is essential during the embryonic development of the vertebrates such as zebrafish. In mouse em-



bryos, knock-out of DNMT3a/b methyltransferase leads to a premature death [337].

### 4.1.2 Non-coding RNA

The second epigenetic mechanism uses the interaction properties of non-coding RNAs (ncRNAs). High-throughput sequencing approaches show that while 90% of our genome's DNA is transcribed into RNAs, less than 5% of the transcribed mRNAs are translated into proteins [338]. The ncRNA family consists of molecules with very different functions, such as rRNA and tRNA (transfer RNA), essential for translating mRNA into proteins and regulatory RNAs. The small ncRNAs, about 20 nucleotides long, and the miRNAs (micro RNA) act on mRNAs to destabilize or inhibit them and thus control the level of a protein in the cell. Other ncRNAs act directly on the DNA, targeting regions of the genome of the same sequence and creating an environment that suppresses gene expression [339]. Finally, much larger ncRNAs, consisting of thousands of nucleotides, such as Xist (X inactive specific transcript) RNA, cover the DNA entirely and inactivate the activity of one of the two X chromosomes of female mammals [340].

### 4.1.3 Histone modifications

Histones are nucleoproteins only found in eukaryotic cells. Mainly found in the cell nucleus, they have an important role in DNA stabilization and packing, enabling it to order into nucleosomes. Histone post-translational modifications (PTM), also called histone modifications, are covalent modifications well known to impact chromatin supercoiling levels. The main changes in histones are acetylation, methylation, ubiquitinylation and phosphorylation. These modifications affect various cellular mechanisms such

as replication, transcription or DNA repair process but can also affect gene expression by altering chromatin structure or by recruiting histone modifier enzymes, responsible for the addition or removal of these groups. Their activity changes the chemical nature of histones, interactions with their partners, their position, and the composition of nucleosides. These enzymes, also referred to as “writers,” are therefore involved in the remodeling of chromatin, and recruited proteins are referred to as “readers” [341]. In 2001, a “histone code” was proposed, and genomic mapping of the different chromatin states allows the prediction of whether a gene is active or not based on changes in the histones present in its regulatory sequences [342]. Among all the histones, modification of histone H3 is the best described. As an example, trimethylation of histone H3 lysine 9 residue is associated with DNA compaction and transcription suppression [343]. In contrast, trimethylation of the same histone H3 on its lysine 4 residue is related to transcription activation [344].

#### 4.1.4 Histone H3

Histone H3 is a basic 15 kDa protein that exists in humans as six different variants. They can be part of a conventional nucleosome, such as H3.3 [345], incorporated into centromeric nucleosomes, such as CENPA [346] or found at DNA damage sites like ASF1A [347]. Two are well-described and present in conventional nucleosomes, H3.1 and H3.3, but are components of different histone assembly complexes [345, 348]. H3.1 and H3.3 are remarkably similar in amino acid sequence but different in their expression patterns or modes of deposition [349]. H3.1 is replication-dependent because it is expressed during DNA synthesis, and H3.3 is replication-independent and expressed throughout the cell cycle.

- **H3 Acetylation/Deacetylation**

Hyperacetylation of histone H3 is generally associated with chromatin decondensation, allowing DNA accessibility to proteins increasing and increasing the area's transcriptional activity [350, 351]. For instance, acetylation on lysine 16 (K16ac) was proved to hinder the formation of 30 nm chromatin fiber [352], and K9ac/K14ac were frequently found in the transcriptional permissive region [353]. As expected, histone hypoacetylation is associated with chromatin condensation and, therefore, on transcriptional repression [354]. One mechanism to increase H3 acetylation is to inhibit histone deacetylases (HDACs), allowing histone acetyltransferase (HAT) to modify histone. Together, HATs and HDACs control the delicate balance on the H3 acetylation level, playing an essential role in transcription activation or repression.

- **H3 Methylation/Demethylation**

The methylation of lysine 9 from H3 (H3K9me) plays a significant role in transcriptional silencing. Proteins from the SUV39H1 family are histone methyltransferases specific to H3K9me and are required both in *Drosophila* and *Schizosaccharomyces pombe* for gene silencing near the centromere [355]. In addition, H3K9me is present in transcriptional inactive areas, such as pericentric heterochromatin, sexual-type silent locus in *S. pombe* [343] or the silent region adjacent to the chicken's  $\beta$ -globin locus [356]. On the contrary, H3K4me, H3K4me<sub>2</sub>, H3K4me<sub>3</sub> are associated with transcription activation [344]. Indeed, methyltransferases responsible for H3K4me are usually transcriptional activators, such as SET1 or ASH1 [357], or are associated with co-activators [358]. Unfortunately, the H3K4me molecular consequence is not yet known well but is believed to promote the recruitment of activating factors or, in contrast, inhibit the interaction of a repressive complex [359]. At euchromatic sites, H3K9me, H3K9me<sub>2</sub>, H3K9me<sub>3</sub>,

H3K27me and H3K27me3 mediate heterochromatin formation participating in silencing gene expression; therefore, they constitute repressive marks [360].

- **H3 Ubiquitination**

Histone mono-ubiquitination plays a crucial role in DNA damage repair (DDR) [361, 362]. Knock-out of CUL4 from the ubiquitin ligase complex CUL4-DDB-ROC1 weakens the interaction between DNA and histone, facilitating the recruitment of repair proteins to the DNA damaged site [363]. H3 ubiquitination also plays a role in tumor formation and glycolysis because H3K23/K36/37Ub mediated by NEDD4 is linked to transcriptional activation of tumor-related genes IL1 $\alpha$ , IL1 $\beta$  and GGLM [364].

- **H3 Phosphorylation**

Upon activation of many cellular pathways, such as ERK or p38 MAPK, H3 phosphorylation plays the role of the intermediate step in the chain of events leading to the target gene induction [365]. Moreover, it is now suggested that H3 phosphorylation is a crucial element linking the MAPK signaling cascade with chromatin remodeling. Stimulation of MAPK by EGF results in elevated levels of H3S10ph, which are found by CHIP-seq in mouse fibroblasts, linked with promoter regions of immediate early genes, such as c-jun, c-fos and c-myc [366]. Moreover, H3S10ph caused changes in HP1 $\gamma$  displacement and Brg1 recruitment, as well as RNA polymerase II [367].

## 4.2 The new era of nucleomodulins

Bacterial proteins can access the eukaryotic cell cytoplasm in different ways: (i) injection by a type III or IV secretion system; (ii) self-transport, like toxins; (iii) secretion by free bacteria into the cytoplasm. Nucleomodulins are secreted bacterial proteins that have an impact on the host cell nucleus [368, 369]. The nucleomodulins' effect can be indirect (through a chain reaction starting in the cytoplasm but eventually having an impact on nucleus activity) or direct (if the nucleomodulin can migrate to the nucleus). We are going to focus on the second possibility in this thesis. The nucleocytoplasmic transport of nucleomodulin is carried out using the host import machinery system [368]. Proteins must cross through the nuclear membrane double lipid layer via a nuclear pore to access the nucleus. Nuclear import of nucleomodulin is subjected to the same constraint as eukaryotic protein nuclear import.

### 4.2.1 Protein nuclear import

#### 4.2.1.1 Nuclear pore structure

The migration of large proteins from the cytoplasm to the nucleus is tightly controlled by nuclear pore complexes (NPCs). The number of NPCs differs from one species to another: around 200 NPCs per nucleus in yeast cells compared to around  $5 \times 10^3$  to  $5 \times 10^7$  NPCs per nucleus in human cells [370]. This structure comprises 50 nucleoporins called Nups, organized in a central channel composed of a central ring, a nuclear basket and cytoplasmic filaments. The central channel comprises nucleoporins that contain an abundant and highly repeated region of phenylalanine (F) and glycine (G). This hydrophobic amino acid region, named FG Nups, leads to a tight, hydrophobic barrier [371]. This channel is made of two rings, one around 150 nm in diameter and

one around 45 nm in diameter, forming a central channel. This particular architecture is complemented by eight cytoplasmic and nucleoplasmic fibrils converging to make a basket-shaped structure called a nuclear basket.

#### 4.2.1.2 Protein import through the nuclear pore

Not all proteins can cross the nuclear membrane because nuclear import is a selective mechanism depending on the molecular weight of the cargo protein. Many studies sought to determine the maximum molecular weight a protein can have to migrate by passive diffusion. Depending on the plasmid construction, the authors concluded that if smaller than 60 kDa, a protein can cross the nuclear membrane without using host machinery [372]. Import of higher-weight proteins requires a specific protein carrier, importins, which recognize specific amino acid sequences in the cargo protein, the nuclear localization signal [373].

#### 4.2.1.3 Key to the nucleus: Nuclear localization signal

To have access to the nucleus, high molecular weight proteins require a specific peptidic sequence called a nuclear localization signal (NLS). This sequence is specifically recognized by the importin system, in particular the importin  $\alpha$  [374]. Many NLSs have been discovered, but unfortunately, there is no true consensus about a specific patterned sequence. However, two types of NLS are well described: classical NLS (cNLS) [374] and non-classical NLS (ncNLS) [375].

How nucleomodulins enter into the nucleus is not clear because they generally lack a classical NLS. However, some possess an ncNLS or can be transported via an unknown host nuclear pathway.

- **cNLS**

The most common and first type of NLS discovered, cNLS, mainly consists of basic residue stretch, like arginine and/or lysine (K and/or R). It is considered monopartite if a basic amino acid sequence is present once, like the NLS from SV40 large T antigen (126 PKKKRRV 132) or bipartite if present twice, like nucleophosmin NLS (155 KRPAATKKAGQAKKKK 170). The gap between the two sequences is set between 10 to 100 amino acids. The NLS is recognized by the adaptor protein importin  $\alpha$ , which directly binds to the NLS, which in turn contains its own bipartite NLS recognized by importin  $\beta$  (karyopherin  $\beta$  or p97) [376].

- **ncNLS**

Interestingly, many nuclear proteins do not possess a cNLS but contain ncNLS, whose import is mediated by the direct binding to a protein from the importin  $\beta$  family. Compared to the cNLS case, these proteins do not depend on importing an  $\alpha$  adapter, but by the straight interaction with ncNLS from the cargo protein and facilitate crossing through the NPC [377]. To date, four classes of ncNLSs exist: the acidic M9 domain of hnRNPA1 [378], the complex signals of U snRNPs [379], the KIPK sequence of the yeast transcription repressor Mat $\alpha$ 2 [380] and proline-tyrosine-NLS (PY-NLS) [381] [382].

Among ncNLSs, proline-tyrosine-NLS or PY-NLS is the best described [381]. The peptidic signal is around 15–100 residues long and usually follows three rules: a structural disorder, overall positive charge, and poorly conserved sequence motifs made of a hydrophobic or basic residue on the N-terminal and a C-terminal R/K/H-X<sub>2</sub>-5P-Y

motif recognized by importin  $\beta 2$  (karyopherin- $\beta 2$  or transportin-1).

- **Case of nucleolar proteins**

Inside the nucleus, nucleoli were discovered by Oscar Miller using electron microscopy. They are crucial places dedicated to ribosome synthesis and assembly [383]. Nucleoli are very dynamic places, resulting from the equilibrium between the level of rDNA expression, rDNA processing and transport of ribosomal proteins (40S and 60S) toward the cytoplasm.

Three hypotheses about how proteins can migrate to the nucleolus exist: the presence of a nucleolar localization sequence (NoLS), a nucleolar detention sequence (NODS) or the interaction with nucleolar proteins. To date, little is known.

- Nucleolar Localization Sequence**

The characterization of a NoLS is still an issue, as most nucleolar localization signals are usually made of a random series of basic residues. Over the years, many motifs mostly composed of arginine and lysine residues have been identified (Annexe Table 3). Some nucleolar proteins are much larger than the size exclusion limit of the NPC and also require an NLS to cross the nuclear pore [384].

- Nucleolar Detention Sequence**

In contrast, a NoDS is better characterized by a consensus sequence consisting of a minimum of one arginine motif RR(I/L) and at least two hydrophobic triplets  $L\psi(V/L)$  ( $\psi$  represents a hydrophobic residue) [385]. An *In silico* analysis and photobleaching experiment led to the discovery of around 20 proteins containing an NoDS called NoDP. NoDPs have various functions in protein folding, proteosomal degradation, ubiquitination, and DNA replication showing they may control many cellular pathways [386, 387].



They are sequestered in the nucleolus and undergo a stimuli [388, 385], demonstrating nucleolar detention as a form of post-translational regulation.

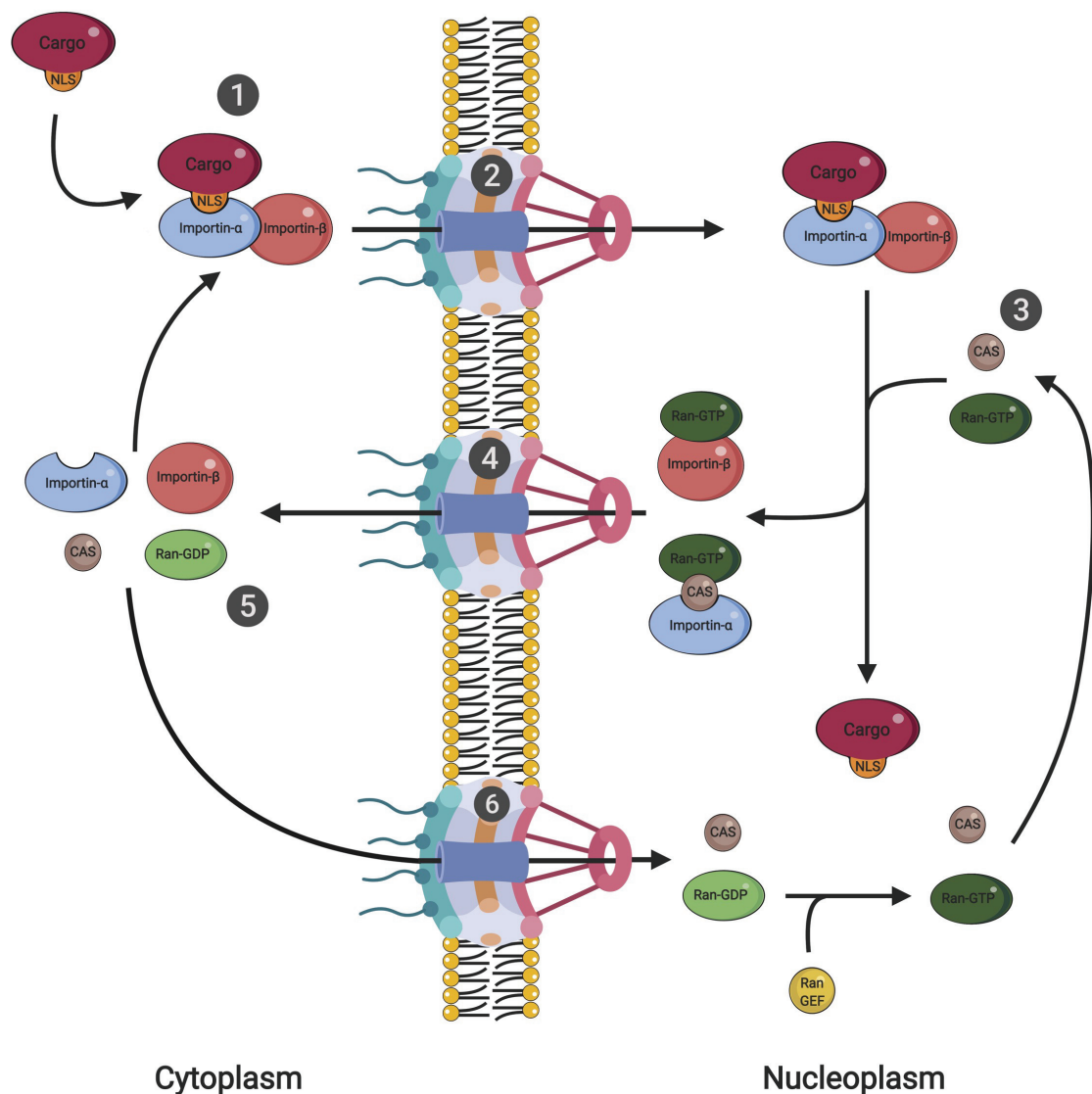
#### **- Interaction with nucleolar protein**

Nucleolar proteins lacking a nucleolar sequence signal (NoLS or NoDS) can also be localized to the nucleolus thanks to their interaction with the protein that contains it. As an example, PP1, a nucleolar protein, does not contain a NoLS/NoDS but is localized in the nucleolus thanks to its interaction with NOM1 and its NoLS [389]. This method of transport is often targeted by viral proteins such as the groundnut rosette virus (GRV) that use host fibrillarin for viral ribonucleoprotein formation [390].

#### **4.2.1.4 Import of the cargo protein mechanism**

Active transport of protein is a finely regulated mechanism mediated by the karyopherin (KAP) family, which includes small protein G Ran and importins. Amongst them, the importin  $\beta$ -1 (karyopherin  $\beta$ -1) binds via an adaptor protein (importin  $\alpha$ ) to the NLS containing the cargo protein. Importins recognize and have a strong affinity to protein cargo NLS. They also possess binding sites to nucleoporins via the FG Nups basic repeats. These interactions lead to the cargo protein transporting from the cytoplasm to the nucleoplasm.

The importin “ $\alpha$ ” is first bound to the cargo protein NLS and acts as an adaptor to the importin  $\beta$  (Figure 15). The complex cargo protein-importin  $\alpha$ -importin  $\beta$  is translocated through nuclear pores due to the importin-pore interaction.



**Figure 15: Nuclear import mechanism and recycling of essential transport components (This study)**

① Formation of the cargo-importin  $\alpha$ -importin  $\beta$  complex in the cytoplasm. ② The complex translocates through the nuclear pore. ③ The complex is dissociated by the CAS exportin and Ran-GTP releasing the cargo protein in the nucleoplasm. ④ Importin  $\alpha$ -CAS-Ran-GTP and importin  $\beta$ -Ran-GTP complexes cross the nuclear pore. ⑤ Hydrolysis of Ran-GTP to Ran-GDP and releasing of importins  $\alpha$  and  $\beta$  in the cytoplasm. ⑥ CAS and Ran-GDP return to the nucleus where Ran GEF (GTP Exchange Factor) activity allows GDP-GTP exchange.

In the nucleoplasm, Ran-GTP and Ran-GTP linked to cellular apoptosis susceptibil-

ity protein (CAS) exportin tapped importin  $\beta$  and  $\alpha$ , respectively. These protein-protein interactions lead to the release of the cargo protein in the nucleoplasm. Both importin  $\beta$ -Ran-GTP and importin  $\alpha$ -CAS-Ran-GTP complexes migrate through the nuclear pore and return to the cytoplasm. Cytoplasmic GAPs (GTPase activating proteins) enhance spontaneous hydrolysis of GTP linked to Ran. The hydrolysis of GTP to GDP changes the molecular conformation of Ran and leads to the release of importins  $\alpha$  and  $\beta$ . Importins are now available and can take charge of another cargo protein.

### 4.2.2 Nucleomodulins

Through co-evolution with their hosts, bacteria have established proteins capable of mimicking the functions of eukaryotic proteins [391]. Although many translocated effectors localize to the *Legionella* replicative vacuole, other effectors can affect remote intracellular sites. Interestingly, some of these eukaryotic-like proteins affect the functions of the host nucleus [246, 263].

Through different molecular mechanisms, the activity of these proteins, called nucleomodulins, allows the bacterium to directly or indirectly control the expression of genes in the host. For intracellular bacteria, the purpose of this control is usually to escape the host defenses while looting its resources. Extracellular bacteria aim to lyse the cell to release nutrients essential for bacterial growth.

Since their discovery about ten years ago, nucleomodulins have been considered major virulence factors for their essential role in the development of the infectious process and are an emerging axis of research.

Immediately after migration to the nucleus, there are mainly two different types of nucleomodulin targets: DNA and proteins such as histones or transcription factors. Moreover, in rare cases, nucleomodulins may target the nucleolar sub-compartment,

targeting the nucleolus.

Three nucleomodulins exist in *L.pneumophila*: RomA, LegAS4 and SnpL.

#### 4.2.2.1 RomA

RomA is a methyltransferase that presents in the Paris strain. After secretion, RomA migrates in the host nucleus, where it trimethylates H3 histone on lysine 14 (K14) residue. Histone post-translational modifications control eukaryotic gene expression and regulate many biological processes, including immunity. Methylation on this residue prevents subsequent post-translational modification, such as acetylation. This methylation leads to global repression of host gene expression, mainly innate immune-related genes [246]. RomA can also methylate other proteins during the infection, but those proteins must contain a G-K-X-(PA) sequence [258], such as AROS (active regulator of SIRT1). AROS participates in ribosome biogenesis and helps to maintain protein synthesis at the maximal rate [259].

#### 4.2.2.2 LegAS4

LegAS4 is the RomA homolog in the Philadelphia-1 strain. It is also a methyltransferase. Unlike RomA, LegAS4 localizes in nucleoli. There, it methylates histone H3 on the lysine 4 (K4) residue, leading to ribosomal RNA (rRNA) transcript increase [260]. For an unknown reason, *L. pneumophila* recruits ribosomes on its LCV and LegAS4, which could participate in producing more ribosomes [260].

### 4.2.2.3 SnpL

During the infection, SnpL localizes in the nucleus of mammalian cells. SnpL ectopic expression leads to a global host gene expression upregulation and cell death [263]. Co-immunoprecipitation assay followed by mass spectrometry led to the identification of SUPT5H, a eukaryotic partner of SnpL. SUPT5H is a transcription elongation factor, part of a protein complex that regulates mRNA processing dependent on RNA polymerase activity.



## Part II

New insights about LegK4 Paris  
from *L. pneumophila*

## INTRODUCTION

*Legionella pneumophila* is the etiological agent of legionellosis, a severe pneumonia in humans.

Pathogenic strains of *Legionella* appear from the environment after intracellular multiplication in amoebae. These bacteria can be disseminated by water aerosols due to human technologies, such as cooling towers. When inhaled, contaminated aerosols lead bacteria into the lungs to the lower respiratory track where alveolar macrophages are present.

Several studies revealed that the Dot/Icm type 4 secretion system (T4SS) and the secretion of approximately 300 proteins, called effectors, are required for *Legionella* to survive lysis [392, 204].

Within environmental phagocytic cells and human macrophages, this secretion of proteins allows *Legionella* to reroute its phagosome and trigger the biogenesis of an LCV.

For many years, significant investigative efforts have aimed to decipher the individual contributions of effectors to the *Legionella* intracellular lifestyle. Due to a high functional redundancy between effectors, single-gene deletions usually result in the absence of phenotype. Unfortunately, to date, only a few T4SS effectors have been functionally characterized (Annexe Table 1).

Among the Dot/Icm effectors secreted by *Legionella*, some display distinctive eukaryotic motifs or domains such as protein kinase domains. *In silico* analysis of effector sequences and *in vitro* phosphorylation assays with purified proteins led to the identification of four secreted kinases. The endemic *Legionella* Paris strain can secrete five kinases, designated LegK1–4 and LegK7 [268] [262].

To date, the roles and functions of LegK1, LegK2, and LegK3 are partly character-



ized, but very little is known about LegK4.

LegK4 is a well-conserved protein found in five *Legionella* strains: Paris (*lpp0267*), Lens (*lpl0283*), Philadelphia (*lpg0208*), Alcoy (*lpa0386*) and Corby (*lpc0283*) (Figure 17).

We mainly studied LegK4 from the Paris strain because it is a recent French strain isolated from a clinical sample [60].

In this thesis, we aimed to decipher the precise role of LegK4 Paris and its phosphorylation activity through three axes.

First, we aimed to localize LegK4 Paris in human infected cells, then to find its eukaryotic interactants and substrates and finally better understand its impact on *Legionella* and on human cells.

# 1 Localization of LegK4

To better understand the role of LegK4 Paris, we first aimed to determine where LegK4 Paris localized in transfected HeLa cells, then when it was secreted in human macrophages and finally where it localizes in infected human cells.

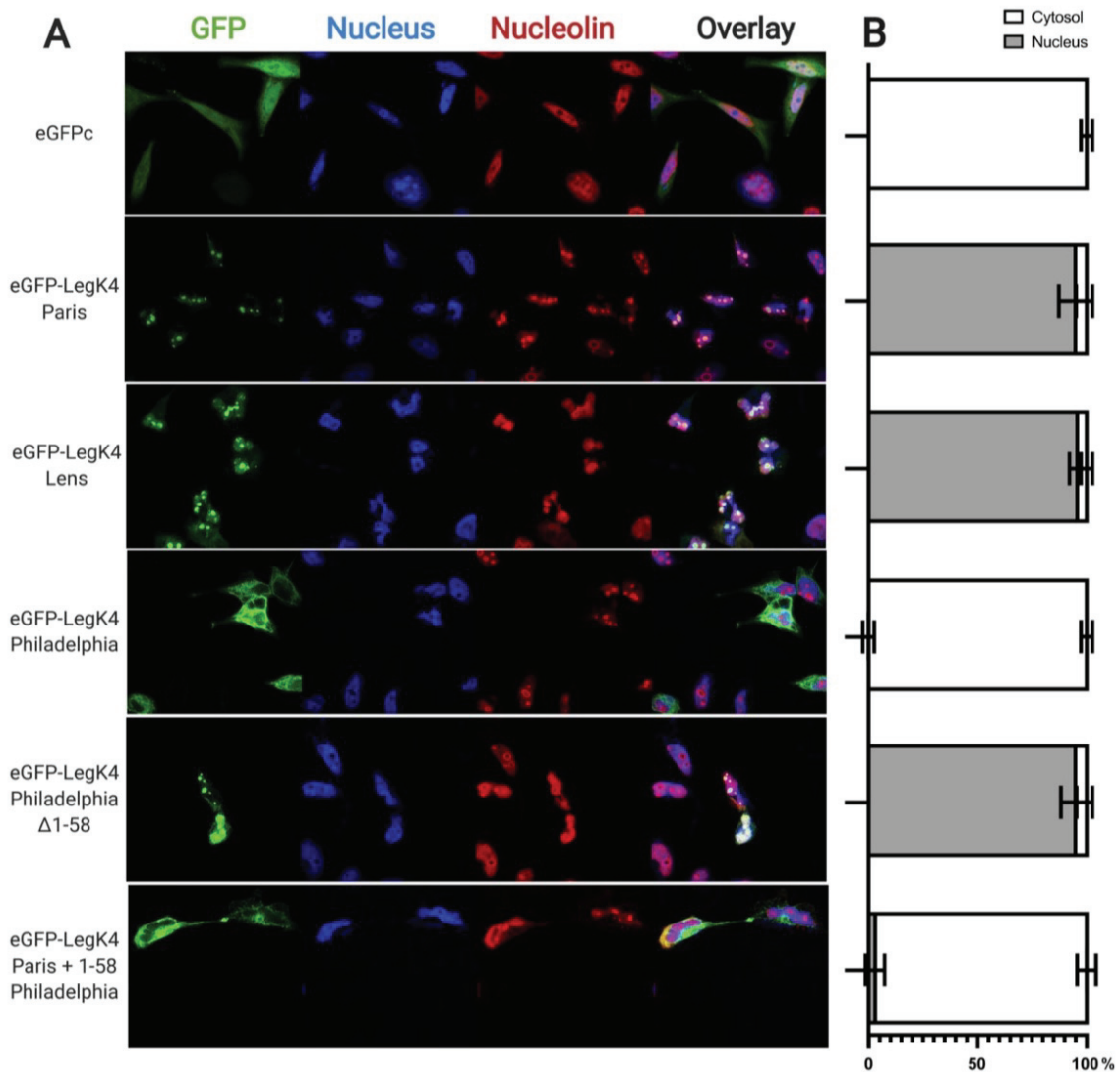
## 1.1 LegK4 Paris localized in the nucleus of transfected HeLa cells

To determine the localization of LegK4 Paris, we have transiently expressed the fusion protein eGFP-LegK4 from Paris or the negative control eGFPc (emerald green fluorescent protein) in HeLa cells for 24 h.

We used the GFP direct fluorescence to detect the protein of interest, and we used 4,6-diamidino-2-phenylindole (DAPI) to stain the nucleus and immunostained nucleolin, a specific nucleolar marker.

As expected in confocal microscopy, eGFPc was localized in the cytosol and the nucleus of transfected cells. However, eGFP-LegK4 Paris was nucleolar in 90% of the cells with a strong colocalization within nucleoli (Figure [16A](#) and B).

In HeLa cells, eGFP-LegK4 Paris localized specifically in nucleoli. However, this result was not consistent with the 2019 study by Moss and collaborators about LegK4 Philadelphia [\[269\]](#). They determined that the localization of LegK4 from the Philadelphia strain was cytosolic but not nucleolar in transfected HeLa cells.



**Figure 16: LegK4 localization depending on *Legionella* strain**

A. HeLa cells were transiently transfected with different eGFP-tagged LegK4 constructs to determine LegK4 localization. Nuclei were stained blue with DAPI, and nucleoli were detected with an anti-C23 (nucleolin) and immunostained using an anti-IgG Dylight 594 antibody (red).

B. The amount of GFP-positive cells in the cytosol or the nucleus was quantified by counting 100 cells per experimental condition. Each condition was done in triplicate. Error bars represent the standard error of the mean (SEM).

## 1.2 LegK4 Philadelphia localized into the cytosol of transfected HeLa cells

Because the result we obtained was different from the result described by Moss *et al.*, we aimed to understand this difference. Therefore, we recreated the same experiment as Moss *et al.*. We made a plasmid to produce the fusion protein eGFP-LegK4 from Philadelphia and identified its localization in transfected HeLa cells.

We used GFP direct fluorescence to detect the protein of interest, and we used DAPI to stain the nucleus and immunostained nucleolin, a specific nucleolar marker.

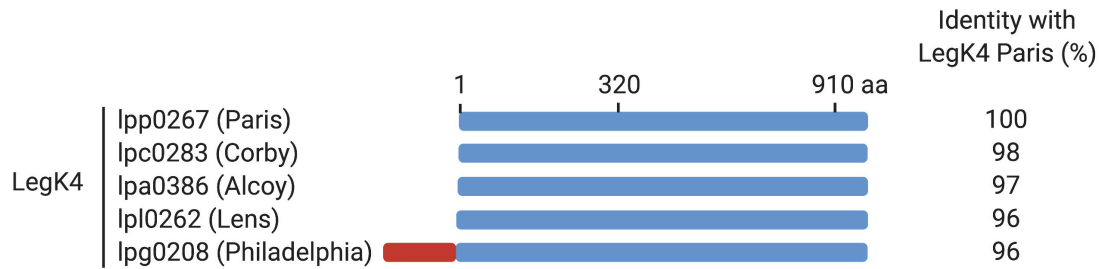
Using confocal microscopy, eGFP-LegK4 Philadelphia was localized in 99% of transfected cells in the cytosol (Figure [16A](#) and B).

This last result was consistent with the results obtained by Moss *et al.*, although they used an anti-GFP antibody to detect eGFP-LegK4 Philadelphia instead of the GFP direct fluorescence [\[269\]](#).

We considered the amino acid sequence of the two proteins to understand this difference of localization of LegK4 between the Paris and Philadelphia strains of *L. pneumophila*. For this purpose, we compared the protein sequence of all LegK4s found in the National Center for Biotechnology Information (NCBI) database.

The sequence of five LegK4 proteins, including LegK4 Paris and LegK4 from other strains, shared a high identity level between 96% to 98% (Figure [17](#)). However, LegK4 Philadelphia is the only one to possess a 58-amino-acid extension in the N-terminal region. Moreover, LegK4 Philadelphia possesses a “TTG” start codon compared to the usual “ATG” start codon found in other LegK4 sequences.

This extension could be responsible for LegK4 Philadelphia’s localization in the cytosol.



**Figure 17: Identity level of LegK4 between the different strains**

LegK4 is present in five different strain of *L. pneumophila* and a high identity level between the sequences is observed. In the Philadelphia strain, the red square represents a 58-amino-acid extension found only in this strain (aa: amino acid position).

To understand the impact of this extension on LegK4 Philadelphia localization, we expressed eGFP-LegK4 Philadelphia without the extension (eGFP-LegK4 Philadelphia  $\Delta$ 1–58) and added the additional sequence to LegK4 Paris (eGFP-LegK4 Paris + 1–58 Philadelphia). Then, we identified the localization of these protein fusions in transfected HeLa cells by confocal microscopy.

In HeLa cells, eGFP-LegK4 Philadelphia without its first 58 amino acids (eGFP-LegK4 Philadelphia  $\Delta$ 1–58) localized like eGFP-LegK4 Paris in the nucleolus (Figure 16A). Furthermore, the addition of these 58 amino acids to the eGFP-LegK4 Paris sequence led to a cytoplasmic localization in HeLa cells like LegK4 Philadelphia (Figure 16A). Moreover, we also identified the localization of LegK4 Lens that does not possess the extension found in LegK4 Philadelphia. Interestingly, this protein localized like LegK4 Paris in the nucleolus, as did LegK4 Paris.

The results we obtained suggested that the difference of localization between LegK4 Philadelphia and the other LegK4 proteins could be due to the additional 1–58 residues present in the N-terminal region of the protein.

### 1.3 LegK4 Philadelphia is wrongly annotated in databases

To better understand the role of this additional sequence in LegK4 Philadelphia, we considered the sequence of the protein for the first time. We verified whether this extension of the protein was due to a mistake of annotation. Indeed, LegK4's genomic organization in Philadelphia and in the other strains was quite different (Figure 18A). Unlike *legK4* Paris, *legK4* Philadelphia started with a TTG start codon.

To determine whether the TTG was the effective start codon of *legK4* Philadelphia, we constructed two plasmids: one containing the first 112 amino acids of LegK4 Philadelphia fused with superfolder-GFP (*p<sub>legK4</sub>-ATG-sfGFP*), the other harboring a mutation on the ATG (*p<sub>legK4</sub>-GCG-sfGFP*) (Figure 18B).

In both plasmids, a putative promoter region of *legK4* Philadelphia (500 base pairs) was placed upstream to the TTG codon. A  $\Delta$ *legK4* defective mutant from the Paris strain was transformed with each plasmid. Finally, fluorescence was observed, and the protein weight was determined by western blot.

The *p<sub>legK4</sub>-ATG-sfGFP* condition was considered as a positive control. If the TTG was not the start codon, the ATG downstream led to a fluorescent protein (sfGFP-LegK4) (Figure 18C). For both plasmids, two independent clones were used to determine the presence of fluorescence.

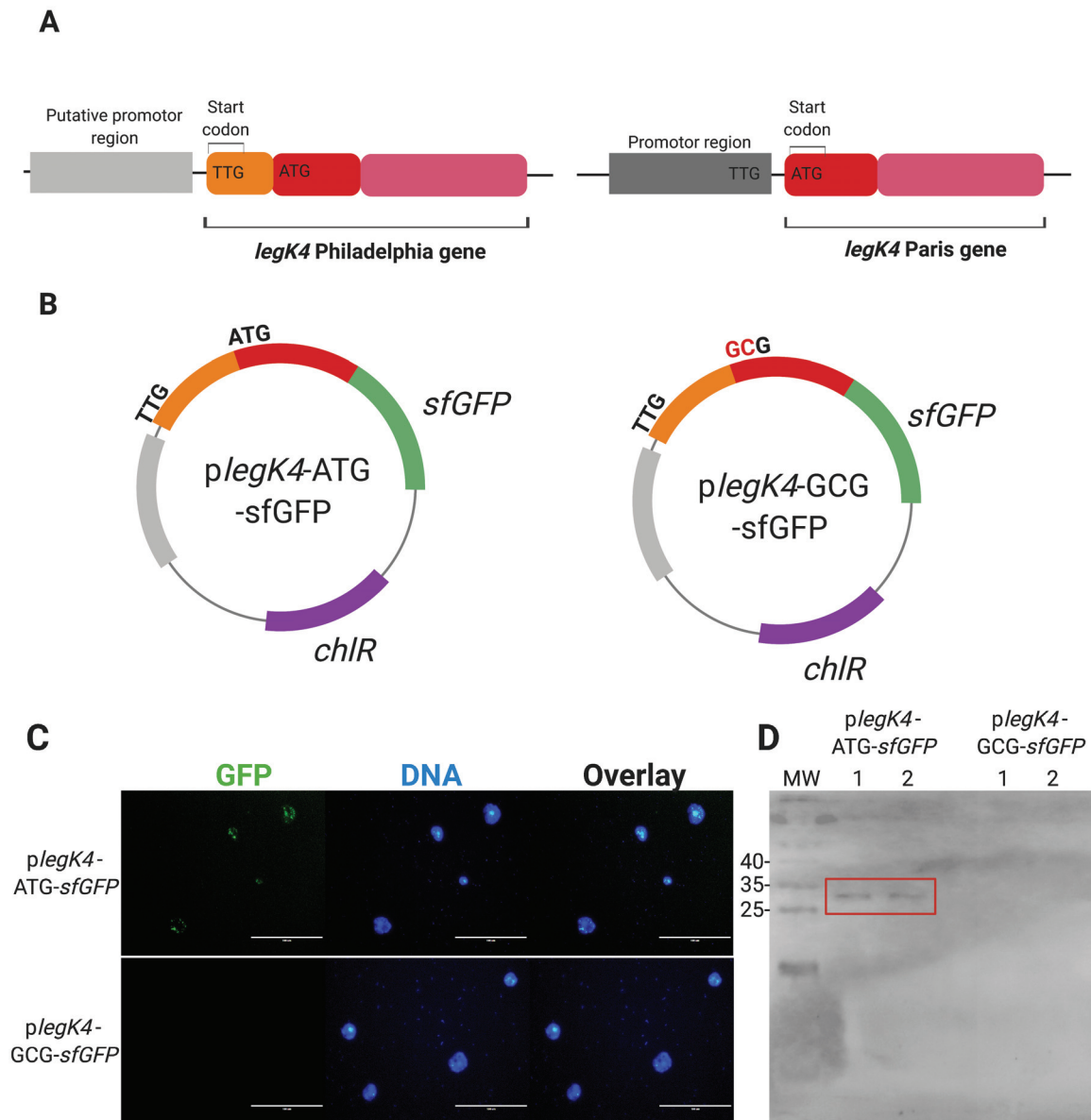
After transformation, bacteria were observed using an epifluorescence microscope, but we could not detect any fluorescence in either case.

We identified and validated by PCR the presence of *legK4* and *sfGFP* gene in transformed bacteria.

In both conditions, we added a putative promoter region of *legK4* Philadelphia, but we did not know what signal triggers LegK4 production.

To determine whether the concentration of bacteria could have an impact, we incubated them overnight in an AYE medium.

Only very few bacteria containing *plegK4*-ATG-sfGFP exhibited fluorescence by epifluorescence microscopy. At the same time, however, bacteria transformed with *plegK4*-GCG-sfGFP were not fluorescent.



**Figure 18: LegK4 Philadelphia is wrongly annotated in database**

A. Genetic environment of *legK4* in Philadelphia and Paris strain.

B. *plegK4*-ATG-sfGFP was constructed using 500 bp before *legK4* Philadelphia and a part of *legK4* were fused with *sfGFP*. *plegK4*-GCG-sfGFP was a *plegK4*-ATG-sfGFP with a mutation on the ATG.

C. After transformation with *plegK4*-ATG-sfGFP or *plegK4*-GCG-sfGFP, bacteria were put in contact with *Acanthamoeba castellanii* for 2 h. The fluorescence from sfGFP was followed by epifluorescence.

D. After transformation with *plegK4*-ATG-sfGFP or *plegK4*-GCG-sfGFP, two clones were incubated overnight in AYE medium and lysed with Laemmli buffer. After determining the difference in molecular weight, the samples were examined by western blot using anti-GFP antibody. Numbers represent protein molecular weights (MW) in kDa.



To increase the signal and the number of bacteria expressing sfGFP, we then placed bacteria in contact with polylysine-coated plates or with *Acanthamoeba castellanii*. We hoped that the presence of a host or a surface to adhere would improve the quantity of sfGFP-LegK4 and, thus, the fluorescence.

Interestingly, these conditions successfully increased the fluorescence emitted by bacteria transformed with *p<sub>legK4</sub>-ATG-sfGFP* (Figure 18C). Moreover, in these conditions, no fluorescence could be detected in the case of the bacteria transformed with *p<sub>legK4</sub>-GCG-sfGFP* (Figure 18C).

All these results suggested that the TTG codon found in *legK4* Philadelphia could not be considered as a start codon.

Interestingly, these results showed also that LegK4 expression was induced by contact with surfaces or with amoebae.

The absence of fluorescence could also be due to a wrongly folded protein. We then investigated the presence and molecular weight of sfGFP-LegK4 by western blot using an anti-GFP antibody. Transformed bacteria were incubated overnight in AYE medium and then lysed with Laemmli buffer.

The two clones containing the *p<sub>legK4</sub>-ATG-sfGFP* expressed a chimeric protein characterized by a molecular weight between 25 and 35 kDa (Figure 18D). In the plasmid, if the ATG was the start codon, the protein would be 32.9 kDa, but if the TTG was the start codon, sfGFP-LegK4 would be a 40-kDa protein. This result showed that the ATG was the start codon and not the TTG. Moreover, no sfGFP was detected in clones transformed with the *p<sub>legK4</sub>-GCG-sfGFP* (Figure 18D) because the absence of ATG led to the absence of sfGFP-LegK4 production.

Together, these results showed that the genome annotation of LegK4 Philadelphia was incorrect and could lead to mistakes.

We demonstrated that the TTG found in *legK4* Philadelphia was not the start codon. In Paris and Philadelphia strains, the start codon for *legK4* is a classical ATG codon. Our results revealed that the differences in localization observed between LegK4 Paris and LegK4 Philadelphia were due to the additional presence of 58 amino acids (Figure 16A and B), and, in this work, we demonstrated that these 58 amino acids were not produced in bacteria.

This annotation mistake could lead to many consequences, such as the use of a protein with the wrong size, thus a wrong localization (Figure 16A eGFP-LegK4 Philadelphia). LegK4 Philadelphia did not localize in the cytoplasm but in the nucleus like LegK4 Paris and Lens (Figure 16A eGFP-LegK4 Philadelphia $\Delta$ 1–58).

## 1.4 Determination of the nuclear localization signal (NLS) of LegK4 Paris

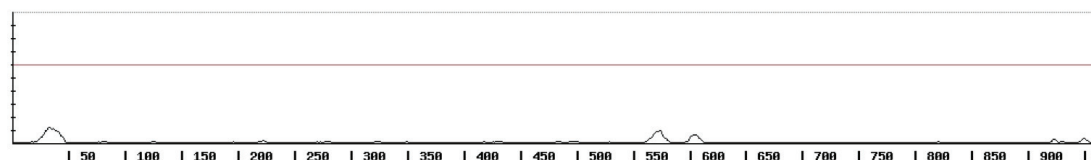
To migrate from the cytosol to the nucleus LegK4, Paris must possess a nuclear localization signal (NLS). Indeed, it is known that proteins over 40 kDa cannot go through nuclear pores by passive diffusion and must have an NLS to migrate to the nucleus. Usually, an NLS is made of a stretch of basic amino acid, also called a classical NLS (cNLS).

Using bioinformatic software, we investigated whether LegK4 Paris exhibited a cNLS. NLStradamus is online software that can predict the presence of a cNLS. This software uses a hidden Markov model (HMM) to compare the protein of interest se-

quence to a database of known NLSs and determine the presence of basic amino acids [393].

Using NLStradamus, no NLS was clearly detected in LegK4 Paris, but three small peaks under the NLStradamus threshold were present. These peaks represented sequences with several basic amino acids but not enough to be characterized as an NLS by the software (Figure 19).

Using a smaller cutoff, three putative NLSs could be then found: 30–KGIRKLIK-SANGK–42, 569–NPSKVKKGE–577 and 602–RGREGKP–608 (Figure 19).



**Figure 19: LegK4 Paris NLS prediction made with NLStradamus**  
Number represents amino acid position, and the red line represents the prediction cutoff.

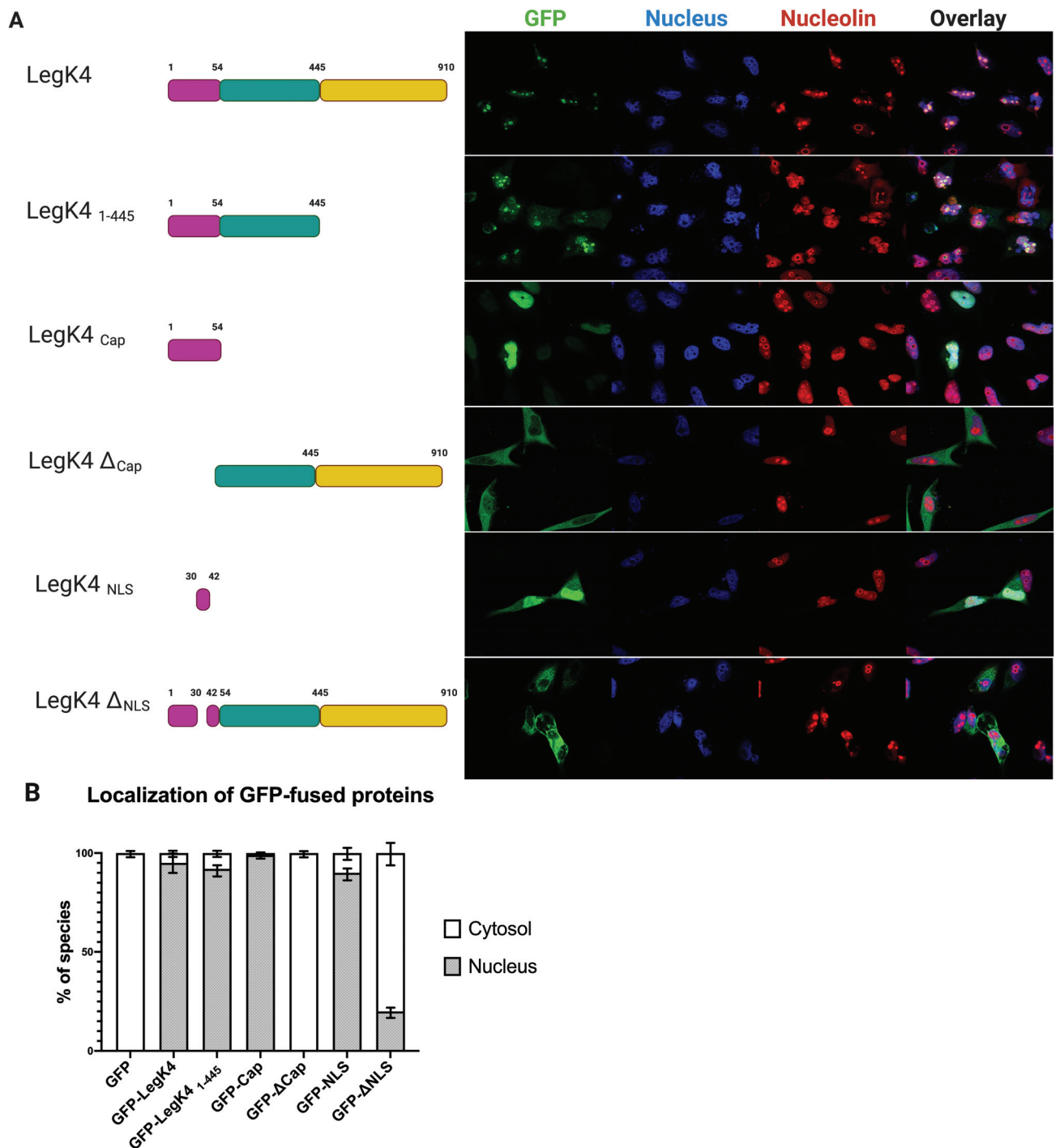
Given this information, we conducted protein truncation experiments and looked for GFP-fused protein localization by confocal microscopy.

We previously demonstrated that LegK4 Paris (LegK4 1–910) is localized in the nucleus of HeLa cells.

The expression of the first 445 amino acids from LegK4 Paris (LegK4 1–445) did not change its nucleolar localization (Figure 20A and B).

Then, we expressed the same construct without the first 54 amino acids, representing the Cap domain and containing the first putative NLS found by NLStradamus (30–KGIRKLIK-SANGK–42). In this last condition, LegK4 was no longer able to migrate to the nucleus and was restricted to the cytosol (Figure 20A and B).

This result suggested that the Cap domain could contain a peptidic signal triggering LegK4 migration from the cytosol to the nucleus.



**Figure 20: LegK4 Paris possesses an atypical NLS in its Cap domain**

A. HeLa cells were transiently transfected with different eGFP-tagged LegK4 constructs to identify LegK4 NLS. LegK4 1–910: LegK4 full length. LegK4 1–445: Part of LegK4 used for the 3D structure ([267]). LegK4 Cap: LegK4 1–54. LegK4  $\Delta$ Cap: LegK4 54–910. LegK4 NLS: LegK4 30–42. LegK4  $\Delta$ NLS: LegK4 without residues 30–42. Nuclei were stained blue with DAPI, and nucleoli were immunostained with an anti-C23 (nucleolin) antibody.

B. Amount of GFP-positive cells in the cytosol or the nucleus was quantified by counting 100 eGFPc positive cells per experimental condition. Each condition was done in triplicate. Error bars represent SEM.

Based on this analysis and to verify this hypothesis, we expressed only the Cap domain fused to eGFPc and, surprisingly, eGFP-Cap co-localized with DAPI staining (Figure 20A and B).

This result indicated that the Cap domain could play an important role in LegK4 nuclear localization.

The Cap domain is made of 4  $\alpha$  helix, but only  $\alpha$  helix 2 and 3 are required to lead eGFP-LegK4 into the nucleus. These helices contain a small basic amino acid enrichment made of arginine and lysine residues, as in classical NLS.

In a second experiment, we used the first sequence identified with NLStradamus to construct *peGFP-legK4*<sub>30-42</sub>. This plasmid expressed LegK4 Paris from 30 to 42 fused to the eGFPc (eGFP-KGIRKLIKSIANGK). We determined its localization in transfected HeLa cells.

The expression of this small sequence led to the accumulation of eGFPc in the nucleus, showing that it was crucial for LegK4 nuclear targeting (Figure 20A and B).

Finally, to prove that this sequence was essential to LegK4 nuclear localization, we constructed and expressed LegK4 without the amino acids 30–42, also called LegK4 $\Delta$ *NLS*.

As expected, eGFP-LegK4 was no longer detected in the nucleus but only in the cytosol (Figure 20B). In addition, the single or multiple mutagenesis of this atypical NLS into alanine did not change the localization of the LegK4 (data not shown).

The particular localization of LegK4 Paris in the nucleus was due to the presence of an NLS between the amino acid 30 to 42, found in the Cap domain. The presence of this sequence was essential to the migration of LegK4 from the cytosol to the nucleus. Because the NLS from all LegK4s protein is well conserved, we hypothesize that LegK4 Corby and LegK4 Alcoy could also migrate to the nucleus.

## 1.5 Migration of LegK4 Paris to the nucleus

LegK4 is a high molecular weight protein over 100 kDa and cannot migrate to the nucleus by passive diffusion. Comparison with previously found NLSs from eukaryotic and prokaryotic proteins revealed that LegK4 exhibited an atypical NLS because it did not look like any of them.

Like any other NLS, this sequence is recognized by eukaryotic machinery before the migration to the nucleus.

Because LegK4 does not possess a classical NLS, we do not know how the migration mechanism of this protein. Three migration paths are commonly described: the common importin  $\alpha/\beta$  pathway, importin  $\beta$  pathway or another unknown pathway.

We used two different strategies to answer this question: drugs inhibiting proteins involved in nuclear migration and mutated yeasts deleted for specific importins.

Two different drugs were used, ivermectin that inhibits the importin  $\alpha/\beta$ -mediated nuclear import and importazole that inhibits the importin  $\beta$ -mediated nuclear import. HeLa cells expressing eGFP-LegK4 or eGFPc were transfected, and 24 h after transfection, they were treated for 1 h with 25  $\mu$ M of ivermectin or 100  $\mu$ M of importazole or DMSO as a negative control. We then investigated whether the drug changed LegK4 localization by confocal microscopy.

Cells treated with the drug and expressing LegK4 Paris were all dead (data not shown). Surprisingly, we did not see this phenomenon with cells that expressed the eGFPc. Even using various concentrations of both inhibitors, these experiments were not conclusive, probably due to the impact of LegK4 combined with the drug effect.

The other strategy was to determine LegK4 Paris localization in a mutant of *Saccharomyces cerevisiae* BY4742 deleted for some importin ( $\alpha/\beta$  or  $\beta$ ).

We constructed plasmids encoding eGFP-LegK4 or eGFPc as a negative control.

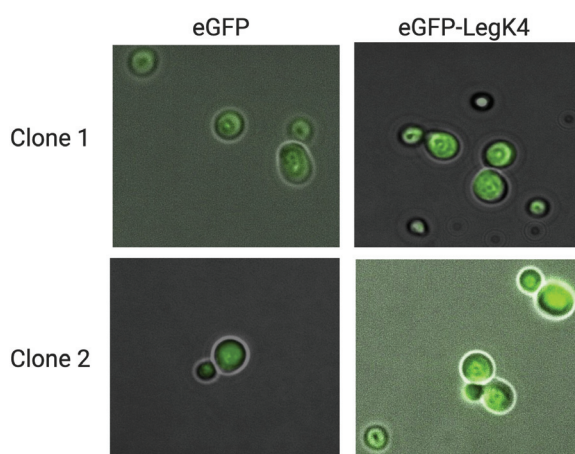
The protein of interest was placed under a constitutive promoter (pMet-YC) or in an inducible way with galactose (pBT3N) to monitor LegK4 cytotoxicity.

Yeasts were then transformed with the different plasmids and selected on a synthetic dextrose (SD) plate without methionine.

Unfortunately, no transformants were found with the constitutive expression plasmid demonstrating the LegK4 cytotoxicity again in eukaryotic cells.

Yeast transformed with the inducible plasmid were observed by epifluorescence microscopy to confirm the absence of basal fluorescence. Then they were cultured in YP (yeast extract peptone) medium in the presence of galactose to induce LegK4 expression.

After 24 h of culture, we verified by epifluorescence microscopy the presence of eGFPc and eGFP-LegK4. We could detect fluorescence in both conditions, but the eGFPc and eGFP-LegK4 were both present in the cytosol and nucleus of yeast (Figure 21).



**Figure 21:** eGFPc and eGFP-LegK4 localization in yeast cells

Plasmid extraction followed by PCR and restriction enzyme digestion were then conducted in parallel to verify the presence of *legK4*, *GFP* and the methionine auxotrophy gene.

For the e*GFP-legK4* plasmid, all the clones had lost *legK4* but still possessed *GFP* and the methionine auxotrophy gene. The green fluorescence in yeast transformed by the e*GFP-legK4* plasmids was therefore due to the eGFPc and not eGFP-LegK4.

This experiment using *S. cerevisiae* demonstrated that LegK4 was toxic to yeast cells and in this condition was not an appropriate model for our study.

## 1.6 Translocation of LegK4 Paris into infected cells

Before this study, a translocation assay of LegK4 into the cytosol of infected macrophages was conducted to obtain insight into LegK4's role during *Legionella* infection.

A translocation assay was conducted in infected human macrophages using  $\beta$ -lactamase fused protein and CCF4-AM [394]. This technique allows determining precisely when the effectors are secreted. Moreover, the secretion timing could help to localize LegK4 during the infection of human cells.

Prior to the infection, macrophages were treated with CCF4-AM, a fluorescence resonance energy transfer (FRET) substrate. Macrophages were then infected with a strain of *L. pneumophila* that expressed LegK4 fused with  $\beta$ -lactamase (BlaM-LegK4). The release of BlaM-LegK4 into the cytoplasm was detected and quantified based on the extent of cleavage of CCF4-AM loaded into cells.

The CCF4-AM consists of the fluorescence donor (coumarin, blue emission) and acceptor (fluorescein, green emission) moieties joined by the cephalosporin linker.



Due to the FRET between the donor and the acceptor, uncleaved substrate excited at 400 nm (to excite coumarin) emits in green at 528 nm . Cleavage of the cephalosporin linker present in CCF4-AM by BlaM-LegK4 in the cytosol resulted in the loss of FRET and a blue shift (460 nm) in the emission spectrum.

The extent of fusion was thus readily measured as the ratio of fluorescence intensities at 460 nm and 528 nm, using a TECAN<sup>®</sup> micro-plate reader [394].

LegK4 was secreted early during *L. pneumophila* infection in macrophages, around 15 min after contact (internal laboratory data).

This result may suggest that LegK4 could probably have an early role during the infection. To localize LegK4 during the infection of human cells, we should verify it after 15 min.

## 1.7 Localization of LegK4 Paris during the infection of human cells

LegK4 localization during the infection could have been easily followed using eGFP-LegK4. Unfortunately, it is well known that substrates of T4SS are not secreted when fused with GFP or any other fluorescent protein in *Legionella* [395].

Moreover, the use of a chimeric protein with a tag could sometimes change protein conformation, thus localization. To avoid that and to be in the most physiological conditions possible, we aimed to detect LegK4 without any associated tag and planned to detect LegK4 by immunofluorescence and by western blot.

For this, we overexpressed and purified 6His-LegK4 from *E.coli* to make polyclonal anti-LegK4 antibodies.

Experiments of infection using wild-type Paris strain and a  $\Delta legK4$  mutant as the negative controls were then performed.

We infected U937 macrophages, A549 pneumocytes and HeLa epithelial cells and tried to detect LegK4 from 15 min to 24 h post-contact (post-infection).

Moreover, we changed the number of bacteria, called multiplicity of infection (MOI), from 10 to 500.

Even with 500 times more bacteria than human cells (MOI 500), we could not detect LegK4 by immunofluorescence with an anti-LegK4 antibody (data not shown). LegK4 endogenic level may be too low to be detected by microscopy.

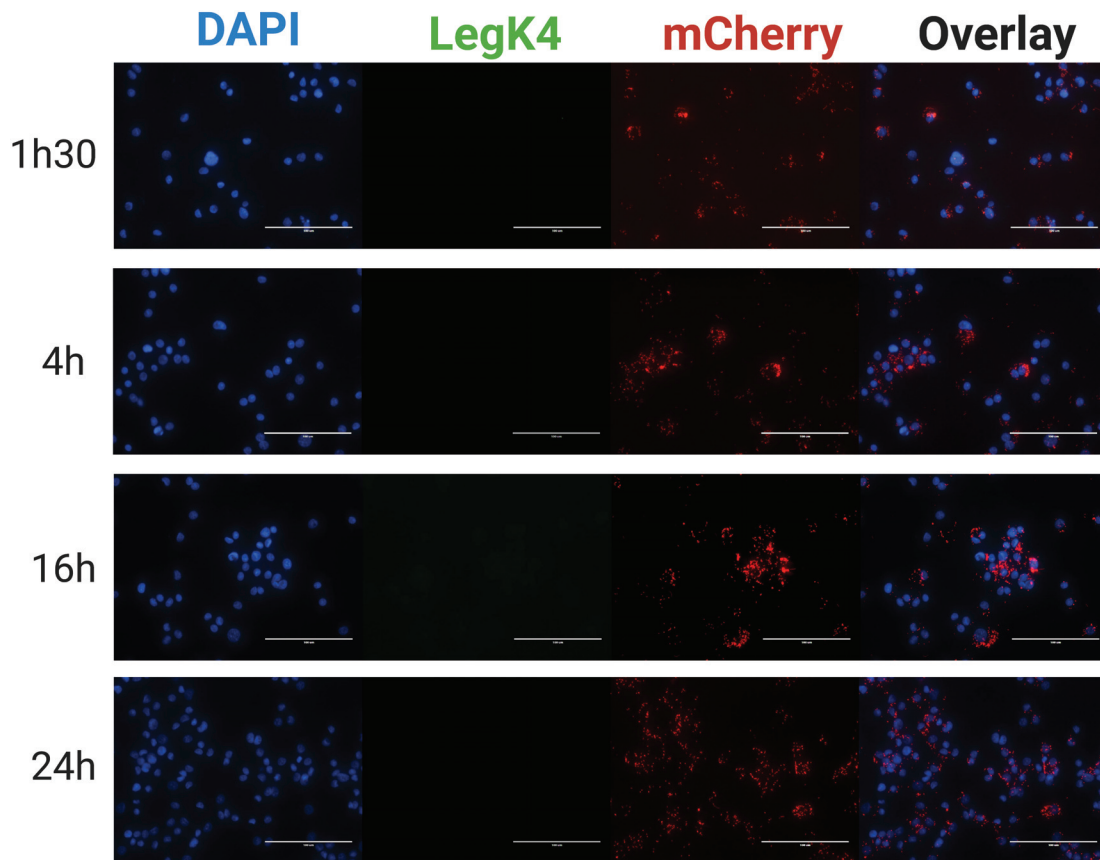
To amplify the signal in immunofluorescence, we transformed *L. pneumophila* with a multicopy plasmid (pXDC50-*legK4*) that allowed the expression of *legK4* and the mCherry in an inducible way. Indeed, the *legK4* and *mCherry* gene were under a *tac* promoter that induced the expression of LegK4 and the mCherry with IPTG.

Prior to the infection, LegK4 was induced by two different methods: overnight on plate CYE + IPTG or overnight in liquid medium AYE + IPTG to reduce the formation of filamentous bacteria. Indeed, *Legionella* can be found filamentous when stressed, and these bacteria were observed to be less infectious (internal data lab).

We infected macrophages and other cells with induced bacteria as previously described and verified the presence of LegK4 by immunofluorescence.

Unfortunately, we were not able to detect LegK4 in all these conditions even if LegK4 was overexpressed (Figure 22).

These results suggested that the sensitivity of the anti-LegK4 antibody was not enough to detect LegK4 in microscopy during the infection.



**Figure 22: Determination of LegK4 localization during macrophages infection**

U937 macrophages were infected at 500 MOI with a  $\Delta legK4$  *Legionella* containing pXDC50-*legK4* from 1 h 30 min to 24 h in RPMI medium supplemented with IPTG. Nuclei were stained blue with DAPI, LegK4 was immunostained with an anti-LegK4, and anti-rabbit IgG was conjugated with DyLight 488 (green). mCherry represents bacteria.

Because the anti-LegK4 antibody could not be used for immunofluorescence, we considered using it to detect LegK4 by western blot after cell fractionation. We infected human macrophages or pneumocytes with a strain of *L. pneumophila* transformed with pXDC50-*legK4* at high MOI (MOI 500) as previously described. After 24 h, cells were collected, lysed to release the nucleus and centrifuged to separate LCVs that contained bacteria from the nucleus.

Several times and speeds of centrifugation were tested, but we were not able to

separate the LCV from the nucleus (data not shown).

In parallel, we generated an isogenic  $\Delta legK4$  strain of *L. pneumophila* transformed with p4HA-*legK4*. This vector encoded LegK4-tagged with four human influenza hemagglutinins (4HA) that were placed under an inducible *tac* promoter.

We infected HeLa cells, human macrophages and pneumocytes from MOI 1 to 500 from 15 min to 24 h postinfection. We then verified the localization of LegK4 by immunofluorescence using an anti-HA antibody this time.

In these experiments, we could detect 4HA-LegK4 inside bacteria but not in human cells during the infection. 4HA-LegK4 seemed to be expressed and accumulated by *L. pneumophila* and not secreted by bacteria.

To understand why we could not detect LegK4 during the infection of human cells, we analyzed the transcription of *legK4*.

We used data from a transcriptome of *L. pneumophila* Paris during the infection of macrophages made by X. Charpentier (INSERM U1111) and collaborators (data not published).

In macrophages infected by *L. pneumophila* Paris, we observed that the LegK4 mRNA was transcribed at a low rate.

This suggested that LegK4 could be secreted in low quantity during the infection and might be difficult to detect. Furthermore, a study by Allombert and collaborators (2019) explained that the secretion level of T4SS substrate was not correlated to protein production. *Legionella* finely regulates the secretion of effectors during the infection [96], so overexpressing LegK4 did not mean that it will be secreted.

In host-pathogen relationship studies, it is usually difficult to localize a bacterial

effector during the infection of human cells, and LegK4 was no exception to the rule. The absence of detection usually depends on many factors, such as secretion timing, the cell line used or the quantity of the protein of interest. In this study, we tried to change the timing of infection and the cell line and also tried to overexpress LegK4 but still could not detect LegK4 during the infection.

LegK4 was found well conserved in five strains of *L. pneumophila*. Our experiments showed that LegK4 from the Paris strain localized in the nucleus of transfected HeLa cells. However, this result contradicted those from Moss *et al.* concerning LegK4 Philadelphia [269]. They demonstrated that LegK4 Philadelphia localized in the cytosol and not in the nucleus. Even if the peptidic sequence of LegK4s is very similar, LegK4 Philadelphia possesses a 58-amino-acid extension at the N-terminal region. We then determined that extension was responsible for the accumulation of LegK4 Philadelphia into the cytosol. LegK4 Philadelphia without this extension migrated to the nucleus like LegK4 Paris. However, we proved that this extension was never produced by *Legionella*; thus, LegK4 Philadelphia, like LegK4 Paris, was nuclear.

We aimed to find in the LegK4 sequence an NLS that is an essential element for the nuclear localization of the protein. An NLS never described before was present in the LegK4 Cap domain. This NLS was found to be very conserved and presented in the other strains of *Legionella*, suggesting that LegK4 Alcoy and LegK4 Corby could migrate to the nucleus as well. Unfortunately, we observed that LegK4 has a strong impact on eukaryotic cells, which did not help to characterize which mechanism LegK4 used to migrate to the nucleus.

Finally, we aimed to understand when LegK4 was secreted by *Legionella* and where it localized during the infection of human cells. A translocation assay in infected

macrophages was conducted to identify when LegK4 was secreted. LegK4 was secreted early during the infection, around 15 min postinfection. This result showed that LegK4 could have a role at the beginning of a *Legionella* infection. The production and secretion of LegK4 are finely regulated by *Legionella*; this regulation could be why we could not localize it during the infection of human cells.

## 2 LegK4 interactant and substrate

To better understand the role of LegK4, we aimed to identify its eukaryotic interactants and substrates. This investigation started with a yeast two-hybrid screen approach. Then we determined the interactants by co-immunopurification and the substrates by western blot and mass spectrometry.

### 2.1 Yeast two-hybrid screening

The yeast two-hybrid screen is a sensitive way to test the direct interaction between two targeted proteins. The protein of interest can also be used as a bait to screen libraries of proteins fragments prepared from the desired cell types, tissues or entire organisms.

Two fusions, also called “hybrids,” are constructed between each protein and either the DNA binding domain (DBD) or the activation domain (AD) of a transcription factor. The protein fused to the DBD is referred to as the “bait,” and the protein fused to the AD as the “prey.”

Upon interaction between the bait and the prey, the DBD and AD are brought close enough to activate the transcription of a reporter gene.

The identity of the interacting partners is then obtained by sequencing the corresponding plasmids in the selected yeast colonies.

The screening was set up in the yeast *S. cerevisiae* by the company Hybrigenics<sup>®</sup>. They fused the AD of LexA or Gal4 transcription factor to LegK4 and used proteins from a cDNA library of human macrophages fused with the DBD of LexA or Gal4 as

prey. A *S. cerevisiae* strain harboring an auxotrophy for histidine was transformed with a plasmid that can express the bait and the prey. If there is an interaction between the prey and the bait, *S. cerevisiae* produces histidine and can grow on a plate without histidine.

The expression of LegK4 in *S. cerevisiae* was reported by Hybrigenics<sup>®</sup> to be cytotoxic because yeast rapidly died during the test, delaying the results.

Once again, LegK4 expression seemed to have a strong impact on yeast cells. To limit this cytotoxicity, Hybrigenics used LegK4<sub>D195N</sub>, a mutated version of LegK4, also called a kinase-dead mutant. The absence of the kinase activity of LegK4 increased the yeast cell viability.

Finally, five eukaryotic interactants candidates were found: TOX4, SP140, RBBP4, PTEN and BiP. TOX4, SP140, and RBBP4 are known to be involved in chromatin remodeling and participate in the epigenetic control of eukaryotic gene expression.

- TOX4 (TOX high mobility group box family member 4) is a component of the PTW/PP1 phosphatase complex that plays a role in the control of chromatin structure and cell cycle progression during the transition from mitosis into interphase [396].
- SP140 is a component of the nuclear body and may play a role in chromatin-mediated regulation of gene expression [397].
- RBBP4 (retinoblastoma binding protein 4) is a component of several complexes that regulate chromatin metabolism and may target chromatin assembly factors, chromatin remodeling factors and histone deacetylases [398].
- PTEN activity is critical for tumor suppression and is involved in cell cycle progression and cell survival [399].
- BiP acts as a chaperone protein in the ER and is a crucial sensor of the ER stress-



inducing the UPR. By recruiting ER vesicles on its LCV, *L. pneumophila* should activate the UPR. However, it was admitted that several *Legionella* effectors interfere with the different UPR pathways to inhibit it [231]. TOX4, SP140, RBBP4 and PTEN retained our attention because of their nuclear localization.

The yeast two-hybrid screen showed that many other proteins interacted with LegK4, such as H3 histone. However, they were eliminated by Hybrigenics<sup>®</sup> because they were considered as part of the contaminant repository for affinity purification (CRAPome). These proteins interacted with low affinity with too many proteins and are usually removed from yeast two-hybrid screen results. We decided to add H3 histone to the list of potential interactants because two out of three nucleomodulins from *L. pneumophila* also interact with H3 histone [246, 245].

Consequently, we retained six candidates: the nuclear proteins TOX4, SP140, RBBP4, PTEN, H3 histone and the cytosolic protein BiP.

## 2.2 Validation of the interactants

The interaction partner found with the yeast two-hybrid screen must be validated with another technique. We decided to perform a co-immunopurification assay in transfected HeLa cells to validate these interactions.

For this, we used the GFP-Trap<sup>®</sup> technology that allows the purification of a GFP-tagged protein and its interactants. GFP-Trap<sup>®</sup> is an easy-to-use technology based on a particular domain  $V_{HH}$  of antibody against GFP derived from Camelids, such as alpacas. This technology can be used to specifically purified GFP-fused proteins and

their interacting factors in animals [400], viruses [401] or plants [402].

To validate our candidates, we transiently transfected HeLa cells with the different constructs: eGFP-LegK4 Paris, eGFP-*LegK4*<sub>D195N</sub> Paris catalytic mutant, eGFPc alone and eGFP-LegK2 as negative controls. eGFP-LegK2 could be considered as a control for LegK4 specificity.

We also determined whether the interaction was strain-dependent using eGFP-LegK4 Philadelphia and eGFP-LegK4  $\Delta$ 1–58 Philadelphia.

Finally, using specific antibodies, we verified the presence of TOX4, SP140, BiP, RBBP4, PTEN and H3 after their interaction with LegK4 or the controls.

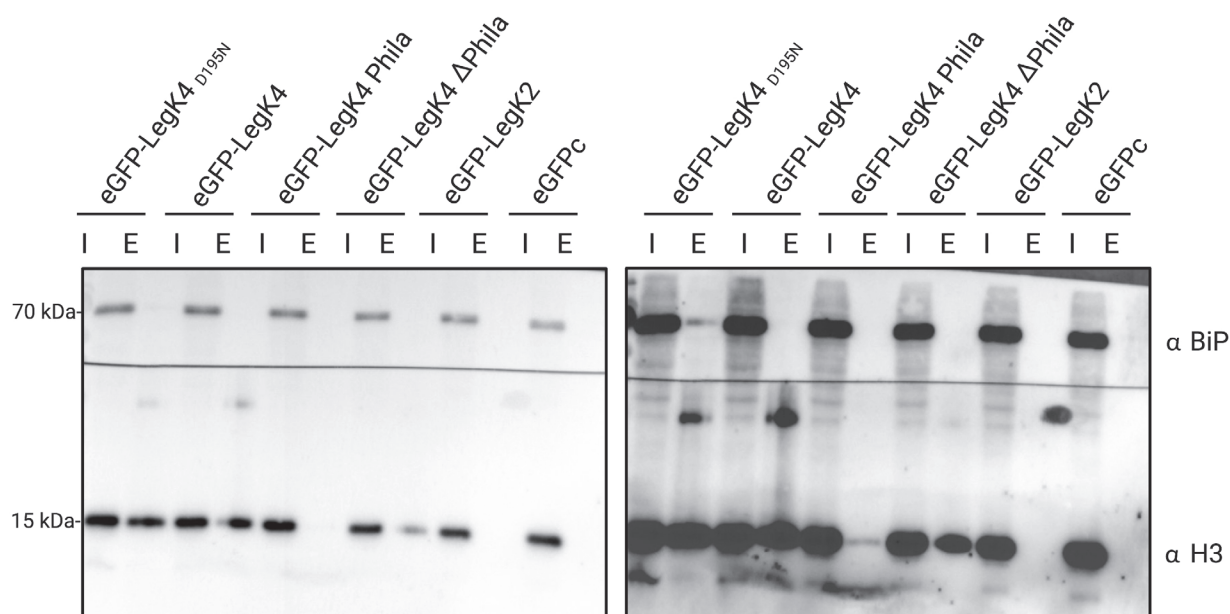
Unfortunately, TOX4, SP140, RBBP4 and PTEN were always present in the cell lysate (input) but were not present after interaction with LegK4 (elution) from any conditions (data not shown). We detected a strong signal, even after short exposure, corresponding to an interaction between H3 histone and LegK4 (Paris and  $\Delta$ 1–58 Philadelphia) or with its catalytic mutant.

However, we could not detect an interaction between H3 and LegK4 Philadelphia, LegK2 and eGFPc, showing that the interaction between H3 and LegK4 was very specific.

BiP was the only protein found after long exposure to interact with *LegK4*<sub>D195N</sub> (Figure 23). The absence of the kinase activity of LegK4 could help to stabilize the interaction. The low transfection rate of eGFP-LegK4 and a labile interaction between LegK4 and BiP could explain why there was no signal with LegK4 that possessed a kinase activity. The interaction with LegK4 and BiP was very specific because we could not detect BiP in the controls.

These results showed that the interactions between H3 and LegK4 and BiP and LegK4 were highly specific (Figure 23). H3 was also detected, after a longer time ex-

posure, in the eGFP-LegK4 Philadelphia condition. However, the signal was very weak compared to eGFP-LegK4 Paris or eGFP- $\Delta$ 1–58 Philadelphia. This result showed again that the 58-amino-acid extension found in LegK4 Philadelphia seemed not to be present. From the yeast two-hybrid screen candidates, BiP was the only protein localized in the cytosol. However, BiP was reported to be also nuclear and can cross-link with DNA to suppress DNA-damaged induced apoptosis [403].



**Figure 23: H3 and BiP interact with LegK4 in HeLa cells**

Co-immunoprecipitation assay from HeLa cells expressing eGFP-*LegK4*<sub>D195N</sub> (LegK4 Paris catalytic mutant), eGFP-LegK4 (LegK4 Paris WT), eGFP-LegK4 Phila (LegK4 Philadelphia), eGFP-LegK4  $\Delta$ Phila (LegK4  $\Delta$ 1–58 Philadelphia), eGFP-LegK2 or eGFPc. eGFP-LegK2, another *Legionella* kinase and eGFPc were used as a specificity control. Input (I) represented whole cell lysate, and Elution (E) represented the result of the interaction. Each condition was revealed using an anti-H3 antibody or anti-BiP. Short (Left) and Long (Right) exposure.

## 2.3 *In vitro* validation of LegK4 substrates

We showed previously that LegK4 specifically interacts with H3 and BiP. However, it was important to verify whether these proteins could be phosphorylated *in vitro* by LegK4 and thus could be considered as substrates.

### 2.3.1 H3 Histone

To determine if H3 histone was a LegK4 substrate, we conducted an *in vitro* phosphorylation assay. We mixed purified 6His-LegK4 with commercial nucleosome, a mixture of histones (H1, H2A, H2B, H3, and H4) isolated from calf thymus or two isoforms of histone H3 (H3.1 and H3.3). Then we used an anti-phosphothreonine to reveal the phosphorylated proteins. As a negative control, experiments were also conducted with the kinase-dead mutant 6His-*LegK4*<sub>D195N</sub>.

Two signals were detected in all the conditions tested with 6His-LegK4 except for the nucleosome. They represented phosphorylated proteins, one around 100 kDa and another one around 15 kDa. The signal around 100 kDa was due to the LegK4 autophosphorylation as previously observed by Flayhan *et al.* [267].

According to the protein size around 15 kDa, the second phosphorylated protein might be H3 histone (Figure 24A). In the conditions using isoforms of H3 histone, we could detect a phosphorylation signal at 15 kDa that represented phosphorylated H3 (Figure 24A).

### 2.3.2 Identification of the H3 phosphorylated residue by LegK4

To determine which residue of H3 was phosphorylated, we then sent our samples from the *in vitro* phosphorylation assay (Figure 24A Well 3 and 5) to the mass spectrometry facility platform.

As expected, according to the molecular weight, H3 histone was found phosphorylated and, more precisely, on the threonine 32 (T32) residue.

We then revealed our *in vitro* phosphorylation assay samples using an anti-H3pT32 antibody to detect the phosphorylation of H3 histone on threonine 32 residue.

We were able to detect this phosphorylation in the histone mix, H3.1, and H3.3 when mixed with 6His-LegK4.

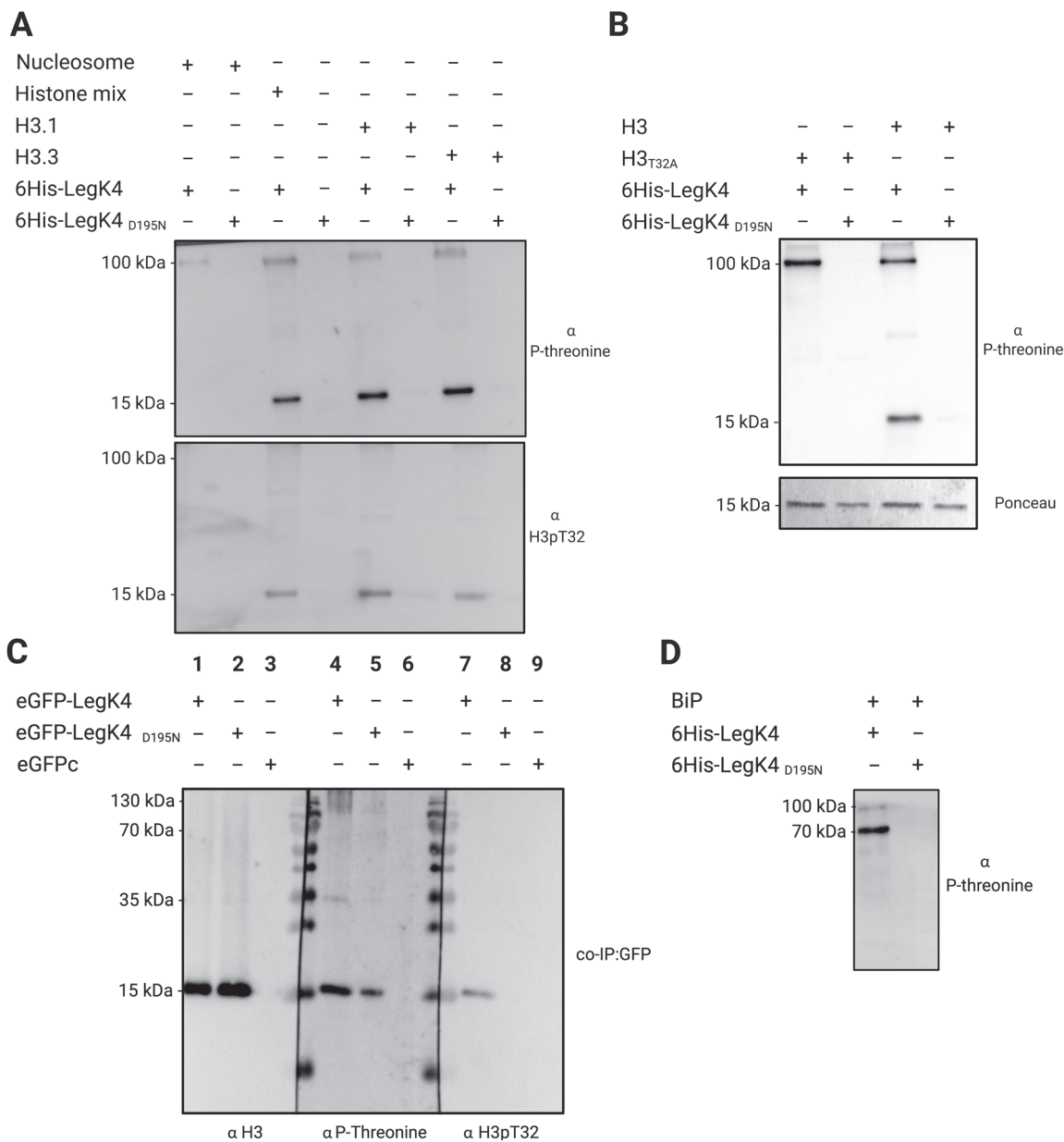
These results showed that LegK4 phosphorylated specifically H3 on T32 residue. In our phosphorylation assay, we could detect almost the same signal when we used H3.1 and H3.3. There are small amino acid differences between H3.1 and H3.3, but T32 is present in both. Unfortunately, we could not detect a signal in the condition using purified nucleosomes. This result suggested that in the nucleosome, H3 histone was probably not accessible for phosphorylation by LegK4.

We then determined whether the single mutation on the H3 T32 phosphorylation site was sufficient to abrogate LegK4-mediated phosphorylation. We mutated T32 from H3 into non-phosphorylatable alanine (T32A). To do another *in vitro* phosphorylation assay, we first overexpressed in *E. coli* H3 and H3 T32A. *E. coli* were transformed with pGEX-*H3* or pGEX-*H3T32A* that allowed the expression of GST-H3 and GST-H3 T32A, respectively. Bacteria were lysed, and GST-tagged proteins were purified using glutathione magnetic beads. GST is known to be phosphorylated by kinase, so

we cleaved it with PreScission Protease to avoid its phosphorylation by LegK4 [404]. Finally, we mixed purified H3 or H3 T32A with 6His-LegK4 or as a negative control 6His-*LegK4*<sub>D195N</sub>.

As previously observed using an anti-phosphothreonine antibody, we detected the autophosphorylation of LegK4 around 100 kDa. However, we could detect a signal corresponding to phosphorylated H3 only when we used H3, not when we used the mutated H3 T32A protein.

This result revealed that this single mutation was sufficient to abrogate LegK4-mediated phosphorylation, suggesting a high specificity for a single site of phosphorylation (Figure 24B).



**Figure 24: H3 and BiP are host proteins targeted by LegK4**

A. *In vitro* phosphorylation assay of the histone mix (H1, H2A, H2B, H3, and H4), H3.1 or H3.3 or nucleosome with 6His-LegK4 or the catalytic mutant 6His-LegK4<sub>D195N</sub>. The results were revealed by western blot using an anti-phosphothreonine or anti-H3pT32 antibody.

B. *In vitro* phosphorylation assay of H3 histone or non phosphorylable H3 (H3T32A) with 6His-LegK4 or 6His-LegK4<sub>D195N</sub>. The gel loading control was done by Ponceau S staining (right after transfer), and the results were revealed by western blot using an anti-phosphothreonine or anti-H3pT32 antibodies.

C. Co-immunoprecipitation (co-IP) assay using GFP-Trap<sup>®</sup> (co-IP:GFP) from cells expressing eGFP-LegK4, eGFP-LegK4<sub>D195N</sub> or eGFP. The co-IPs were revealed using an anti-H3, anti-phosphothreonine or anti-H3pT32 antibody.

D. *In vitro* phosphorylation assay mixing BiP with 6His-LegK4 or 6His-LegK4<sub>D195N</sub>. The results were revealed by western blot using an anti-phosphothreonine antibody.

### 2.3.3 BiP

To determine whether BiP was also a LegK4 substrate, we conducted an *in vitro* phosphorylation assay using BiP.

Purified BiP was incubated with 6His-LegK4 or the kinase-dead mutant 6His-LegK4<sub>D195N</sub>. Proteins were separated by SDS-PAGE and probed with an anti-phosphothreonine antibody.

BiP was phosphorylated *in vitro* by 6His-LegK4, while no or weak phosphorylation was observed with the kinase-dead mutant (Figure [24D](#)).

### 2.3.4 Identification of the BiP phosphorylated residue by LegK4

To determine which residue of BiP was phosphorylated, we then sent our sample produced during the *in vitro* phosphorylation assay to the mass spectrometry facility platform.

Surprisingly, the mass spectrometry could not detect either BiP or a phosphopeptide.

Moss *et al.* showed in 2019 that LegK4 from Philadelphia strain was able to phosphorylate BiP on T518 residue [\[269\]](#). However, to date, there is no other article about the BiP phosphorylation on this residue, but its phosphorylation could impact the UPR [eMoss2019](#). Moreover, no antibody against BiP T518P was available to detect this modification by western blot or flow cytometry.



## 2.4 Validation of LegK4 substrates in transfected cells

Finally, we aimed to determine whether H3 and BiP were phosphorylated by LegK4 in transfected HeLa cells. H3 and BiP are both described as being phosphorylated on serine/threonine residues by cellular kinase [405, 269].

### 2.4.1 H3

To determine whether H3 was phosphorylated in HeLa cells, we used the elution fraction obtained from the co-immunopurification by GFP-Trap<sup>®</sup> with eGFP-LegK4, eGFP-*LegK4*<sub>D195N</sub> and eGFPc. We used an anti-H3 antibody to detect H3 histone, an anti-phosphothreonine antibody to detect all the phosphorylated proteins, and, finally, an anti-H3T32 antibody to detect the phosphorylation of H3 on T32 residue (Figure 24C).

As expected, we could detect a signal corresponding to H3 in the well with eGFP-LegK4 and eGFP-*LegK4*<sub>D195N</sub> (Figure 24C Well 1 and 2) but not with eGFPc (Figure 24C Well 3). Using an anti-phosphothreonine antibody, we detected a phosphorylated protein around 15 kDa that was phosphorylated H3 (Figure 24C Well 4 and 5). Using ImageJ software, we measured and compared the band intensity of H3 versus phosphorylated H3. For LegK4, the ratio intensity of phosphorylated H3/H3 was 1.16 compared to 0.34 for the catalytic mutant *LegK4*<sub>D195N</sub>. Finally, using an anti-H3 T32P antibody, we confirmed that H3 was phosphorylated by LegK4 on the threonine 32 residue (Figure 24C Well 7).

H3 was phosphorylated on the threonine 32 residue by LegK4 in HeLa cells. These results were consistent with the mass spectrometry result from the *in vitro* phosphory-

lation assay.

### 2.4.2 BiP

To determine whether BiP was phosphorylated in HeLa cells, we used the elution fraction from the co-immunopurification by GFP-Trap<sup>®</sup> (Figure 24C Well 4, 5 and 6). Because an anti-BiP T518P antibody or another tool was not available, we only investigated the presence of phosphorylated BiP with an anti-phosphothreonine antibody (Figure 24C Well 4).

We detected many signals but, unfortunately, no phosphorylated protein around 70 kDa, corresponding to BiP. Fewer signals were detected in the elution fraction with *LegK4*<sub>D195N</sub> because it does not have kinase activity. The anti-phosphothreonine antibody might not be sensitive enough, or there might not have been enough phosphorylated BiP in our sample.

We sent samples from the interaction test to the mass spectrometry to detect the presence of phosphorylated BiP. Surprisingly, the mass spectrometry could not detect either BiP or a phosphopeptide, probably due to the same reason as the *in vitro* phosphorylation.

We could not determine if BiP was phosphorylated by LegK4 in HeLa cells, and the lack of available tools to detect any phosphorylation was a handicap. However, by changing the treatment protocol of our samples before the mass spectrometry, we could validate that BiP was phosphorylated on threonine 518.

### 2.4.3 Identification of new substrates

The interaction test followed by a western blot using an anti-phosphothreonine antibody led to the identification of H3 histone as a substrate of LegK4. However, several others phosphorylated proteins were also observed (Figure 24C Well 4).

To identify these new substrates of LegK4, we sent the elution fraction from the interaction test with LegK4 or *LegK4<sub>D195N</sub>* to mass spectrometry.

We discovered new interactants of LegK4, but only six proteins were phosphorylated by LegK4 (Annexe 4). All these proteins are known to be nuclear proteins, which was consistent with our previous results. Moreover, many of them are known to be phosphorylated and involved in the cell cycle or transcription/translation (Annexe 4).

Heterogeneous nuclear ribonucleoprotein U (HNRNPU) and heterogeneous nuclear ribonucleoproteins A2/B1 (HNRNPA2B1) are both RNA-binding proteins that complex with heterogeneous nuclear RNA (hnRNA). These proteins are associated with pre-mRNAs in the nucleus and appear to influence pre-mRNA processing and other aspects of mRNA metabolism and transport. While all the hnRNPs are present in the nucleus, some seem to shuttle between the nucleus and the cytoplasm. Mass spectrometry revealed that LegK4 phosphorylated HNRNPU on serine 4 (Ser4) and serine 271 (Ser271) residues and HNRNPA2B1 on serine 189 (S189) residue. HNRNPU is phosphorylated by cellular kinase during mitosis [406].

Histone lysine acetyltransferase 7 (KAT7) is a histone acetyltransferase from HBO1 complexes that specifically mediate acetylation of histone H3 on lysine 14 residues (H3K14ac), thereby regulating various processes, such as gene transcription, protein

ubiquitination and immune regulation [407, 408]. Mass spectrometry revealed that LegK4 phosphorylated this protein on threonine 88 (T88). Phosphorylation by CDK1 on T88 is associated with mitosis [409].

Nucleolin (NCL) is a nucleolar phosphoprotein involved in the synthesis and maturation of ribosomes. It is located mainly in dense fibrillar regions of the nucleolus [410]. NCL is the major nucleolar protein of growing eukaryotic cells. It induces chromatin decondensation by binding to histone H1. It is thought to play a role in pre-rRNA transcription and ribosome assembly. Mass spectrometry revealed that NCL was phosphorylated by LegK4 on serine 67 residue (Ser67). NCL is known to be phosphorylated by cellular CDK2 on Ser67 during mitosis [410].

Nucleophosphomin (NPM1) possesses a bipartite NLS and a nucleolar localization signal at the C-terminal region. Mass spectrometry revealed that NPM1 was phosphorylated by LegK4 on the serine 4 residue (Ser4). When phosphorylated on Ser4, NPM1 is known to mediate the mitotic function of polo-like kinase triggering centriole duplication [411]. The centriole duplication occurs at the end of the G1 phase of the cell cycle. Together, this information could indicate that LegK4 might have a role in the G1 to S transition of the cell cycle.

Zinc finger protein 546 (ZNF546) is a potential nucleic acid-binding protein and transcriptional regulator. Mass spectrometry revealed ZNF546 was phosphorylated by LegK4 on the serine 373 (Ser373) residue; however, ZNF546 was not described as being phosphorylated on this residue by cellular kinase.

Our experiments showed that LegK4 interacted specifically with many nuclear proteins such as H3 histone or NPM1.

We validated that H3 histone was phosphorylated by LegK4 on the T32 residue *in vitro* and in transfected HeLa cells and confirmed that this protein is a substrate of LegK4. Unfortunately, very little is known about H3 T32P.

BiP was also validated to interact specifically with LegK4 in transfected HeLa cells and a LegK4 substrate *in vitro*. However, we could not validate that it was a substrate in HeLa cells, probably due to a technical problem with the mass spectrometry.

Thanks to mass spectrometry, we were able to find new interactants and substrates of LegK4. Surprisingly, all these proteins were nuclear and possessed activity related to the cell cycle.

### 3 Role of LegK4 Paris

The identification of LegK4 substrates was not enough to completely decipher its impact on *Legionella* and human cells. To better understand its role, we first determined whether LegK4 could have an impact on *Legionella* replication during the infection of macrophages. Then, to investigate the role of this kinase in the eukaryotic cell, we aimed to create a biologic tool consisting of stable cell lines that expressed LegK4 after induction. As seven of eight substrates of LegK4 were related to transcription, we also determined whether LegK4 could modulate the transcription of infected macrophages. Finally, we considered the role of LegK4 on the last substrate, BiP.

#### 3.1 LegK4 does not have an impact on *Legionella* intracellular replication

To determine whether the absence of LegK4 could have an impact on *L. pneumophila* Paris, we created a  $\Delta legK4$  strain and compared its intracellular replication to a wild-type strain of *Legionella*.

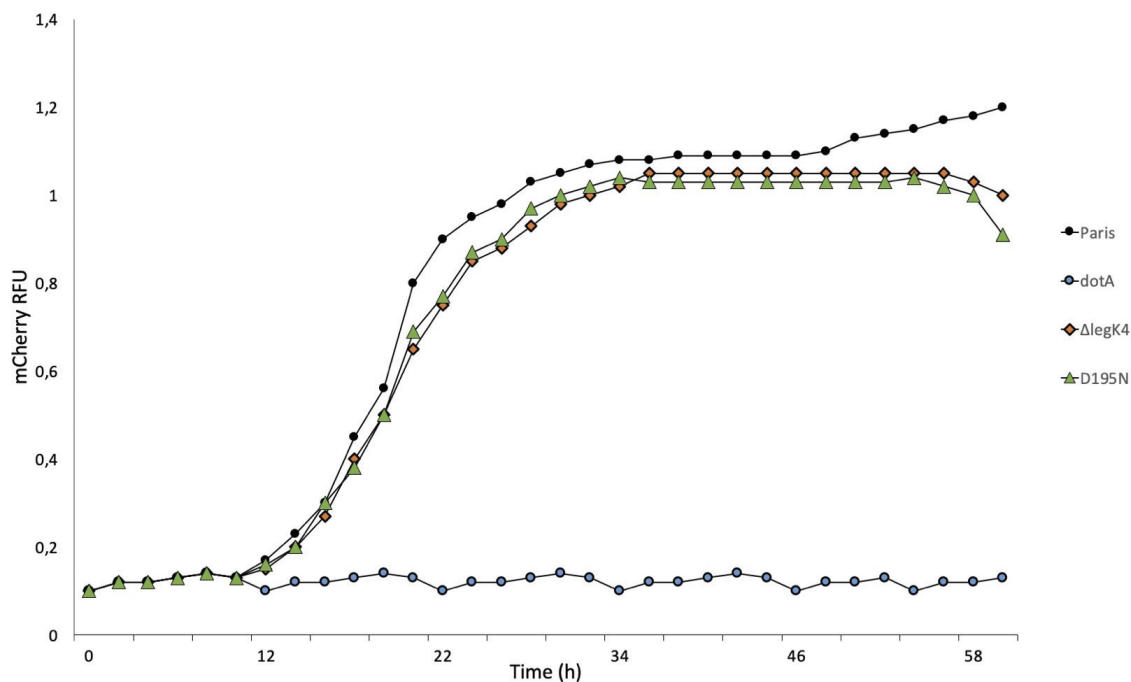
We also created another strain that expressed the inactive catalytic form of LegK4 (*LegK4*<sub>D195N</sub>), also called D195N, to see the impact of the kinase activity of LegK4.

Finally, a  $\Delta dotA$  strain was created as a negative control of infection. Indeed, the deletion of this protein from the Dot/Icm T4SS was known to induce complete avirulence [412]. We then transformed all these strains with a plasmid (pXDC50) that allowed the expression of the mCherry. Prior to the infection, we first investigated whether all the strains created could grow in an axenic medium with the same efficiency. For the

infection assays, we used the *L. pneumophila* environmental host *A. castellanii* (data not shown) or U937 human macrophages. Bacterial multiplication was monitored with a TECAN<sup>®</sup> plate reader by measuring the fluorescence emitted by the mCherry.

As expected, the  $\Delta dotA$  mutant was avirulent and, thus, was not able to infect or replicate into either macrophages or amoebae. On the other hand, all the other strains could replicate both in macrophages and amoebae with the same efficiency as the wild-type strain (Figure 25).

These results showed that LegK4 might not have a role in *L. pneumophila* intracellular replication in amoebae or macrophages.



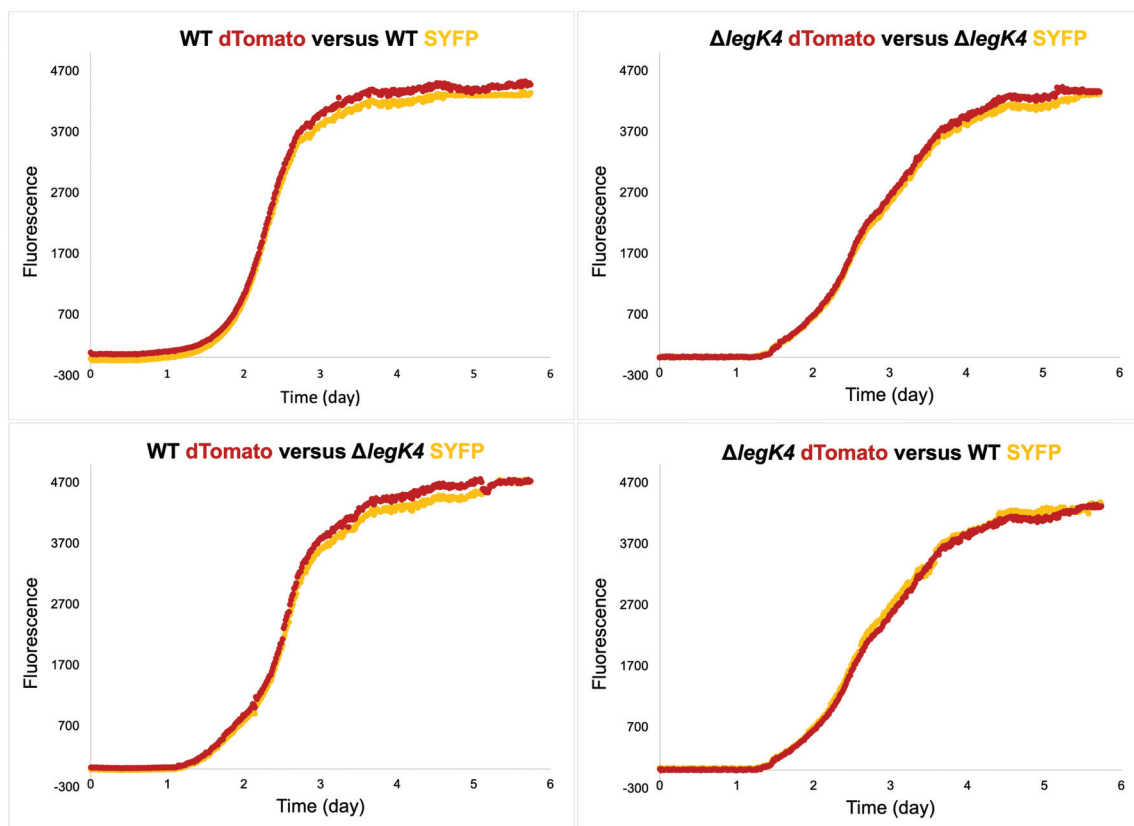
**Figure 25: Intracellular growth of *L. pneumophila* in U937 macrophages**

Cells were infected at an MOI of 10 with *L. pneumophila* expressing the *mCherry* gene on pDXC50. Bacterial multiplication was monitored with TECAN<sup>®</sup> by measuring the fluorescence of mCherry at an excitation of 587 nm and an emission of 610 nm every 2 h for 60 h. Fluorescence data were subjected to background subtractions using uninfected cells as a blank. We compared the intracellular replication of Paris wild type (Paris), a  $\Delta dotA$  strain (dotA), a  $\Delta legK4$  strain ( $\Delta legK4$ ) and a *legK4*<sub>D195N</sub> strain that expressed the kinase-dead mutant of LegK4 (D195N).

Because of functional redundancy, the deletion of one effector usually has no impact on *Legionella*'s intracellular cell growth.

To try to detect a small difference in *Legionella* replication, we decided to do a competition between a wild-type Paris strain versus a  $\Delta legK4$  at very low MOI. For this, these strains were tagged with two different fluorescent markers. By natural transformation and homologous recombination, we created bacteria that possessed in their chromosome either the *sYFP* (“super” yellow fluorescent protein) gene or the *dTomato* gene that allowed the production of two fluorescent proteins. These genes were inserted in a region known to have no impact on *Legionella* growth (internal laboratory data by E. Kay).





**Figure 26: Intracellular growth of fluorescent *L. pneumophila* in U937 macrophages**

Macrophages were infected with a mix of *L. pneumophila* expressing *dTomato* and *SYFP* both at MOI 0.005 (final MOI at 0.01). Bacterial multiplication was monitored with TECAN<sup>®</sup> by measuring the fluorescence every 20 min for 6 days. Fluorescence data were subjected to background subtractions using uninfected cells as a blank.

The strategy was to mix one SYFP strain with another *dTomato* at the same MOI, to infect macrophages and to monitor the fluorescence emitted by the SYFP and the *dTomato* with a TECAN<sup>®</sup> plate reader.

First, we determined whether the integration of these genes in the *L. pneumophila* genome affected the intracellular replication of the bacteria of the wild-type or the  $\Delta\textit{legK4}$  mutant strains (Figure 26 WT *dTomato* VS WT SYFP and  $\Delta\textit{legK4}$  *dTomato* VS  $\Delta\textit{legK4}$  SYFP). Then, we compared the intracellular replication of a wild-type strain (*dTomato* or SYFP) versus a  $\Delta\textit{legK4}$  (SYFP or *dTomato*) and could not detect

any significant differences.

In these conditions, even at a very low MOI (0.01), the strains behaved the same way in macrophages (Figure 26). These results suggested that LegK4 was not essential to infect human macrophages.

## 3.2 Impact of LegK4 on human cells

Because LegK4 targeted eukaryotic proteins such as H3 histone or BiP, our goal was to better understand LegK4 impact on human cells. For this, we created a biological tool that expressed LegK4 on demand. Then, we investigated whether LegK4 could change the gene expression of infected macrophages. Finally, we aimed to better characterize the relationship of LegK4 with another substrate BiP, using a siRNA approach.

### 3.2.1 Creation of LegK4 stable cell lines

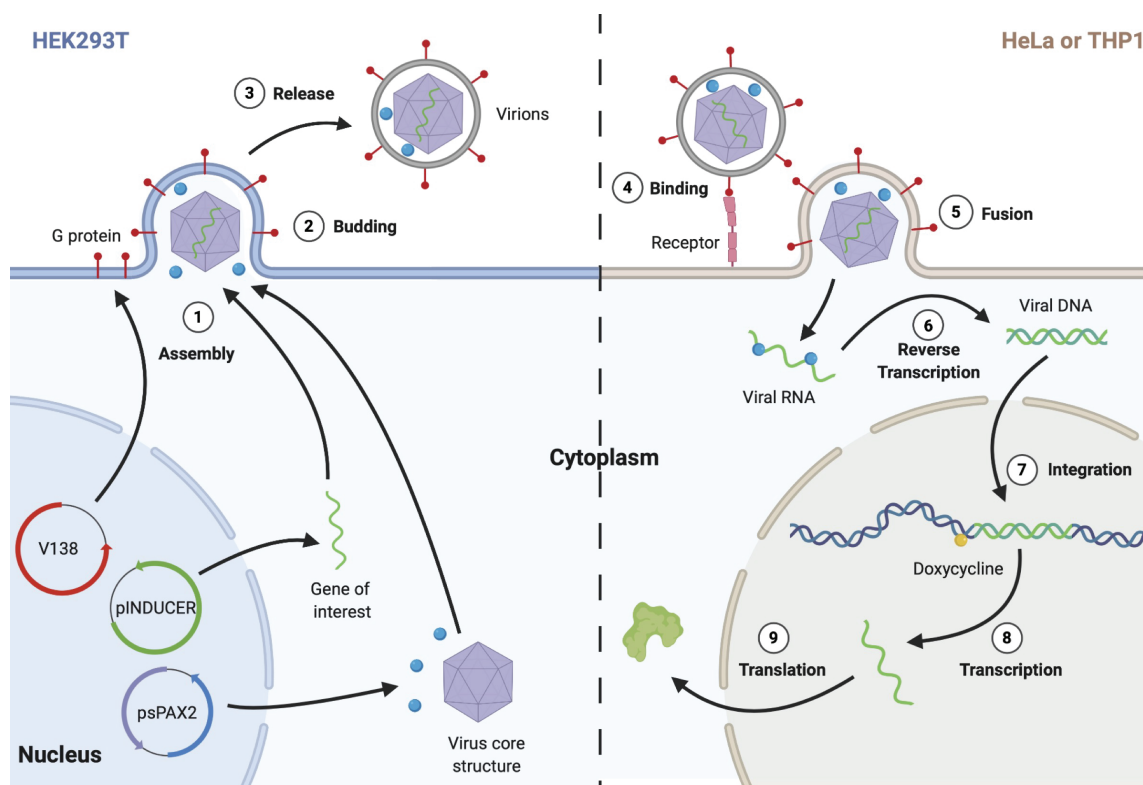
Experiments to find LegK4 localization highlighted crucial problems that complicated the search of LegK4 function in eukaryotic cells. The transfection rate of a plasmid that allowed the expression of LegK4 was always very low, around 20%. We tried to change the cell lines, the plasmids and the transfection technique to increase the transfection rate without success. Moreover, the expression of LegK4 was showed to be cytotoxic in eukaryotic cells many times (internal lab data). The creation of stable cell lines was conducted to improve LegK4 expression in mammalian cells, reduce cytotoxicity, and thus help us to better understand the LegK4 role.

### 3.2.1.1 Construction of lentiviral vector

To create a stable cell line expressing LegK4, we decided to use lentivirus because they lead to a random stable integration in the eukaryotic genome of the transgene to be expressed [413]. Moreover, the expression of LegK4 would be induced using doxycycline to try to reduce its cytotoxicity.

The first step was to construct three lentiviral vectors derived from HIV-1, each crucial for lentivirus production. The plasmid psPAX2 vector allowed the expression of HIV gag-pol proteins, the plasmid V138, and the G protein of the VSV (vesicular stomatitis virus), and the plasmid pINDUCER expressed the gene of interest with HA tag. We created two different pINDUCERs: pINDUCER-*legK4* that contained *legK4* gene and pINDUCER-*legK4*<sub>D195N</sub> that expressed LegK4 without its kinase activity (kinase-dead mutant) (Figure 27).

HEK293T cells were then co-transfected with the three plasmids to produce the lentivirus (Figure 27). We used this specific cell line because of the presence of the SV40 T-antigen in the cells that rendered them more efficient for vector production. In addition, HEK293T cells showed a better cell growth and transfection efficiency than other cells. Only HEK293T cells transfected with the three plasmids could produce virions because each plasmid possessed an essential element to lentivirus production. Virions were produced 72 h post-transfection. They were collected and purified on a sucrose gradient.



**Figure 27: Creation of a stable cell lineage**

In HEK293T cells, the V138 plasmid allows the production of the VS G protein; psPAX2 allows the production of the proteins necessary for the nucleocapsid, reverse transcription and integration; and finally, pINDUCER allows the production of an mRNA encoding HA-LegK4 and GFP. After production of viral proteins and RNA, virions can assemble ①. Only cells with all three plasmids can produce virions ② that will be released and collected in culture medium ③. Virions were used to infect HeLa cells or THP1 monocytes. VSV G proteins on the surface of viruses allow their adhesion to the G protein-coupled receptor (GPCR) family of transmembrane proteins ④ and their fusion with the plasma membrane ⑤. Viral RNA was released into the cytoplasm where it was retro-transcribed into cDNA ⑥ and then randomly integrated into the genome ⑦ thanks to an integrase. Under induction with doxycycline, integrated genes were transcribed ⑧, and the protein of interest was produced ⑨.

### 3.3.1.2 Transduction and selection

We infected HeLa cells and THP-1 monocytes with lentivirus because these cells are the primary cells targeted by *Legionella* thanks to a collaboration with T. Henry's team (CIRI, Inserm U1111, CNRS UMR5308, UCBL, ENSL). Transduced cells were

then cultivated and sorted with fluorescence-activated cell sorting (BioAster, France). We collected only cells that expressed the GFP and, therefore, had integrated *legK4* (Figure 28A). Indeed, every cell that integrated the gene of interest expressed constitutively the GFP contained in the pINDUCER.

After sorting, an average of 70,000 HeLa cells and 4,000 THP-1 cells were harvested (Figure 28A). Nonadherent cells like THP-1 are usually harder to transduce than adherent cells like HeLa cells, which could explain the difference in transduction rate. Independently of the cell type, a much smaller number of cells integrated the *legK4* gene encoding a functional kinase, compared to the mutated gene *legK4<sub>D195N</sub>* (Figure 28A).

### 3.3.1.3 Expression of LegK4 in stable cell lines

We observed cells that have integrated *legK4* revealed a lower capacity for division and growth (internal lab data). LegK4 was probably expressed at low concentration, even without induction. Moreover, transduced THP-1 monocytes rapidly died after 24 h. For the rest of the experiment, we used HeLa cells that expressed HA-*LegK4<sub>D195N</sub>* because it was the only condition under which the cells did not die.

The first step was to determine the optimal concentration of the inducer, the doxycycline, to express LegK4. The optimal concentration was the lower concentration of doxycycline, which allowed the expression of HA-*LegK4<sub>D195N</sub>* without inducing excessive cytotoxicity.

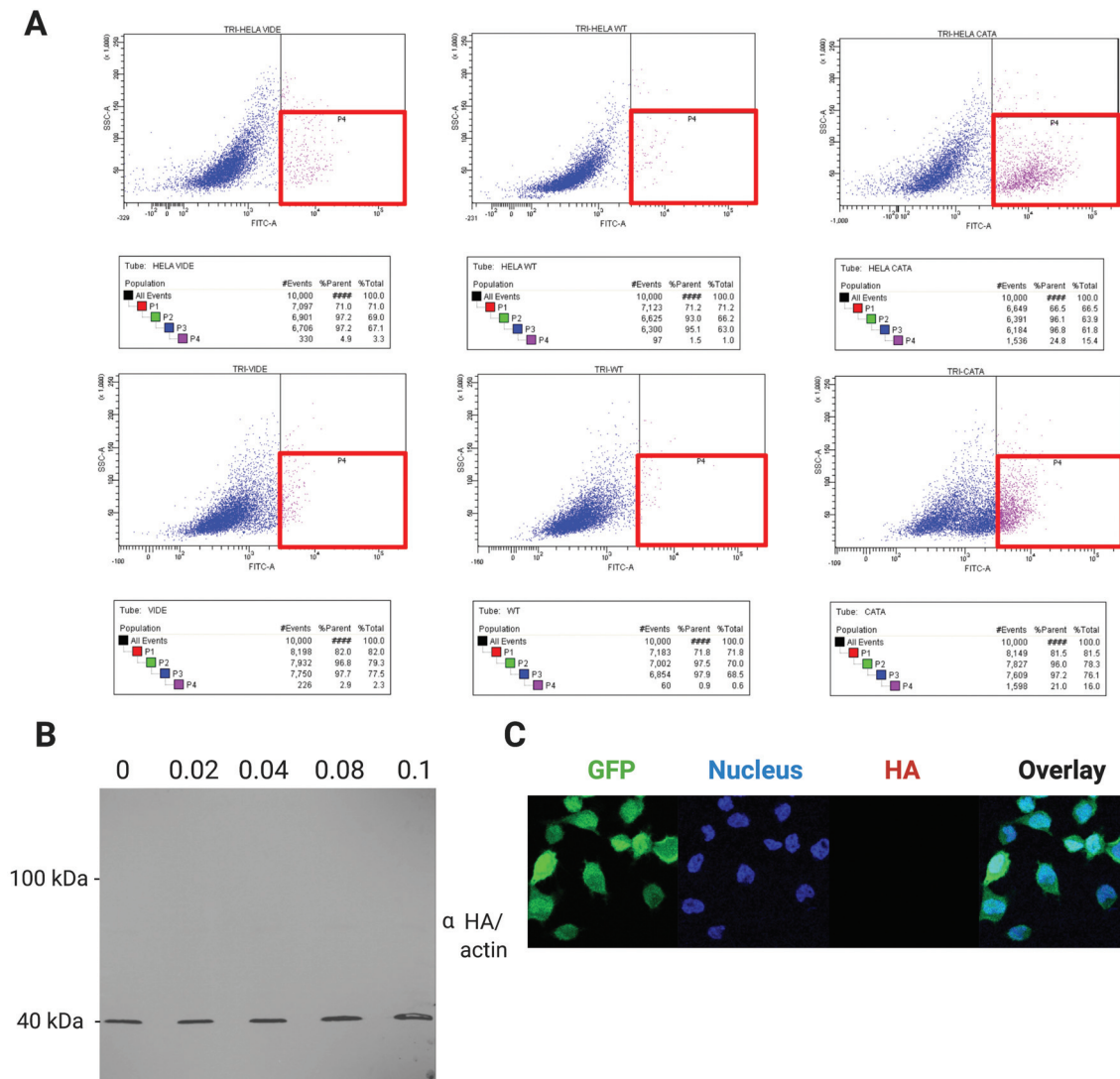
HeLa cells were incubated with 0 to 0.1  $\mu\text{g}/\text{mL}$  of doxycycline for 16 h (Figure 28B). Then, cells were lysed, and proteins were separated by SDS-PAGE followed by a western

blot. To identify the optimal concentration of doxycycline, an anti-HA antibody was used to detect HA-*LegK4*<sub>D195N</sub>. We also used an anti-actin antibody to detect actin as a loading control (Figure 28B). This control showed that the same amount of protein was loaded into each well.

In these conditions, we could not detect HA-*LegK4*<sub>D195N</sub> (109 kDa). Supplementary tests were also conducted with a new batch of doxycycline to verify whether the lack of detection was related to a failure of induction by the drug that was no longer active. Unfortunately, this did not change the result.

The transduced cells used were a polyclonal population sorted based on their GFP expression. The lack of signal could be due to an insufficient number of cells that expressed HA-*LegK4*<sub>D195N</sub>. The level of expression of the gene of interest could vary for many reasons. As an example, the number of inserts integrated into the genome and the site of integration of the transgene may have an impact on the gene of interest expression. Cells that efficiently expressed HA-*LegK4*<sub>D195N</sub> could be in the minority: these cells could be mixed with a large number of cells that expressed it at a lower level.

To identify the optimal concentration of doxycycline, we then determine the expression of HA-*LegK4*<sub>D195N</sub> cell by cell using immunofluorescence.



**Figure 28: HA-*LegK4*<sub>D195N</sub> induction with doxycycline**

A. HeLa cells and THP1 monocytes were sorted and collected if they expressed the GFP and thus contained the transgene (represented by the red box). Cell debris and clusters were removed. WT: pINDUCER21-*legK4* WT; Cata: pINDUCER21-*legK4*<sub>D195N</sub>; Blank: pINDUCER21.

B. Western blot after 16 h of induction with 0 to 0.1  $\mu\text{g}/\text{mL}$  of doxycycline. Samples were revealed with an anti- $\alpha$ -actin as a loading control (40 kDa) and an anti-HA antibody to detect HA-*LegK4* (100 kDa). C. Immunofluorescence after 16 h of induction with 0 to 0.1  $\mu\text{g}/\text{mL}$  of doxycycline. Nuclei were stained with DAPI (in blue) and HA-*LegK4*<sub>D195N</sub> with a primary anti-HA antibody and a secondary antibody Dylight 594-conjugated (in red).

Cells were treated as previously described with 0 to 0.1  $\mu\text{g}/\text{mL}$  of doxycycline for 16 h, and the expression of HA-*LegK4*<sub>D195N</sub> was observed by confocal microscopy (Figure 28C). As expected, all the cells expressed the GFP, and this confirmed the integration of the vector into their genome. Moreover, these cells exhibited different levels of green fluorescence, due probably to a difference in the GFP expression (Figure 28A and C). However, in our conditions, no signal was detected for HA-*LegK4*<sub>D195N</sub> (Figure 28C).

This absence of signal in western blot and immunofluorescence could be due to a failure to induce the expression of HA-*LegK4*<sub>D195N</sub> or the absence of *legK4* in transduced cells. We could increase the concentration of doxycycline and incubate cells even longer. Unfortunately, over 0.1  $\mu\text{g}/\text{mL}$  of doxycycline, cells died very quickly. Finally, to confirm the presence of *legK4*, we could consider sequencing the genome of the transduced cells.

### 3.2.2 Impact on gene expression

LegK4 nuclear localization and its substrates suggested a putative function in the modulation of the transcription by epigenetic control. RNA sequencing was performed during infection to determine if LegK4 could impact the gene transcription of infected macrophages.

To be in the most physiological experimental condition possible, we decided to examine the infection at low MOI (MOI 10). Macrophages were infected by the wild-type Paris strain (WT), the  $\Delta\textit{legK4}$  strain and the *legK4*<sub>D195N</sub> that expressed the kinase-dead mutant of LegK4 (*LegK4*<sub>D195N</sub>). All the infections were done in triplicate.



Because LegK4 was secreted around 15 min, we chose arbitrarily to do the RNA extraction at 45 min postinfection to give LegK4 enough time to migrate to the nucleus. RNAm purification and RNA sequencing were done by France g enomique (Strasbourg, France). Finally, we compared the results between each condition and kept only transcripts with a p-value lower than 0.05. Data sorting and analysis were done thanks to an pipeline made by T. Ohlmann (CIRI, INSERM U1111, ENS Lyon) and the help of C. Ginevra (INSERM U1111, CNRS, UM5308). We compared the conditions  $\Delta legK4$  versus  $legK4_{D195N}$ , WT versus  $\Delta legK4$  and WT versus  $legK4_{D195N}$ .

When we compared the results for  $\Delta legK4$  versus  $legK4_{D195N}$ , many transcripts presented exactly the same change fold. These results showed clearly that the kinase activity of LegK4 was responsible for the modulation of transcription in infected macrophages.

We then compared the results from  $\Delta legK4$  versus WT and, as expected, many genes were impacted.

The results showed that 289 genes were up-regulated more than two times, 193 were downregulated less than 0.5 times, and 20,972 were unchanged. Unfortunately, more than half of the transcripts found upregulated were lncRNA or ncRNA.

lncRNAs are a type of RNA defined as transcripts with lengths exceeding 200 nucleotides that are not translated into protein-like ncRNA. Unfortunately, to date, the role of these particular ncRNAs is not well known.

LegK4 seemed to impact many genes, so we focused on transcripts related to specific cellular pathways. We first inspected the transcripts related to the UPR because LegK4 could impact the UPR according to Moss *et al.* [269]. We also examined those related to apoptosis because LegK4 was observed to have a cytotoxic impact on eukaryotic cells,

and finally, arbitrarily, we examined those related to glycolysis.

### 3.3.2.1 UPR:

Overall, LegK4 did not have a critical impact on UPR-related transcripts: among 79 transcripts, only five were affected (Annexe 5). It negatively affected four of them: EXOSC8, KDELR3, SYVN1 and DDIT3 with a fold-change of 0.54, 0.49, 0.46 and 0.6, respectively. However, the transcript for the gene SERP1 was found to be expressed 1.76 more times in the wild-type strain than in  $\Delta legK4$ . SERP1 is known to protect unfolded target proteins against degradation during ER stress. It could help to decrease the ER stress provoked by the *Legionella* infection.

EXOSC8 encodes a 3'-5' exoribonuclease that specifically interacts with mRNAs containing AU-rich elements. This protein is part of the exosome complex that is important for the degradation of numerous RNA species.

KDELR3 encodes a member of the KDEL (Lys-Asp-Glu-Leu) ER protein retention receptor family. The retention of resident soluble proteins in the lumen of the ER is achieved by their continual retrieval from a cis-Golgi or a pre-Golgi compartment.

SYVN1 encodes a protein involved in ER-associated degradation. This protein removes unfolded proteins accumulated during ER stress by retrograde transport to the cytosol from the ER. It also uses the ubiquitin-proteasome system for additional degradation of unfolded proteins.

DDIT3 encodes a member of the CCAAT/enhancer-binding protein (C/EBP) family of transcription factors. It acts as a dominant-negative inhibitor by forming heterodimers with other C/EBP members, such as C/EBP and LAP (liver activator protein), and preventing their DNA binding activity. DDIT3 is activated by ERS and promotes apoptosis. Inhibiting the expression of DDIT3 could help *Legionella* to inhibit

apoptosis.

### 3.3.2.2 Apoptosis:

*Legionella* was described as negatively affecting apoptosis [241]. Indeed, LegK1 activates the host NF- $\kappa$ B signaling and plays important roles in modulating macrophage defense and inflammatory responses. Reducing apoptosis could be a strategy to keep infected cells alive as long as possible (Annexe 7).

NFKBIA, also called NFKB inhibitor alpha, was found 0.47 times less in the wild-type strain than in  $\Delta$ legK4. NFKBIA inhibits the activity of dimeric NF-kappa-B/REL complexes by masking the REL nuclear localization signal. Moreover, CD55/DAF (decay acceleration factor) transcript was found 0.32 times less when LegK4 was expressed.

DAF inhibits early complement activation by accelerating the degradation of C3 convertase, which is a central molecule that regulates the complement cascade. In addition, DAF prevents the assembly of C3 convertase, inhibiting complement activation. To sum, CD55 regulates both innate and adaptive immune responses [414]. LegK4 could, like LegK1, activate NF- $\kappa$ B signaling or inhibit complement activation to better modulate the macrophage defense and reduce inflammatory responses.

PIK3R3 (phosphatidylinositol 3-kinase regulatory subunit  $\gamma$ ) was found 1.7 more times while PIK3CD (phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit  $\Delta$ ) was reduced by 0.49.

PI3K phosphorylates phosphatidylinositol to produce PtdIns(3)P and similar compounds that then serve as second messengers in growth signaling pathways. Moreover, PI3K plays an important role in the regulation of cellular lipid metabolism. Increasing the production of PtdIns(3)P could be favorable for *Legionella* growth by improving the accessibility of particular lipids. However, PtdIns(3)P was also shown to be a crucial

element for the phagolysosomal pathway.

LegK4 could participate in the decrease of the endosome PtdIns(3)P level and lead to the recruitment inhibition of crucial proteins for membrane fusion such as EEA1 [143, 144].

### 3.3.2.3 Carbohydrate metabolism:

Finally, we determined if the expression of genes related to carbohydrate metabolism was modulated by LegK4 (Annexe 6).

We found that the 6-phosphofructo-2-kinase/ fructose-2,6-biphosphatase 2 (PFKFB2) and the glucosidase  $\alpha$  (GANC) encoding genes were up-regulated with a fold change of 3.43 and 2.99, respectively. PFKFB2 is involved in both the synthesis and degradation of fructose-2,6-bisphosphate. The fructose-2,6-biphosphate is a regulatory molecule that controls glycolysis in eukaryotes.

GANC is a key enzyme in glycogen metabolism. It hydrolyzes the glycosidic bond between two or more carbohydrates or between a carbohydrate and a non-carbohydrate moiety. This enzyme hydrolyses terminal, non-reducing 1,4-linked  $\alpha$ -D-glucose residues and releases  $\alpha$ -D-glucose.

At the same time, ALDOA (aldolase, fructose-bisphosphate A), LDHA (lactate dehydrogenase A) and LDHB (lactate dehydrogenase B) transcripts were found to be half their initial level. ALDOA is a glycolytic enzyme that catalyzes the reversible conversion of fructose-1,6-bisphosphate to glyceraldehyde 3-phosphate and dihydroxyacetone phosphate.

LDHA and LDHB are both lactate dehydrogenases (LDH) and catalyze the conversion of L-lactate and NAD to pyruvate and NADH in the final step of anaerobic glycolysis. Moreover, LDH enhances the production of the proinflammatory cytokine

IL-6 (interleukin 6) and  $\text{TNF}\alpha$  by J774.1 murine macrophages [415]. LegK4 could reduce the activation of the inflammatory response and apoptosis by reducing the release of proinflammatory cytokines.

### 3.2.3 Relationship between LegK4 and BiP

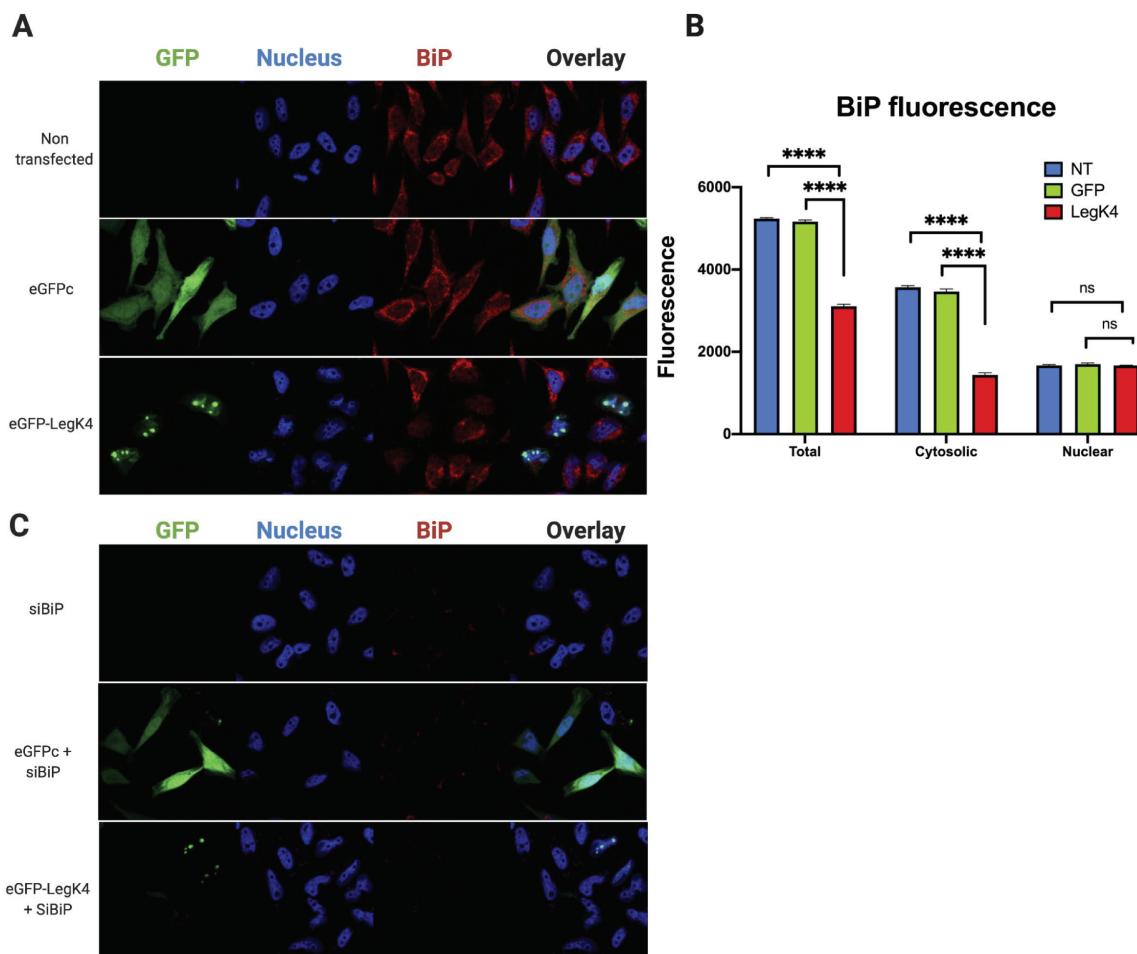
We previously demonstrated that LegK4 and BiP interacted in HeLa cells and aimed to better understand the impact of BiP on LegK4.

We transiently expressed the eGFP-LegK4 or the eGFPc in HeLa cells as a negative control. Then, we verified the presence of BiP by immunofluorescence.

Surprisingly, BiP vanished in cells expressing eGFP-LegK4 but was present in cells expressing the eGFPc and non-transfected cells (Figure 29A). These results showed that LegK4 seemed to have a strong impact on BiP.

Using confocal microscopy images and ImageJ software, we aimed to understand if BiP really vanished or was relocalized elsewhere. We quantified cytosolic and nuclear BiP using the Basic Intensity Quantification function of ImageJ (Figure 29B).

In non-transfected cells and those expressing the eGFPc, the same quantity of BiP was present in cells. However, the total quantity of BiP in cells expressing eGFP-LegK4 was significantly lower compared to the other conditions (Figure 29B). More precisely, it was the cytosolic and not nuclear BiP that was impacted by the presence of LegK4.



**Figure 29: Relationship between LegK4 and BiP**

A. HeLa cells were transiently transfected to produce eGFPc or eGFP-LegK4. Nuclei were stained blue with DAPI, and BiP was immunostained with an anti-BiP antibody (red).

B. Quantification using ImageJ of BiP fluorescence level in HeLa. BiP was quantified in the cytosol or the nucleus from 200 eGFPc-positive cells or non-transfected cells for the negative control. Each condition was done in triplicate. Error bars represent SEM. \*\*\*\* represents a highly significant difference ( $p < 0.001$ ).

C. HeLa cells were first transiently transfected to produce eGFPc or eGFP-LegK4 and then treated with siRNA anti-BiP. Nuclei were stained blue with DAPI, and BiP was detected with an anti-BiP antibody and immunostained using anti-IgG Dylight 594 (red).

These experiments showed that the expression of LegK4 has a strong impact only on the cytosolic BiP level but not on the nuclear BiP. Moreover, the expression of LegK4

did not have an impact on BiP mRNA during the infection of macrophages (Annexe 5). This result suggested that the BiP disappearance was not due to a lower transcription of BiP but rather a post-transcriptional mechanism induced by LegK4.

Interestingly, a protein without an NLS can migrate to the nucleus simply by interacting with another protein that has an NLS. BiP can be found in the cytosol and the nucleus, so we investigated whether LegK4 localization was related to BiP presence.

HeLa cells were treated with siRNA that targeted BiP mRNA (BiP siRNA) to substantially reduced BiP transcription on its intracellular level. As a negative control, we treated HeLa with a siRNA control that did not impact BiP. We then determined that the quantity of BiP was indeed decreased when treated with BiP siRNA by western blot. Finally, we transfected these cells with *peGFP-legK4* or *peGFP* as a negative control and determined LegK4 localization by confocal microscopy.

As expected, the BiP siRNA treatment had an impact on BiP quantity but did not change LegK4 localization in HeLa cells (Figure 29C). Even with a low concentration of BiP, LegK4 migrated to the nucleus with the same efficiency.

These results showed that BiP could not be responsible for LegK4 nuclear import.

We proved that LegK4 was not essential for *Legionella*'s growth and intracellular replication in amoebae or macrophages. Indeed, the deletion of the *legK4* gene into the *Legionella* genome did not impact its ability to infect human macrophages or amoebae.

Then, we created human monocytes and HeLa cells to express LegK4 under induction with doxycycline. Unfortunately, our stable cell line either died due to the LegK4 cytotoxicity or did not express LegK4. Until these cell lines are validated, we are unable to use them to better understand the impact of LegK4 on human cells.

By interacting and phosphorylating its substrates, LegK4 changed the gene expression of infected macrophages. Indeed, 289 genes were up-regulated more than two times, 193 were downregulated less than 0.5 times, 20.972 were unchanged, but many transcripts did not have a characterized function.

However, *Legionella* seemed to modulate the gene expression to its own benefit by inhibiting host defenses such as apoptosis and inflammation and by boosting the host metabolism to improve the production of nutrients.



## Part III

# Discussion and Conclusion

## Localization of LegK4

LegK4 is a kinase secreted by the pathogen *L. pneumophila*. This protein was found well conserved in five strains of *L. pneumophila*: Paris, Philadelphia, Lens, Corby and Alcoy. Here, we have demonstrated that LegK4 was addressed to the nucleolus of transfected mammalian cells. However, this result contradicted those from Moss *et al.* concerning LegK4 Philadelphia [269]. They demonstrated that LegK4 Philadelphia localized in the cytosol and not in the nucleus. Even though the peptidic sequences of LegK4s are very similar, LegK4 Philadelphia possesses a 58-amino-acid extension at the N-terminal region. We then determined that this extension was responsible for the accumulation of LegK4 Philadelphia into the cytosol. Without this extension, LegK4 Philadelphia migrated to the nucleus like LegK4 Paris. However, we proved that this extension was never produced by *Legionella*; thus, LegK4 Philadelphia as LegK4 Paris was nuclear. LegK4 is the fourth *L. pneumophila* effector to be localized in the nucleus and the second localizing to the nucleolus. Other nucleomodulins from *L. pneumophila* RomA and LegAS4 are histone methylases, and SnpL targets RNA polymerase. These proteins did not present sequence homology with LegK4 and do not yet have a characterized NLS.

The N-terminal Cap domain of LegK4 contained the following NLS **KVLGNKGI RKLKKSANGK**. This peptide sequence allowed LegK4 to be addressed to the nucleus. The expression of LegK4 without this NLS led to a change in its localization: instead of being localized in the nucleus, LegK4 was cytosolic. Unfortunately, the direct mutagenesis of one or all the basic amino acids presented in the NLS did not change the LegK4 localization.

Interestingly, this NLS allowed LegK4 to be nuclear but not nucleolar. Using bioin-

formatic softwares, such as NLStradamus [416] or NOD [417] (nucleolar localization sequence detector), no nucleolar localization signal (NoLS) could be found. However, LegK4 localization could probably be explained by the interaction and the phosphorylation of nucleolar substrates, allowing its nucleolar detention [418]. As an example, PP1, a nucleolar protein, does not contain a NoLS/NoDS localized in the nucleolus thanks to its interaction with NOM1 and its NoLS [389].

Many experiments were conducted to learn more about the migration of LegK4 to the nucleus. We inhibited the nuclear import machinery with drugs or used yeast mutants deleted for importins. Globally, these experiments were not conclusive, probably due to the highly cytotoxic effect of LegK4 expression in eukaryotic cells. This toxicity was also observed at the beginning of this project by Hybrigenics<sup>®</sup> during the yeast hybrid screening.

As Cassany *et al.* [419], we could consider reconstituting the nuclear import mechanism in permeabilized cells to determine how LegK4 migrates to the nucleus.

In permeabilized NRK (normal rat kidney epithelial) cells, we could add some element of the nuclear transport such as purified importin  $\alpha$  and/or  $\beta$ , Ran and eGFP-LegK4. Then, we could monitor LegK4 migration and localization by confocal microscopy. This experiment could provide the means to characterize the molecular mechanisms of LegK4 nuclear or nucleolar import.

LegK4 is known to be a dimeric, Hanks-like kinase able to dimerize and autophosphorylate [267]. Some nuclear proteins do not exhibit a cNLS, but their dimerization or the autophosphorylation allows their migration to the nucleus [420].

To determine whether LegK4 autophosphorylation is crucial for its localization, we could mutate the amino acids responsible for the dimerization and autophosphoryla-

tion. The catalytic mutant of LegK4, *LegK4<sub>D195N</sub>*, did not migrate to the nucleus with the same efficiency as LegK4. In transfected HeLa cells, eGFP-*LegK4<sub>D195N</sub>* could be localized in the nucleus, in the cytosol or in both. This could be a clue about the implication of the LegK4 kinase activity on its localization.

To complete LegK4 localization, we tried to establish it during the infection of human macrophages, pneumocytes and amoebae.

For this, we tried to detect LegK4 during the infection by immunofluorescence or western blot using an anti-LegK4 antibody we made. The identification of LegK4 localization by confocal microscopy during infection of human macrophages, pneumocytes or amoebae was not conclusive, probably due to the low sensitivity of the anti-LegK4 antibody. To increase the signal, we transformed *L. pneumophila* with a plasmid to produce more LegK4 protein, but that transformation did not change the result.

We also built a strain of *L. pneumophila* transformed with an inducible plasmid that expressed 4HA-tagged LegK4. After the infection of amoebae or macrophages, we could not detect LegK4 by immunofluorescence using an anti-HA antibody. To localize LegK4 during the infection of human cells, changing the timing of infection or the MOI was not successful. Identification of LegK4 localization by western blot was not conclusive because, when using a conventional centrifuge, the LCV containing bacteria and the nucleus pelleted at the same rotation speed, polluting our result.

To prove the nuclear localization of LegK4 during the infectious process, we could infect a consistent number of cells, then do a cellular fractionation followed by a western blot. If available, we could use ultracentrifugation to better separate the LCV from nuclei. Another option could be to lyse infected cells, collect them and catch LCVs using anti-LCV antibody on magnetic beads [421].

Before this study, a translocation assay determined that LegK4 was secreted around 15 min postinfection. Even though we knew when LegK4 was secreted, we did not know how long it would take for LegK4 to migrate to the nucleus.

Until we have more answers about the timing of infection, it will be challenging to do more experiments about LegK4 localization during the infection.

## LegK4 interactants and substrates

We demonstrated by co-immunoprecipitation that LegK4 could interact and phosphorylate BiP, a chaperon from the HSP70 family involved in the UPR. These results were consistent with those described by Moss *et al.* [269].

To go further, we investigated which residue of BiP was phosphorylated by LegK4. For this, we sent BiP from an *in vitro* phosphorylation assay and from the interaction test by GFP-Trap<sup>®</sup> to mass spectrometry. Surprisingly, even in the *in vitro* phosphorylation assay sample that contained purified BiP, we could not detect either a phosphopeptide or BiP. The treatment before the mass spectrometry probably led to a bad protein coverage, thus making it unable to detect BiP.

We must change the enzymatic treatment to produce larger peptides that could be detected by mass spectrometry.

In parallel, Moss *et al.* identified that BiP was phosphorylated by LegK4 on threonine 518 (T518) [269]. No antibody against this modification on BiP was available, but we could produce one like in Moss' *et al.* study [269].

In HeLa cells, we demonstrated that the expression of LegK4 has a strong impact on the BiP quantity in cells. When LegK4 was expressed in HeLa cells, the total quantity of BiP decreased significantly. More precisely, the cytosolic fraction of BiP vanished. Using ImageJ software, we determined that it was not a relocalization of BiP in the nucleus. In HeLa cells, the same quantity of nuclear BiP was detected when LegK4 was present or absent.

We also expressed the dead-kinase mutant of LegK4, *LegK4<sub>D195N</sub>*, in HeLa cells. *LegK4<sub>D195N</sub>* exhibited a hybrid localization and could be detected in the cytosol and in the nucleus. When nuclear *LegK4<sub>D195N</sub>* was found, we detected a lower quantity of cytosolic BiP. However, when *LegK4<sub>D195N</sub>* was found in the cytosol, BiP quantity did not change. This was proof that the catalytic activity of LegK4 seemed not to be responsible for the disappearance of BiP. More precisely, the localization of LegK4 or its interaction with BiP was related, to a certain extent, to BiP disappearance.

A comparison of the transcriptome from infected macrophages with a wild-type *Legionella* versus a  $\Delta$ *legK4* was conducted. The sequencing of mRNA proved that the BiP mRNA level did not change between these two conditions.

BiP disappearance was not due to transcriptional control by LegK4. It was more probably a post-transcriptional control caused by LegK4. Interaction tests in HeLa cells proved that LegK4 and BiP could interact. It would be interesting to decipher whether LegK4 interacted with the cytosolic or nuclear fraction of BiP. After cell fractionation, we could try to co-immunoprecipitate cytosolic or nuclear BiP with eGFP-LegK4.

To date, there are no known differences in terms of amino acids sequence between cytosolic and nuclear BiP. LegK4 is secreted straight to the host cytoplasm, where it could interact and phosphorylate cytosolic BiP before it migrates to the nucleus.

BiP is known to be a major ER stress sensor involved in the UPR. However, two

other proteins named CHOP and XBP1 are crucial for UPR. In this study, we showed that the transcription of these proteins was not impacted by the presence of LegK4 in infected macrophages. However, it could be interesting to follow the post-translational role of LegK4 on these proteins during the infection. After infection of macrophages with *L. pneumophila* WT and a  $\Delta legK4$ , we could extract proteins for a western blot analysis. Compared to the RNAseq, we could easily change many parameters, such as the MOI or the timing of infection.

We demonstrated that LegK4 could interact with H3 histone using co-immunoprecipitation.

We sent our sample to mass spectrometry to identify the phosphorylation site on H3 histone after its interaction with LegK4. Samples were treated with trypsin that cleaved the peptide chain at the carboxyl side of the amino acids lysine or arginine. Histones are basic proteins that contain many lysine residues. The treatment with trypsin produced many small peptides and decreased the protein coverage. The presence of too many small peptides led to the impossibility to detect a phosphopeptide

Then, an *in vitro* phosphorylation assay using purified H3 histone and 6His-LegK4 was conducted. By western blot, a phosphorylation signal was detected when H3 was mixed with LegK4. This phosphorylation was due to the kinase activity of LegK4 because no signal was detected with the kinase-dead LegK4 (*LegK4*<sub>D195N</sub>).

We sent a sample of H3 histone after its phosphorylation with LegK4 to mass spectrometry. This time, we changed the treatment protocol to produce larger fragments using chymotrypsin, thus increased protein coverage.

H3 histone was phosphorylated on threonine 32 residue (H3 T32P) by LegK4. Using an anti-H3 T32P antibody, a signal corresponding to a phosphorylated protein around 15 kDa was detected in the H3 phosphorylation assay samples.

Moreover, we also detected H3 T32P in the interaction test using eGFP-LegK4. No

signal was detected in the condition using eGFPc or eGFP-*LegK4*<sub>D195N</sub>.

To determine whether the phosphorylation signal we observed was only due to the threonine 32 residue, we mutated it into a non-phosphorylatable alanine (H3 T32A). We then overexpressed H3 and H3 T32A in *E. coli* and did a phosphorylation assay.

As expected, a signal was detected for H3 using an anti-phosphothreonine antibody.

However, no signal was detected when we used the mutated H3 T32A.

These different results suggest that the T32 residue seems to be the only residue phosphorylated by LegK4.

Finally, to complete this last result, we aimed to detect this phosphorylation in macrophages infected by *Legionella*.

Using confocal microscopy, no significant difference was observed between cells infected with a wild-type strain versus a  $\Delta$ *legK4*. The same result was observed by western blot.

A very little is known about H3 T32P, but it occurs during mitosis and meiosis in plants [405] and animals [422].

In addition to BiP and H3, we also found six new substrates of LegK4 by interaction test in HeLa cells.

Interestingly, these substrates were all nuclear proteins and related to transcription or RNA processing. They could be phosphorylated by LegK4 like they are by eukaryotic kinase.

Heterogeneous nuclear ribonucleoprotein U (HNRNPU) and heterogeneous nuclear ribonucleoproteins A2/B1 (HNRNPA2B1) are RNA-binding proteins that interact with heterogeneous nuclear RNA (hnRNA). HNRNPU and HNRNPA2AB1 are associated with pre-mRNAs in the nucleus and appear to influence pre-mRNA processing and other aspects of mRNA metabolism and transport. The mass spectrometry analysis revealed that LegK4 phosphorylated HNRNPU on the serine 4 (Ser4) and serine 271



(Ser271) residues and HNRNPA2B1 on the serine 189 (S189) residue. HNRNPU is phosphorylated by cellular kinase during mitosis [406].

Histone lysine acetyltransferase 7 (KAT7) is a histone acetyltransferase from HBO1 complexes that specifically mediate acetylation of histone H3 on lysine 14 residues (H3K14ac), thereby regulating processes such as gene transcription, protein ubiquitination, immune regulation [407, 408]. Mass spectrometry revealed that LegK4 phosphorylated this protein on threonine 88 (T88). Phosphorylation by CDK1 on T88 is associated with mitosis [409].

Nucleolin (NCL) is the major nucleolar protein in eukaryotic cells nucleolar and is involved in the synthesis and maturation of ribosomes [410]. It is thought to play a role in pre-rRNA transcription and ribosome assembly. Mass spectrometry analysis revealed that NCL was phosphorylated by LegK4 on the serine 67 residue (Ser67). NCL is known to be phosphorylated by cellular CDK2 on Ser67 during mitosis [410].

Nucleophosphomin (NPM1) possesses a bipartite NLS and a nucleolar localization signal at the C-terminal region. Mass spectrometry revealed that NPM1 was phosphorylated by LegK4 on the serine 4 residue (Ser4). When phosphorylated on Ser4, NPM1 is known to mediate the mitotic function of polo-like kinase, triggering centriole duplication [411]. The centriole duplication occurs at the end of the G1 phase of the cell cycle. Altogether, this information could indicate that LegK4 might have a role in the G1-to-S transition of the cell cycle.

Zinc finger protein 546 (ZNF546) is a potential nucleic acid-binding protein and transcriptional regulator. Mass spectrometry revealed that ZNF546 was phosphorylated by

LegK4 on the serine 373 (Ser373) residue; however, ZNF546 was not described as being phosphorylated on this residue by cellular kinase.

## The LegK4 and MavR relationship

We created a  $\Delta legK4$  mutant to better characterize the role of LegK4 in the infectious cycle of *L. pneumophila*. Unfortunately, the intracellular replication in macrophages or amoebae of *Legionella* was not impacted by the lack of *legK4*.

In the case of *Legionella*, traditional genetic approaches usually do not work because of functional redundancy. The simple deletion of an effector gene usually did not affect the *Legionella* infectious cycle.

Interestingly, the genomic organization of *legK4* could be a crucial clue to determine LegK4 function. The *legK4* gene (*lpp0267*) is located immediately upstream of another gene, *mavR* (*lpp0268*). Both are under the same promoter, forming an operon. MavR is an unknown protein with no described catalytic activity, but it possesses a high degree of similarity (around 41% of identity) with the C-terminal region of LegK4 (323 to 910 residue) (Figure 30).

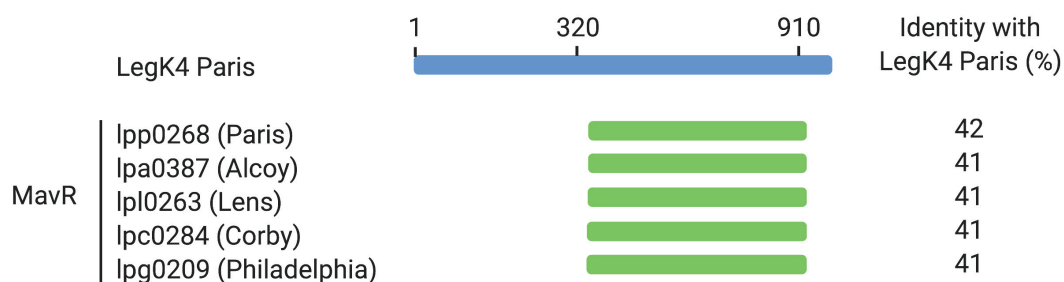


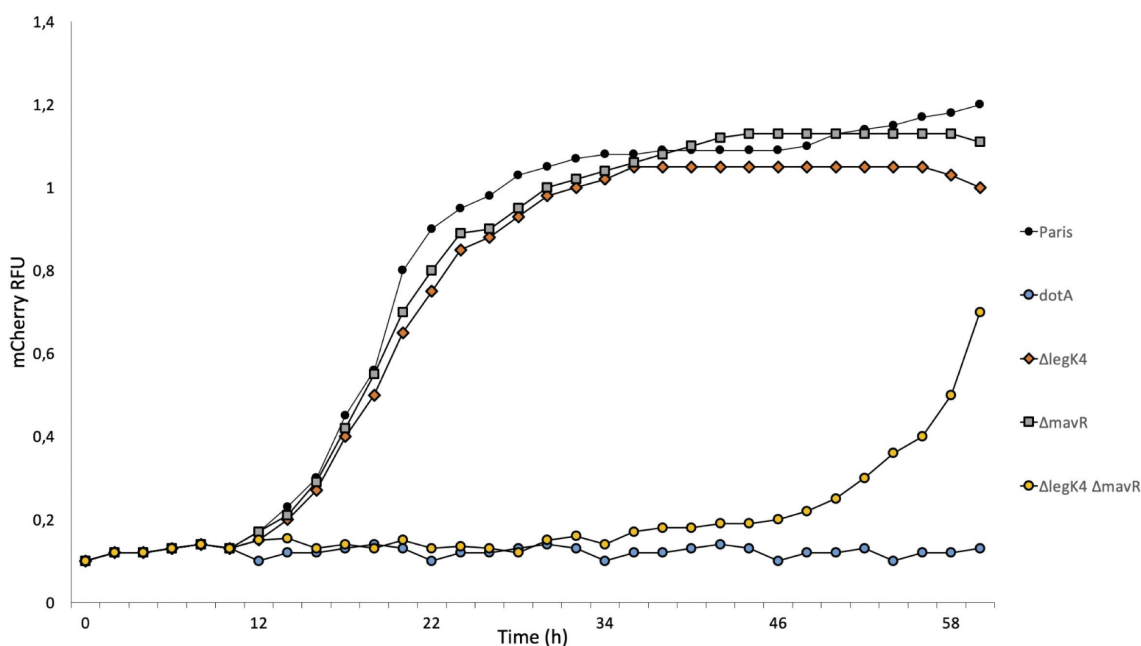
Figure 30: Comparison between LegK4 and MavR

Moreover, this genetic organization was found present in the five sequenced strains

of *L. pneumophila* Sg1. We infected macrophages and amoebae with  $\Delta legK4$ ,  $\Delta mavR$  or the double mutant  $\Delta legK4 \Delta mavR$ .

While a single  $\Delta legK4$  and  $\Delta mavR$  mutant had an intracellular replication similar to the wild-type strain, the  $\Delta legK4 \Delta mavR$  showed a significant replication delay (Figure 31).

These results were compared to the controls represented by the avirulent  $\Delta dotA$  mutant and the Paris wild-type strain. The phenotype exhibited by  $\Delta legK4 \Delta mavR$  was particularly interesting considering the high functional redundancy between *L. pneumophila* effectors and the low probability of observable phenotype for an effector mutant. This could mean that the *legK4-mavR* region would play a major role in the virulence of *L. pneumophila* during the infection.



**Figure 31: Intracellular growth of *L. pneumophila* in U937 macrophages**  
 Cells were infected at an MOI of 10 by *L. pneumophila* expressing the *mCherry* gene on a plasmid. Bacterial multiplication was automatically monitored with TECAN<sup>®</sup> by measuring the fluorescence of mCherry at an excitation of 587 nm and an emission of 610 nm every 2 h for 60 h. Fluorescence data were subjected to background subtractions (uninfected cells).

To determine the relationship between LegK4 and MavR, transfection assays were performed to localize these proteins in HeLa cells. While LegK4 was localized to the nucleolus, MavR stayed in the cytoplasm.

We also performed interaction tests between LegK4 and MavR, but they were unsuccessful. Even though they shared 41% of identity, MavR and LegK4 did not interact *in vitro* with each other.

Finally, the precise relationship between LegK4 and MavR could contribute to the replication of *L. pneumophila* in eukaryotic cells, but this remained to be established.

It would be interesting to determine precisely which stage of the *Legionella* infectious cycle is controlled by LegK4 and MavR. As an example, we could follow the entry of bacteria, the LCV biogenesis, the endocytic pathway escape or the bacterial egress. Moreover, the potential role of the LegK4/MavR tandem in UPR inhibition could also be investigated, given the interaction of LegK4 with BiP.

## Impact of LegK4 on human cells

To better understand the role of LegK4, we created HeLa cells and THP1 monocytes that expressed LegK4 on demand. Unfortunately, we could not validate whether these cells could produce LegK4.

The infection efficiency of the lentiviral used to express HA-LegK4 protein was cell-dependent, with much less transduced THP-1 than HeLa cells. Moreover, consistent with our previous results, the number of cells transduced by a vector expressing *legK4* was lower than those expressing the kinase-dead *LegK4<sub>D195N</sub>*. This was probably due to the cytotoxicity effect of the functional kinase, even in the absence of induction.

Similarly, the growth of cells expressing LegK4 was considerably reduced compared to those expressing the catalytic mutant.

The lack of signal could be due to a problem of revelation with the anti- HA antibody. However, we have previously validated the efficacy of our anti-HA using another HA-tagged protein.

Our protein may also lack the HA tag due to a shift in the reading frame during cloning. The input plasmid pENTRY1A-*legK4* was verified by sequencing; however, the final vector pINDUCER21-*legK4* was not. It was conceivable that a shift in the reading frame did not allow LegK4 to be tagged during the recombination reaction. Finally, the absence of a signal might not be related to the lack of LegK4 but to a high instability of LegK4 or its mRNA. LegK4 could be degraded as soon as it was expressed after induction in mammalian cells.

During the induction assays of HA-*LegK4*<sub>D195N</sub>, we observed that all the cells expressed the GFP. However, they exhibited variable levels of fluorescence. These differences could be explained by the number of copies of the insert and thus of the *GFP* gene that cells have integrated during transduction. Surprisingly, we failed to induce the production of the catalytic mutant of LegK4 in our preliminary trials. It should be noted that the expression of LegK4 in mammalian cells resulted in high cytotoxicity. In particular, LegK4 induced significant deformations of the nucleus forming many lobes (internal lab data). This phenotype appeared to be highly dependent on the nucleolar location of LegK4 and its kinase activity. The cytotoxicity due to LegK4 expression and its impact on the eventual initiation of apoptosis or the cell cycle must be characterized.

Considering the nuclear localization of LegK4 and its ability to phosphorylate proteins involved in gene transcription, it was conceivable that LegK4 could control the host transcription. It should be noted that LegAS4, another nucleomodulin from *L. pneumophila*, activates rDNA transcription when expressed in mammalian cells [245]. Unfortunately, after the RNA extraction, we discarded the rRNA. We only focused on

mRNA in our conditions, but it could be interesting to investigate other types of RNAs. The transcriptomic analysis of infected cell by *L. pneumophila* or a  $\Delta legK4$  gave us new information about LegK4 function.

First, LegK4 negatively impacted the expression of DDIT3, a protein involved in apoptosis. This protein is activated during ERS and can induce apoptosis and autophagy [423].

Second, LegK4 increased the transcription of SERP1. SERP1 protects unfolded proteins from degradation after ERS.

To create a replicative niche, *Legionella* recruits membranes from the ER and mitochondria on the LCV.

This phenomenon could be a stress event triggering the UPR, thus apoptosis. To survive and replicate, *Legionella* must inhibit the activation of both UPR and apoptosis. By inhibiting the expression of DDIT3 and increasing SERP1 expression, LegK4 could help *Legionella* to reduce ERS and inactivate apoptosis. It should be noted that between 5 and 79 genes related to the UPR were affected by LegK4.

Finally, some transcripts related to glycolysis or nutrition were upregulated. At the end of the *Legionella* cell cycle, the lack of nutrients and, in particular, the lack of amino acids is detected by bacteria [424].

Nutrient starvation induces the activation of the stringent response and morphological changes. Bacteria express transmissive traits such as motility (flagella) and become cytotoxic. These infectious bacteria can lyse the vacuolar membrane and are released in the extracellular environment. LegK4 could boost the metabolism of carbohydrates or lipids to avoid starvation, thus killing its host precociously.

LegK4 was secreted around 15 min postinfection; therefore, it could have a role early during macrophages infection. To create a replicative niche and thrive, *Legionella* must inhibit the host's defense, especially apoptosis. The secretion of LegK4 could help

reduce stress and inflammation at the beginning of the infection. Moreover, *Legionella* could modulate the host metabolism to improve the production of nutrients such as lipids or carbohydrates.

LegK4 is part of a group of effectors called nucleomodulins, which is a very active emerging research axis. A significant proportion of these proteins exhibit enzymatic activity that modifies nuclear targets, impacting the transcription of host genes. Since their discovery about ten years ago, they have been considered major virulence factors. They have a crucial role in the development of the infectious process of many pathogenic bacteria, such as *Salmonella* and *L. pneumophila*. This study led to a better characterization of the serine-threonine kinase LegK4, which is the first kinase from *L. pneumophila* to be identified as a nucleomodulin . Moreover, this study contributes to a better understanding of nucleomodulin and the molecular mechanism of the *L. pneumophila* infection.

# ANNEXES



**Table 1: List of characterized *Legionella pneumophila* effectors**

Effector	Alias	Activity	Localization	Target	Involved in	Ref.
AnkB	Ankyrin B, LegAU13, Ceg27	?	LCV	SKP1? PARVB?	Protein degradation, nutritional uptake?	[181] [180] [182] [183]
AnkX	Ankyrin X, LegA8, AnkN	Rab1-phosphocholine transferase	LCV	Rab1 Rab35	LCV biogenesis	[165] [162]
Ceg4	Coregulated with Efc factor Genes 4	Phosphotyrosine phosphatase	ER	Hog Fus3	Attenuate MAPK pathway activation	[425]
Ceg9	Coregulated with Efc factor Genes 9	?	?	Rtn4	Vesicle trafficking?	[219]
Ceg18	Coregulated with Efc factor Genes 18	?	Mitochondria	?	Caspase-3, activation NF-kB pathway	[230] [426]
Ceg19	Coregulated with Efc factor Genes 19	?	LCV, Late endosomes	Rab GTPases	Transport between Golgi, late endosomes and lysosomes?	[106]
CegC1	Coregulated with Efc factor Genes C1, PlcC	Zinc metallo phospholipase	?	Phospholipid	Virulence	[427]
GobX	Golgi-localized U-box E3 ligase, Lpg2455	E3 ubiquitin ligase	Golgi	S-palmitoylation and ubiquitination pathway	?	[275]
LecE	Lethal effect on yeast growth	?	LCV	PAH1	DAG synthesis	[136]
LegA9	AnkY, Ceg12	?	?	LCV	LCV recognition by autophagy	[428] [69]
LegAS4	Eukaryotic huntingtin interacting protein B, AnkI, RomA	Methyltransferase	Nucleus Nucleolus	histone H3 AROS	Chromatin remodelling, transcription repression	[258] [260] [246]
LegC2	Legionella eukaryotic-like gene C2, Y1FB	?	LCV?	SNARE-like protein	Endosomal pathway escape?	[216]

Effector	Alias	Activity	Localization	Target	Involved in	Ref.
LegC3	Legionella eukaryotic-like gene C3 PpeA	?	LCV?	SNARE complexes	Membrane fusion	[216] [270]
LegC7	Legionella eukaryotic-like gene C7, Y1fA	?	LCV?	VSP27? Emp46p/Emp47p	Endocytic evasion and alters ER morphology	[216] [109] [429]
LegG1	PieG, PeIE	Ran-GEF	LCV	Ran GTPases	Microtubules polymerization	[208] [209] [210]
LegK1	Legionella eukaryotic-like gene Kinase 1	Kinase	Cytosol	I $\kappa$ B $\alpha$ , p100	NF- $\kappa$ B pathway activation, apoptosis inhibition	[241] [240]
LegK2	Legionella eukaryotic-like gene Kinase 2	Kinase	LCV	ARP2/3 complex	Inhibition of actin polymerization	[268] [201]
LegK3	Legionella eukaryotic-like gene Kinase 3	Kinase	?	?	Vesicles trafficking pathway	[430]
LegK4	Legionella eukaryotic-like gene Kinase 4, Pkn5	Kinase	Nucleus Nucleolus	Hsp70s	Protein synthesis inhibition	[267] [268] [269]
LegK7	Legionella eukaryotic-like gene Kinase 7	Kinase	Cytosol	MOB1	Hippo pathway, promote infection	[431] [262]
LegL1		?	?	RavJ	?	[270]
LegS2	LpSL	Sphingosin-1-phosphate lyase	Mitochondria	Sphingolipids	Autophagy inhibition, caspase-3 activation, NF- $\kappa$ B pathway	[428] [229] [230]
LegU1		?	?	SKP1	Host ubiquitylation machinery modulation, Increase cell resistance	[182]
Lem3	Legionella Effector identified by Machine learning 3	Rab1-dephosphocholination	LCV	Rab1	LCV biogenesis, ER recruitment	[165]

Effector	Alias	Activity	Localization	Target	Involved in	Ref.
Lem4	Legionella Effector identified by Machine learning 4	Tyrosine phosphatase	Lysosome, Plasma membrane	Protein from LCV or plasma membrane	?	[432]
Lem12	Legionella Effector identified by Machine learning 12	?	Mitochondria	?	Autophagy inhibition, caspase-3 activation and NF- $\kappa$ B pathway	[230]
Lem21	Legionella Effector identified by Machine learning 21, LotA	Deubiquitinase	LCV	Several	Removal of ubiquitin from the LCVs, intracellular bacterial growth	[433] [434]
LepA	Legionella effector protein A	?	LCV?	?	Bacterial egress? Vesicular trafficking?	[257] [435]
LepB	Legionella effector protein B	Phosphoinositide kinase Rab1-GAP	LCV?	PtdIns(3)P, Rab1	<i>PtdIns</i> (3, 4) $P_2$ production, LCV biogenesis	[120] [121] [150] [163] [257] [435]
Lgt1	Legionella glucosyltransferase 1, LpGT	Glucosyl transferase	Cytosol	eEF1A	Protein synthesis inhibition NF- $\kappa$ B pathway, cell cycle arrest	[264] [172] [265] [231] [254] [242]
Lgt2	Legionella glucosyltransferase 2, LegC8	eEF1A glycosidase	Cytosol	eEF1A	Protein synthesis inhibition NF- $\kappa$ B pathway, cell cycle arrest	[264] [172] [231] [254] [242]
Lgt3	Legionella glucosyltransferase 3, LegC5	Glucosyltransferase	Cytosol	eEF1A	NF- $\kappa$ B pathway, cell cycle arrest	[264] [172] [254]
LicA		?	?	SKP1	Host ubiquitylation machinery modulation	[182]
LidA		Rab modulation on active state	LCV	Rab1 Rab6 GTPase	LCV biogenesis	[155] [153] [148] [160]

Effector	Alias	Activity	Localization	Target	Involved in	Ref.
LnaB	Legionella NF- $\kappa$ B activator B	?	Cytosol?	?	NF- $\kappa$ B pathway	[240]
LncP	Legionella nucleotide carrier protein	Nucleotide carrier	Mitochondria	?	Evacuate ATP from the membrane lumen (matrix)?	[196]
LpdA	Lipoamide dehydrogenase	PLD	LCV	Phospholipid	LCV biogenesis	[136] [137]
Lpg0393		Rab GEF	Golgi	Rab 5, Rab21 and Rab22	Vesicular trafficking modulation	[171]
Lpg1137		Serine protease	Mitochondria	STX17	Disrupt ER-mitochondria communication	[227] [228]
Lpg2160		?	?	SKP1	Modulation of host ubiquitylation machinery	[182]
Lpg2603		Kinase	LCV	?	? [436] //	
LpnE	<i>L. pneumophila</i> entry	?	LCV	OCRL1	PtdIns(4)P accumulation	[139] [42]
LppA		Phosphoinositide phosphatase	?	<i>PtdIns</i> (3, 4) $P_2$ , <i>PtdIns</i> (4, 5) $P_2$ , and <i>PtdIns</i> (3, 5) $P_2$	Phytate degradation, PtdIns(4)P accumulation	[126] [124] [47] [127]
LseA	Legionella SNARE effector A	?	Golgi-associated membranes	SNAREs	Regulation of fusion events?	[125] [217]
LtpD	<i>Legionella</i> translocated protein D	?	LCV	IMPA1	Lipid modulation ?	[146]
LtpM	<i>Legionella</i> translocated protein M	Glucosyl transferase	LCV, Rab5 and Rab7 containing endosome	?	Reduce microtubules mobility	[215]
LubX	<i>Legionella</i> U-box protein X, LegU2	E3-ubiquitin-ligase	?	SidH and host Clk1	Protein degradation Metaeffector	[272] [274] [273]
LupA	Legionella ubiquitin-specific protease A	Ubiquitin hydrolase	?	LegC3	?	[270]

Effector	Alias	Activity	Localization	Target	Involved in	Ref.
MavC	More regions allowing vacuole colocalization C	Transglutaminase	?	E2 UBE2N-Ub	Inhibits immune signaling	403
MavN	More regions allowing vacuole colocalization	Metal transporter	LCV	?	Iron acquisition	[194] [195]
MvcA	MavC paralog A, Lpg2148	Ubiquitin deamidase	?	E2 UBE2N-Ub	Promotes intracellular replication	[437]
PieE	Lem18	?	ER	Multiple Rab GTPases	Vesicular trafficking?	[438]
RalF	Recruitment of ARF to the <i>Legionella</i> phagosome	ARF-GEF	LCV	ARF1	Activation of PLD, LCV biogenesis	[128]
RavD	Region allowing vacuole colocalization D	?	LCV	PtdIns(3)P	Suppress endolysosomal maturation, contributes to lysosomal avoidance	[218]
RavK	Region allowing vacuole colocalization K	Metalloprotease	?	Actin	Inhibition of actin polymerization	[204]
RavN	Region allowing vacuole colocalization N	E3 ubiquitin ligase	?	?	Host ubiquitylation pathway.	[439]
RavJ	Region allowing vacuole colocalization J	?	?	Septin and elongator complexes	?	[270]
RavZ	Region allowing vacuole colocalization Z	Protease	Autophagosomes	ATG8 LC3	Autophagy inhibition, ubiquitin de-conjugation?	[184] [223] [225] [224]
RidL	Ceg28	?	LCV	VPS29	Retrograde trafficking inhibition	[212] [214]
SdcA	SdcC paralog	E3 ubiquitin ligase	LCV	Rab1	ER recruitment	[188]
SdeA	Sde substrate of the Dot/Icm system	Ubiquitylating/Deubiquitylating ADP-Ribosyltransferase	Autophagosomes	Rab1, Rab6a, Rab30 and Rab33b		[185] [186] [187]

Effector	Alias	Activity	Localization	Target	Involved in	Ref.
SdhA	Sdh succinate dehydrogenase	?	?	?	Innate immune defense control, Apoptosis inhibition	[440] [102] [441]
SetA	Su(var)3-9, Enhancer-of-zeste and Trithorax	Glycosyl transferase	LCV		Vesicle trafficking	[172] [106] [173]
SidC	Substrate of Icm/Dot transporter C	E3 ubiquitin ligase	LCV	Rab1	ER recruitment	[81] [188] [189] [190] [188]
SidD	Substrate of Icm/Dot transporter D	deAMPylase	LCV	GTPase RAB1	LCV biogenesis, ER recruitment, apoptosis and autophagy inhibition	[164] [165] [442]
SidE	Substrate of Icm/Dot transporter E, LatD	deubiquitination	LCV	mTORC1	mTORC1 inhibition	[104] [169] [443]
SidF	Substrate of Icm/Dot transporter F	Phosphoinositide phosphatase	?	BNIP3 et BCL-RAMBO (BCL2 family) PI(3,4)P2 et PI(3,4,5)P3	Inhibition de l'apoptose de l'hôte Pt-dIns4P accumulation	[122] [123]
SidG	Substrate of Icm/Dot transporter G	?	Membrane-associated?	?	?	[81] [93]
SidH	Substrate of Icm/Dot transporter H	?	?	?	LCV integrity	[81] [273] [272]
SidI	Substrate of Icm/Dot transporter I, Ceg32	?	Cytosol?	eEF1A and eEF1B $\gamma$	Innate immune defense control	[243] [242] [254]
SidJ	Substrate of Icm/Dot transporter J	Glutamylase	?	SidE family members	Inhibition of bacterial ubiquitin ligases by SidJ-calmodulin	[169] [276]
SidK	Substrate of Icm/Dot transporter K	?	LCV	VatA, VMA1	Inhibition of vacuole acidification Endocytic evasion	[115] [114]
SidL	Substrate of Icm/Dot transporter L Ceg14	Glycosyl Hydrolase	Cytosol?	F-Actin?	Actin polymerization inhibitor, translocation inhibition <i>in vitro</i>	[242] [266] [200] [254]

Effector	Alias	Activity	Localization	Target	Involved in	Ref.
SidM	Substrate of Icm/Dot transporter M, DrrA	RAB1-GEF, RAB1-GDF and RAB1-AMPylase	LCV	GTPase RAB1	LCV biogenesis ER recruitment	[148] [149] [150] [151] [152]
SidP	Substrate of Icm/Dot transporter P	PtdIns(3)-phosphatase	LCV?	PtdIns(3)P and PtdIns (3,5)P <sub>2</sub>	LCV biogenesis	[123] [270]
SusF	Suppressor of SidI Function, MesI	Metaeffector	?	SidI effector	SidI activity regulation	[277] [278] [279]
SnpL	Lpw27461, Lpg2519, Lpp2587	?	Nucleus	SUP15H/Spt5	Global upregulation of gene expression	[263]
VipA	VPS inhibitor proteins	Actin nucleator	Early endosomes	F-Actin	Hijack of actin cytoskeleton	[199] [444] [170] [198]
VipD	VipD VPS inhibitor proteins D	?	Mitochondria membrane, endosome	Rab5/Rab22	Inhibition of the fusion LCV-early endosomes, caspase-3 activation	[230] [143] [144] [145] [445]
VipF	VPS inhibitor proteins F	O and N-acetyltransferase	?	?	Endosomal degradation escape? Antibiotic resistance?	[170] [220]
VpdA		Phospholipase A	?	?	Membrane trafficking	[145]
VpdC		Phospholipase A	?	?	Membrane trafficking	[145]
WipA	WASP interacting protein A	Unconventional protein tyrosine phosphatase	?	N-WASP ARP2/3	Inhibition of actin polymerization	[202] [203]
WipB	WASP interacting protein B	Serine/Threonine phosphatase	?	LAMTOR1 v-ATPase	Cellular nutrient sensing and control of energy metabolism	[192]

Table 2: List of human HSP70

Protein name	Other protein names	Stress-inducible	Localization	Gene names	Ref.
HSPA1A/B	Heat shock 70 kDa protein 1A/1B, heat shock 70 kDa protein 1/2 (HSP70-1/HSP70-2, HSP70.1/HSP70.2), HSP72, HSP27, HSP28	Yes	Cytosol, Cell membrane, Nucleus, Lysosomes	HSPA1A, HSPA1, HSPA1B	[446] [299] [447] [448]
HSPA1L	Heat shock 70 kDa protein 1-like, heat shock 70 kDa protein 1L (HSP70-1L), heat shock 70 kDa protein 1-HOM (HSP70-HOM)	No	Cytosol, Nucleus	HSPA1L	[449]
HSPA2	Heat shock-related 70 kDa protein 2, heat shock 70 kDa protein 2 (HSP70-2)	No	Cytosol, Cell membrane, Nucleus	HSPA2	[450] [451]
HSPA4	Heat shock 70 kDa protein 4 (HSP70-4), HSP70RY, heat shock 70-related protein APG-2	No	Cytosol, Nucleus	HSPA4, APG2	[452] [453]
HSPA4L	Heat shock 70 kDa protein 4L, heat shock 70-related protein APG-1, osmotic stress protein 94 (OSP94)	Yes	Centrosome, Cytosol	HSPA4L, APG1, OSP94	[454]
HSPA5	78 kDa glucose-regulated protein (Grp78), heat shock 70 kDa protein 5 (HSP70-5), Immunoglobulin heavy chain-binding protein (BiP)	Yes	ER, Exosomes	HSPA5, GRP78	[455] [456] [457] [458]
HSPA6	Heat shock 70 kDa protein 6 (HSP70-6), heat shock 70 kDa protein B' (HSP70B')	Yes	Cytosol, Nucleus, Exosomes	HSPA6, HSP70B'	[447] [446]
HSPA7	Putative heat shock 70 kDa protein 7 (HSP70-7), heat shock 70 kDa protein B (HSP70B)	Yes	Blood microparticles, Exosomes	HSPA7, HSP70B	[459]
HSPA8	Heat shock cognate 71 kDa protein, heat shock 70 kDa protein 8, HSC70, HSP73	No	Cytosol, Cell membrane, Nucleus	HSPA8, HSC70, HSP73	[447] [293] [460]
HSPA9	Stress-70 protein, mitochondrial 75 kDa glucose-regulated protein (Grp75/mtHSP70), heat shock 70 kDa protein 9 (HSP70-9), mortalin (MOT), PBP74	No	Mitochondria, Nucleus	HSPA9, GRP75, HSPA9B	[303] [302]
HSPA13	Heat shock 70 kDa protein 13 (HSP70-13), microsomal stress-70 protein ATPase core, STCH	No	ER, Exosomes, Microsomes	HSPA13, STCH	[461] [462]



HSPA14	Heat shock 70 kDa protein 14 (HSP70-14), HSP70-like protein 1 (HSP70L1), heat shock protein HSP60	Yes	Mitochondria, Cytosol	HSPA14, HSP60, HSP70L1	<a href="#">463</a> <a href="#">464</a>
--------	---	-----	-----------------------	------------------------	---

Table 3: List of characterized NoLS

Protein name	NoLS sequence	Ref.
Angiogenin	IMRRRGL	<a href="#">465</a>
ApLLP	MAKSIRSKHRRQMRMMKRE	<a href="#">466</a>
ARF	QLRRPRHSHPTRARRCP	<a href="#">467</a>
B23	QDLWQWRKSL	<a href="#">468</a>
DEDD	KRPARGRATLGSQPKRRKSV	<a href="#">469</a>
DNA topo I	NKKKKPKKE	<a href="#">470</a>
DRIM	KKKMKKHKNKSEAKKRK	<a href="#">471</a>
FGF2	RSRKYTSWYVALKR	<a href="#">472</a>
FGF3	GKGVQPRRRRQKQSPDNLEP	<a href="#">473</a>
FXR2P	RPQRNRNRNRNRNR	<a href="#">474</a>
G2E3	RKHDDCPNKYGEKKTKEK	<a href="#">475</a>
GGNNV $\alpha$	RRRANNRRR	<a href="#">476</a>
GNL3	KRPKLKASKRMTCHKRYKIQKKVREHHRKLRLEAKKQGHKKPRK	<a href="#">477</a>
GNL3L	MMKLRHKNNKPGEGSKGHKKISWPYPQPAKQNGKATSKVPSAPHFVHPN	<a href="#">478</a>
HIC p40	GRCRRLANFGPRKRRRRR	<a href="#">479</a>
HIV-1 REV	RRNRNRNRWRERQRQI	<a href="#">480</a>
HIV-1 TAT	RKKRRQRRAHQ	<a href="#">481</a>
hLa	QESLNKWKSKGRRFKGGKGNKAAQPGSGKGGK	<a href="#">482</a>
HSP70	FKRKHKKDISQNKRAVRR	<a href="#">483</a>
HSV type 1 $\gamma(1)34.5$	MARRRRHRGPRRPRPP	<a href="#">472</a>
HSV ORF57	KRPR and RRPSRPFRKP	<a href="#">484</a>
hTERT	MPRAPRCRAVRSLLR	<a href="#">485</a>
I14L	MSRRNKRNRNRNRKPLNTIQ	<a href="#">486</a>
IBV N protein	WRRQARFK	<a href="#">487</a>
ING1b	DKPNSKRSRRQRNNENR and TPKEKKAKTSKSKKRSKAKA	<a href="#">477</a>
Inh3	HRKGRRR	<a href="#">488</a>
LA	QESLNKWKSKGRRFKGGKGNKAAQPGSGKGGK	<a href="#">482</a>
LIMK2	KKRTLKNDRKKR	<a href="#">420</a>
L1 ORF2	RLKIKGQRKIYQANGKQKK	<a href="#">489</a>
LYRIC	KSKKKKKKKKQGE and KQIKKKKKARRET	<a href="#">490</a>
MDM2	KKLKKRNK	<a href="#">491</a>
MDV MEQ protein	RRRKRNRDARRRRRQ	<a href="#">492</a>
MIDN	QQKRLRRKARRDARGPYHWSRKAARS	<a href="#">493</a>
MLF-1	MAPRGRRRPRPHRSEGARRSKNTLERTHS	<a href="#">494</a>
NIK	RKKRKKK	<a href="#">495</a>
NOLP	KEKIQAIIDSCRRQFPEYQERAR	<a href="#">496</a>

NOM1	1 to 269	<a href="#">389</a>
Nop25	<b><u>K</u>RKHP<b><u>R</u></b>RAQDS<b><u>S</u></b><b><u>K</u></b>PP<b><u>R</u></b>AP<b><u>R</u></b>TSKA<b><u>Q</u></b><b><u>R</u></b><b><u>R</u></b></b>	<a href="#">497</a>
NRG1	MSE <b><u>R</u></b> <b><u>K</u></b> <b><u>E</u></b> GRG <b><u>K</u></b> <b><u>G</u></b> <b><u>K</u></b> <b><u>G</u></b> <b><u>K</u></b> <b><u>K</u></b> ERGS <b><u>G</u></b> <b><u>K</u></b>	<a href="#">498</a>
Nuclear VCP-like protein	<b><u>K</u></b> <b><u>R</u></b> <b><u>K</u></b> <b><u>G</u></b> <b><u>K</u></b> L <b><u>K</u></b> N <b><u>K</u></b> <b><u>G</u></b> <b><u>S</u></b> <b><u>K</u></b> <b><u>R</u></b> <b><u>K</u></b> <b><u>K</u></b>	<a href="#">499</a>
Nucleolin	RGGGGGGGDF <b><u>K</u></b> <b><u>P</u></b> <b><u>Q</u></b> <b><u>G</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>T</u></b> <b><u>K</u></b> <b><u>F</u></b> <b><u>R</u></b>	<a href="#">500</a>
Nucleostemin	<b><u>K</u></b> <b><u>R</u></b> <b><u>P</u></b> <b><u>K</u></b> <b><u>L</u></b> <b><u>K</u></b> <b><u>K</u></b> AS <b><u>K</u></b> RM <b><u>T</u></b> CH <b><u>K</u></b> <b><u>R</u></b> <b><u>Y</u></b> <b><u>K</u></b> <b><u>I</u></b> <b><u>Q</u></b> <b><u>K</u></b> <b><u>V</u></b> RE <b><u>H</u></b> <b><u>R</u></b> <b><u>K</u></b> <b><u>L</u></b> <b><u>R</u></b> <b><u>L</u></b> <b><u>E</u></b> <b><u>A</u></b> <b><u>K</u></b> <b><u>Q</u></b> <b><u>G</u></b> <b><u>H</u></b> <b><u>K</u></b> <b><u>P</u></b> <b><u>R</u></b> <b><u>K</u></b>	<a href="#">501</a>
p120	<b><u>S</u></b> <b><u>K</u></b> <b><u>R</u></b> <b><u>L</u></b> <b><u>S</u></b> <b><u>S</u></b> <b><u>R</u></b> <b><u>A</u></b> <b><u>R</u></b> <b><u>K</u></b> <b><u>R</u></b> <b><u>A</u></b> <b><u>A</u></b> <b><u>K</u></b> <b><u>R</u></b> <b><u>R</u></b> <b><u>L</u></b> <b><u>G</u></b>	<a href="#">502</a>
P14ARF	<b><u>R</u></b> <b><u>R</u></b> <b><u>G</u></b> <b><u>A</u></b> <b><u>Q</u></b> <b><u>L</u></b> <b><u>R</u></b> <b><u>R</u></b> <b><u>P</u></b> <b><u>R</u></b> <b><u>H</u></b> <b><u>S</u></b> <b><u>H</u></b> <b><u>P</u></b> <b><u>T</u></b> <b><u>R</u></b> <b><u>A</u></b> <b><u>R</u></b> <b><u>R</u></b> <b><u>C</u></b> <b><u>P</u></b>	<a href="#">503</a>
PAPA-1	HGHGVH <b><u>K</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>H</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>H</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>H</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>H</u></b> <b><u>H</u></b>	<a href="#">504</a>
Parafibromin	<b><u>R</u></b> <b><u>R</u></b> <b><u>A</u></b> <b><u>A</u></b> <b><u>T</u></b> <b><u>E</u></b> <b><u>N</u></b> <b><u>I</u></b> <b><u>P</u></b> <b><u>V</u></b> <b><u>V</u></b> <b><u>R</u></b> <b><u>R</u></b> <b><u>P</u></b> <b><u>D</u></b> <b><u>R</u></b> <b><u>K</u></b> and <b><u>K</u></b> <b><u>K</u></b> <b><u>Q</u></b> <b><u>G</u></b> <b><u>C</u></b> <b><u>Q</u></b> <b><u>R</u></b> <b><u>E</u></b> <b><u>N</u></b> <b><u>E</u></b> <b><u>T</u></b> <b><u>L</u></b> <b><u>I</u></b> <b><u>Q</u></b> <b><u>R</u></b> <b><u>R</u></b> <b><u>K</u></b>	<a href="#">505</a>
PI4KA230	<b><u>S</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>T</u></b> <b><u>N</u></b> <b><u>R</u></b> <b><u>G</u></b> <b><u>S</u></b> <b><u>Q</u></b> <b><u>L</u></b> <b><u>H</u></b> <b><u>K</u></b> <b><u>Y</u></b> <b><u>M</u></b> <b><u>K</u></b> <b><u>R</u></b> <b><u>R</u></b> <b><u>T</u></b> <b><u>L</u></b>	<a href="#">477</a>
PML	<b><u>D</u></b> <b><u>R</u></b> <b><u>P</u></b> <b><u>L</u></b> <b><u>V</u></b> <b><u>F</u></b> <b><u>F</u></b> <b><u>D</u></b> <b><u>L</u></b> <b><u>K</u></b> <b><u>I</u></b> <b><u>D</u></b> <b><u>N</u></b>	<a href="#">506</a>
PPP1R11	<b><u>H</u></b> <b><u>R</u></b> <b><u>K</u></b> <b><u>G</u></b> <b><u>R</u></b> <b><u>R</u></b> <b><u>R</u></b>	<a href="#">477</a>
PRRSV N protein	<b><u>P</u></b> <b><u>G</u></b> <b><u>K</u></b> <b><u>N</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>N</u></b> <b><u>P</u></b> <b><u>E</u></b> <b><u>K</u></b> <b><u>P</u></b> <b><u>H</u></b> <b><u>F</u></b> <b><u>P</u></b> <b><u>L</u></b> <b><u>A</u></b> <b><u>T</u></b> <b><u>E</u></b> <b><u>D</u></b> <b><u>D</u></b> <b><u>V</u></b> <b><u>R</u></b> <b><u>H</u></b> <b><u>H</u></b> <b><u>F</u></b> <b><u>T</u></b> <b><u>P</u></b> <b><u>S</u></b> <b><u>E</u></b> <b><u>R</u></b>	<a href="#">507</a>
PTHrP	<b><u>G</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>G</u></b> <b><u>K</u></b> <b><u>P</u></b> <b><u>G</u></b> <b><u>K</u></b> <b><u>R</u></b> <b><u>R</u></b> <b><u>E</u></b> <b><u>Q</u></b> <b><u>E</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>R</u></b> <b><u>R</u></b> <b><u>T</u></b>	<a href="#">508</a>
PRV	<b><u>R</u></b> <b><u>R</u></b> <b><u>R</u></b> <b><u>R</u></b> <b><u>G</u></b> <b><u>G</u></b> <b><u>R</u></b> <b><u>G</u></b> <b><u>G</u></b> <b><u>R</u></b> <b><u>A</u></b> <b><u>A</u></b> <b><u>R</u></b>	<a href="#">384</a>
RECQL4	<b><u>K</u></b> <b><u>Q</u></b> <b><u>A</u></b> <b><u>W</u></b> <b><u>K</u></b> <b><u>Q</u></b> <b><u>K</u></b> <b><u>W</u></b> <b><u>R</u></b> <b><u>K</u></b> <b><u>K</u></b>	<a href="#">509</a>
RELA	<b><u>E</u></b> <b><u>Q</u></b> <b><u>P</u></b> <b><u>K</u></b> <b><u>Q</u></b> <b><u>R</u></b> <b><u>G</u></b> <b><u>M</u></b> <b><u>R</u></b> <b><u>F</u></b> <b><u>R</u></b> <b><u>Y</u></b> <b><u>K</u></b> <b><u>C</u></b> <b><u>E</u></b> <b><u>G</u></b> <b><u>R</u></b> <b><u>S</u></b> <b><u>A</u></b> <b><u>G</u></b> <b><u>S</u></b> <b><u>I</u></b> <b><u>P</u></b> <b><u>G</u></b> <b><u>E</u></b> <b><u>R</u></b>	<a href="#">510</a>
Rex	<b><u>P</u></b> <b><u>K</u></b> <b><u>T</u></b> <b><u>R</u></b> <b><u>R</u></b> <b><u>R</u></b> <b><u>P</u></b> <b><u>R</u></b> <b><u>R</u></b> <b><u>S</u></b> <b><u>Q</u></b> <b><u>R</u></b> <b><u>K</u></b> <b><u>R</u></b> <b><u>P</u></b>	<a href="#">511</a>
RPS7	<b><u>R</u></b> <b><u>R</u></b> <b><u>I</u></b> <b><u>L</u></b> <b><u>P</u></b> <b><u>K</u></b> <b><u>P</u></b> <b><u>T</u></b> <b><u>R</u></b> <b><u>K</u></b> <b><u>S</u></b> <b><u>R</u></b> <b><u>T</u></b> <b><u>K</u></b> <b><u>N</u></b> <b><u>K</u></b> <b><u>Q</u></b> <b><u>K</u></b> <b><u>R</u></b> <b><u>P</u></b> <b><u>R</u></b>	<a href="#">512</a>
RRP9	GQE <b><u>H</u></b> <b><u>R</u></b> <b><u>L</u></b> <b><u>G</u></b> <b><u>R</u></b> <b><u>W</u></b> <b><u>R</u></b> <b><u>I</u></b> <b><u>K</u></b> <b><u>E</u></b> <b><u>A</u></b> <b><u>R</u></b> <b><u>N</u></b> <b><u>S</u></b> <b><u>V</u></b> <b><u>C</u></b> <b><u>I</u></b> <b><u>P</u></b> <b><u>L</u></b> <b><u>R</u></b> <b><u>R</u></b> <b><u>V</u></b> <b><u>P</u></b> <b><u>P</u></b> <b><u>P</u></b> <b><u>A</u></b> <b><u>A</u></b> <b><u>G</u></b> <b><u>S</u></b>	<a href="#">477</a>
RPP29	<b><u>R</u></b> <b><u>H</u></b> <b><u>K</u></b> <b><u>R</u></b> <b><u>K</u></b> <b><u>E</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>A</u></b> <b><u>K</u></b> <b><u>L</u></b> <b><u>S</u></b> <b><u>A</u></b> <b><u>R</u></b> <b><u>Q</u></b> <b><u>t</u></b> <b><u>e</u></b> <b><u>x</u></b> <b><u>t</u></b> <b><u>f</u></b> <b><u>R</u></b> <b><u>R</u></b> <b><u>E</u></b> <b><u>L</u></b> <b><u>R</u></b>	<a href="#">513</a>
RPP38	<b><u>K</u></b> <b><u>I</u></b> <b><u>K</u></b> <b><u>L</u></b> <b><u>I</u></b> <b><u>P</u></b> <b><u>N</u></b> <b><u>P</u></b> <b><u>N</u></b> <b><u>K</u></b> <b><u>I</u></b> <b><u>R</u></b> <b><u>K</u></b> <b><u>P</u></b> <b><u>P</u></b> <b><u>K</u></b> <b><u>S</u></b> <b><u>K</u></b> <b><u>A</u></b> <b><u>T</u></b> <b><u>P</u></b> <b><u>K</u></b>	<a href="#">513</a>
SAP30L	<b><u>R</u></b> <b><u>R</u></b> <b><u>Y</u></b> <b><u>K</u></b> <b><u>R</u></b> <b><u>H</u></b> <b><u>Y</u></b> <b><u>K</u></b>	<a href="#">514</a>
STT3-B	<b><u>K</u></b> <b><u>Q</u></b> <b><u>K</u></b> <b><u>Y</u></b> <b><u>L</u></b> <b><u>S</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>T</u></b> <b><u>T</u></b> <b><u>K</u></b> <b><u>R</u></b> <b><u>K</u></b> <b><u>R</u></b> <b><u>G</u></b> <b><u>Y</u></b> <b><u>I</u></b> <b><u>K</u></b> <b><u>N</u></b> <b><u>K</u></b> <b><u>L</u></b> <b><u>V</u></b> <b><u>F</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>G</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>I</u></b> <b><u>S</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>T</u></b> <b><u>V</u></b>	<a href="#">515</a>
Survivin	M <b><u>Q</u></b> <b><u>R</u></b> <b><u>K</u></b> <b><u>P</u></b> <b><u>T</u></b> <b><u>I</u></b> <b><u>R</u></b> <b><u>R</u></b> <b><u>K</u></b> <b><u>N</u></b> <b><u>L</u></b> <b><u>R</u></b> <b><u>L</u></b> <b><u>R</u></b> <b><u>R</u></b> <b><u>K</u></b>	<a href="#">516</a>
TCOF1	<b><u>K</u></b> <b><u>R</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>D</u></b> <b><u>K</u></b> <b><u>E</u></b> <b><u>K</u></b> <b><u>E</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>A</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>A</u></b> <b><u>S</u></b> <b><u>T</u></b> <b><u>K</u></b> <b><u>D</u></b> <b><u>S</u></b> <b><u>E</u></b> <b><u>S</u></b> <b><u>P</u></b> <b><u>S</u></b> <b><u>Q</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>T</u></b> <b><u>A</u></b> <b><u>E</u></b> <b><u>Q</u></b> <b><u>T</u></b> <b><u>V</u></b>	<a href="#">517</a>
U3-55K	GQE <b><u>H</u></b> <b><u>R</u></b> <b><u>L</u></b> <b><u>G</u></b> <b><u>R</u></b> <b><u>W</u></b> <b><u>R</u></b> <b><u>I</u></b> <b><u>K</u></b> <b><u>E</u></b> <b><u>A</u></b> <b><u>R</u></b> <b><u>N</u></b> <b><u>S</u></b> <b><u>V</u></b> <b><u>C</u></b> <b><u>I</u></b> <b><u>P</u></b> <b><u>L</u></b> <b><u>R</u></b> <b><u>R</u></b> <b><u>V</u></b> <b><u>P</u></b> <b><u>P</u></b> <b><u>P</u></b> <b><u>A</u></b> <b><u>A</u></b> <b><u>G</u></b> <b><u>S</u></b>	<a href="#">518</a>
Werner	V <b><u>S</u></b> <b><u>R</u></b> <b><u>Y</u></b> <b><u>N</u></b> <b><u>K</u></b> <b><u>F</u></b> <b><u>M</u></b> <b><u>K</u></b> <b><u>I</u></b> <b><u>C</u></b> <b><u>A</u></b> <b><u>L</u></b> <b><u>T</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>G</u></b> <b><u>R</u></b> <b><u>N</u></b> <b><u>W</u></b> <b><u>L</u></b> <b><u>H</u></b> <b><u>K</u></b> <b><u>A</u></b> <b><u>N</u></b> <b><u>T</u></b> <b><u>E</u></b> <b><u>S</u></b>	<a href="#">519</a>
ZNF330	<b><u>M</u></b> <b><u>P</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>T</u></b> <b><u>G</u></b> <b><u>A</u></b> <b><u>R</u></b> <b><u>K</u></b> <b><u>K</u></b> <b><u>A</u></b> <b><u>E</u></b> <b><u>N</u></b> <b><u>R</u></b> <b><u>R</u></b> <b><u>E</u></b> <b><u>R</u></b> <b><u>E</u></b> <b><u>K</u></b> <b><u>Q</u></b> <b><u>L</u></b> <b><u>R</u></b> <b><u>A</u></b> <b><u>S</u></b> <b><u>R</u></b> <b><u>S</u></b> <b><u>T</u></b> <b><u>I</u></b> <b><u>D</u></b> <b><u>L</u></b> <b><u>A</u></b> <b><u>K</u></b>	<a href="#">520</a>

Letter in bold and underlined letters represent respectively basic amino acid and succession of two or more basic residue.

Protein	Phosphorylated residue by LegK4	Molecular weight	Localization	Protein function	Phosphorylation role
p120 / Heterogeneous nuclear ribonucleoprotein U	[S4] / [S271]	102 kDa	Nucleoplasm, Centrosome	Nuclear chromatin organization, DNA- and RNA-binding protein involved in several cellular processes such as nuclear chromatin organization, transcription, mRNA alternative splicing and stability and mitotic cell progression	Found phosphorylated during cytokinesis / Found phosphorylated during mitosis
Heterogeneous nuclear ribonucleoproteins A2/B1	[S189]	42 kDa	Nucleoplasm	Condense and stabilize the transcripts and minimize tangling and knotting	Found in cancer cell phosphoproteome
Histone acetyltransferase KAT7	[T88]	73 kDa	Nucleoplasm, Cytosol	Component of the HBO1 complex which has a histone H4-specific acetyltransferase activity, a reduced activity toward histone H3. May regulate DNA replication and acts as a coactivator of the licensing factor CDT1	Phosphorylate on T88 by CDK1 during mitosis
Nucleolin	[S67]	85 kDa	Nucleoli, Nucleoplasm	Chromatin decondensation by binding to histone H1. Could play a role in pre-rRNA transcription and ribosome assembly	Phosphorylate on S67 by CDK2 during mitosis
Nucleophosmin	[S4]	35 kDa	Nucleoli, Nucleoplasm	Chaperone, RNA-binding	Phosphorylate on S4 by GRK5, PLK2, PLK3, PLK1, mediates mitotic functions of polo-like kinase triggering centriole duplication
Zinc finger protein 546	[S373]	100 kDa	Nucleoplasm, Cytosol	May be involved in transcriptional regulation	nd

Table 4: *In cellula* LegK4 substrates

**Table 5: Modulation of UPR-related genes**

Gene Identifier	Name	p value	Fold-change
ACADVL	acyl-CoA dehydrogenase, very long chain	0,562353964	0,859069803
ADD1	adducin 1 ( $\alpha$ )	0,643981545	0,913129449
ARFGAP1	ADP-ribosylation factor GTPase activating protein 1	0,778214192	0,900021835
ASNS	asparagine synthetase (glutamine-hydrolyzing)	0,240620862	0,69911795
ATF3	activating transcription factor 3	0,243564472	1,418035087
ATF4	activating transcription factor 4	0,362910987	0,777533305
ATF6	activating transcription factor 6	0,511455688	1,202897249
ATP6V0D1	ATPase, H <sup>+</sup> transporting, lysosomal 38kDa, V0 subunit d1	0,54325404	1,257796432
CALR	calreticulin	0,133244222	0,773186294
CCL2	chemokine (C-C motif) ligand 2	0,695279167	1,108323322
CTDSP2	CTD (carboxy-terminal domain) small phosphatase 2	0,114764576	1,759672222
CUL7	cullin 7	0,899832371	0,973830756
CXXC1	CXXC finger protein 1	0,936706935	1,030895867
DCP2	decapping mRNA 2	0,527980015	0,787373737
DCTN1	dynactin 1	0,213802607	0,622429412
DDIT3	DNA-damage-inducible transcript 3	0,005422048	0,610552256
DDX11	DEAD/H (Asp-Glu-Ala-Asp/His) box helicase 11	0,661517272	0,879625912
DIS3	DIS3 exosome endoribonuclease and 3'-5' exoribonuclease	0,134890211	0,786539204
DNAJB11	DnaJ (Hsp40) homolog, subfamily B, member 11	0,571245374	0,841384721
DNAJB9	DnaJ (Hsp40) homolog, subfamily B, member 9	0,384600025	0,723298093
DNAJC3	DnaJ (Hsp40) homolog, subfamily C, member 3	0,486962095	1,255931799
EDEM1	ER degradation enhancer, mannosidase $\alpha$ -like 1	0,625562905	1,197506998
EIF2AK3	eukaryotic translation initiation factor 2- $\alpha$ kinase 3	0,535979035	1,256837532
EIF2S1	eukaryotic translation initiation factor 2, subunit 1 $\alpha$ , 35kDa	0,361037613	0,759507319
ERN1	endoplasmic reticulum to nucleus signaling 1	0,570407441	1,20896575
EXOSC1	exosome component 1	0,986467786	0,994072287
EXOSC2	exosome component 2	0,67651671	0,853888589
EXOSC3	exosome component 3	0,184094199	0,703353469
EXOSC4	exosome component 4	0,793241073	0,952000501
EXOSC5	exosome component 5	0,435827493	0,778271059
EXOSC6	exosome component 6	0,961279535	0,987451512
EXOSC7	exosome component 7	0,531755565	0,793736761
<b>EXOSC8</b>	<b>exosome component 8</b>	<b>0,000912531</b>	<b>0,540665162</b>
EXOSC9	exosome component 9	0,632447603	0,92136244
EXTL3	exostosin-like glycosyltransferase 3	0,647207773	1,179601988
FKBP14	FK506 binding protein 14, 22 kDa	0,571758653	1,191143407

GFPT1	glutamine-fructose-6-phosphate transaminase 1	0,9029824	0,956960448
GOSR2	golgi SNAP receptor complex member 2	0,973609065	0,994034229
GSK3A	glycogen synthase kinase 3 $\alpha$	0,491738893	0,836690294
HDGF	hepatoma-derived growth factor	0,3001073	0,754000878
HERPUD1	homocysteine-inducible, endoplasmic reticulum stress-inducible, ubiquitin-like domain member 1	0,74473197	0,882735091
HSP90B1	heat shock protein 90kDa $\beta$ (Grp94), member 1	0,959132615	0,988902302
HSPA5	heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa)	0,643981545	0,913129449
HYOU1	hypoxia up-regulated 1	0,647742679	1,189482856
IGFBP1	insulin-like growth factor binding protein 1	NA	NA
IL27	myeloid-derived growth factor	0,136179464	0,572344315
IL8	chemokine (C-X-C motif) ligand 8	0,92702641	1,017159852
<b>KDELR3</b>	<b>KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 3</b>	<b>0,000945612</b>	<b>0,490990847</b>
KHSRP	KH-type splicing regulatory protein	0,22262546	0,686978562
KLHDC3	kelch domain containing 3	0,328936263	0,690290412
LMNA	lamin A/C	0,712497303	1,108423368
MBTPS1	membrane-bound transcription factor peptidase, site 1	0,574672466	0,825115592
MBTPS2	membrane-bound transcription factor peptidase, site 2	0,932758529	0,972956724
NFYA	nuclear transcription factor Y, $\alpha$	0,247222228	1,250930133
NFYB	nuclear transcription factor Y, $\beta$	0,918748582	0,961724266
NFYC	nuclear transcription factor Y, $\gamma$	0,504656889	1,290342678
PARN	poly(A)-specific ribonuclease	0,807071919	1,09763229
PDIA5	protein disulfide isomerase family A, member 5	0,970747426	0,988129458
PDIA6	protein disulfide isomerase family A, member 6	0,878001314	0,959579973
PLA2G4B	phospholipase A2, group IVB (cytosolic)	NA	NA
PPP2R5B	protein phosphatase 2, regulatory subunit B', $\beta$	NA	NA
PREB	prolactin regulatory element binding	0,486651528	0,766660896
<b>SEC31A</b>	<b>SEC31 homolog A (<i>S. cerevisiae</i>)</b>	<b>0,536621305</b>	<b>1,251488277</b>
<b>SERP1</b>	<b>stress-associated endoplasmic reticulum protein 1</b>	<b>0,010468141</b>	<b>1,759188674</b>
SHC1	SHC (Src homology 2 domain containing) transforming protein 1	0,835736521	1,076165982
SRPR	signal recognition particle receptor (docking protein)	0,179714209	1,254264698
SRPRB	signal recognition particle receptor, B subunit	0,74366347	0,882517384
SSR1	signal sequence receptor, $\alpha$	0,738019377	1,099940099
SULT1A3	sulfotransferase family, cytosolic, 1A, phenol-preferring, member 3	NA	NA
<b>SYVN1</b>	<b>synovial apoptosis inhibitor 1, synoviolin</b>	<b>0,000156233</b>	<b>0,462722605</b>

TATDN2	TatD DNase domain containing 2	0,271651394	1,282328416
TLN1	talin 1	0,712085779	0,893726271
TPP1	tripeptidyl peptidase I	0,441866837	0,785001493
TSPYL2	TSPY-like 2	0,801144685	1,063595847
WFS1	Wolfram syndrome 1 (wolframin)	0,973609065	0,994034229
WIP1	WD repeat domain, phosphoinositide interacting 1	0,345776812	1,23639039
XBP1	X-box binding protein 1	0,631289884	1,160436279
YIF1A	Yip1 interacting factor homolog A	0,500827683	0,779915276
ZBTB17	zinc finger and BTB domain containing 17	0,911280242	1,043436274

Gene expression with a significant fold-change (p-value < 0.05) are represented in bold

**Table 6: Modulation of Glycolysis-related genes**

Gene Identifier	Name	p value	Fold-change
ACSS1	Acyl-CoA Synthetase Short Chain Family Member 1	0,256519913	1,489876639
ACSS2	Acyl-CoA Synthetase Short Chain Family Member 2	0,449742452	1,266688063
ADH1A	Alcohol Dehydrogenase 1A (Class I), $\alpha$ Polypeptide	NA	NA
ADH1B	Alcohol Dehydrogenase 1A (Class I), $\beta$ Polypeptide	NA	NA
ADH1C	Alcohol Dehydrogenase 1A (Class I), $\gamma$ Polypeptide	NA	NA
ADH4	Alcohol Dehydrogenase 4 (Class II), Pi Polypeptide	NA	NA
ADH5	Alcohol Dehydrogenase 5 (Class III), Chi Polypeptide	0,924535841	1,036956376
ADH6	Alcohol Dehydrogenase 6 (Class V)	0,517852205	0,847576165
ADH7	Alcohol Dehydrogenase 7 (Class IV), Mu Or Sigma Polypeptide	NA	NA
ADPGK	ADP Dependent Glucokinase	0,646082339	1,122903249
AKR1A1	Aldo-Keto Reductase Family 1 Member A1	0,176563972	0,615132198
ALDH1A3	Aldehyde Dehydrogenase 1 Family Member A3	0,172569932	1,678445022
ALDH1B1	Aldo-Keto Reductase Family 1 Member B	0,97437948	0,991405664
ALDH2	Aldehyde Dehydrogenase 2 Family Member	0,63869946	0,91858333
ALDH3A1	Aldehyde Dehydrogenase 3 Family Member A1	NA	NA
ALDH3A2	Aldehyde Dehydrogenase 3 Family Member A2	0,973609065	0,994034229
ALDH3B1	Aldehyde Dehydrogenase 3 Family Member B1	0,68023377	1,170579102
ALDH3B2	Aldehyde Dehydrogenase 3 Family Member B2	0,973609065	0,994034229
ALDH7A1	Aldehyde Dehydrogenase 7 Family Member A1	NA	NA
ALDH9A1	Aldehyde Dehydrogenase 9 Family Member A1	NA	NA
<b>ALDOA</b>	<b>Aldolase, Fructose-Bisphosphate A</b>	<b>0,03399129</b>	<b>0,593636434</b>
ALDOB	Aldolase, Fructose-Bisphosphate B	0,506428062	1,2571072
ALDOC	Aldolase, Fructose-Bisphosphate C	0,696626141	0,899130436
BPGM	Bisphosphoglycerate Mutase	0,716123288	0,876243219
DLAT	Dihydrolipoamide S-Acetyltransferase	0,68124522	0,884388435
DLD	Dihydrolipoamide Dehydrogenase	0,991065626	0,99601258
ENO1	Enolase 1	0,185877028	0,739393682
ENO2	Enolase 2	0,686688105	1,075682538
ENO3	Enolase 3	0,184348743	0,689225166
FBP1	Fructose-Bisphosphatase 1	0,853122582	1,034052833
FBP2	Fructose-Bisphosphatase 2	NA	NA
G6PC1	Glucose-6-Phosphatase Catalytic Subunit 1	NA	NA
G6PC2	Glucose-6-Phosphatase Catalytic Subunit 2	NA	NA
GALM	Galactose Mutarotase	0,059658222	0,487064558
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase	0,054381996	0,62533619
GAPDHS	Glyceraldehyde-3-Phosphate Dehydrogenase, Spermatogenic	0,593490751	1,177244668



GCK	Glucokinase	0,532857657	1,249186739
GPI	Glucose-6-Phosphate Isomerase	0,831304962	0,933207437
HK1	Hexokinase 1	0,591572794	1,148836044
HK2	Hexokinase 2	0,114030196	0,627460004
HK3	Hexokinase 3	0,466917453	0,776502954
HKDC1	Hexokinase Domain Containing 1	0,973609065	0,994034229
<b>LDHA</b>	<b>Lactate Dehydrogenase A</b>	<b>0,01477692</b>	<b>0,486288772</b>
LDHAL6A	Lactate Dehydrogenase A Like 6A	0,92937402	0,966631435
<b>LDHAL6B</b>	<b>Lactate Dehydrogenase A Like 6B</b>	<b>0,01994434</b>	<b>1,971526432</b>
<b>LDHB</b>	<b>Lactate Dehydrogenase B</b>	<b>0,00649496</b>	<b>0,46097376</b>
LDHC	Lactate Dehydrogenase C	0,853122582	1,034052833
PANK1	Pantothenate Kinase 1	0,698018242	0,901029684
PCK1	Phosphoenolpyruvate Carboxykinase 1	NA	NA
PCK2	Phosphoenolpyruvate Carboxykinase 2, Mitochondrial	0,929230109	1,03456083
PDHA1	Pyruvate Dehydrogenase E1 Subunit $\alpha$ 1	0,782957989	0,93934376
PDHA2	Pyruvate Dehydrogenase E1 Subunit $\alpha$ 2	NA	NA
PDHB	Pyruvate Dehydrogenase E1 Subunit $\beta$	0,374621795	0,874852214
PFKFB1	6-Phosphofructo-2-Kinase/Fructose-2,6-Biphosphatase 1	NA	NA
<b>PFKFB2</b>	<b>6-Phosphofructo-2-Kinase/Fructose-2,6-Biphosphatase 2</b>	<b>1,25E-05</b>	<b>3,430422203</b>
PFKFB3	6-Phosphofructo-2-Kinase/Fructose-2,6-Biphosphatase 3	0,30722001	1,478430003
PFKFB4	6-Phosphofructo-2-Kinase/Fructose-2,6-Biphosphatase 4	0,518479163	0,805786678
PFKL	Phosphofructokinase, Liver Type	0,299049645	0,680856989
PFKM	Phosphofructokinase, Muscle	0,26951517	1,448236813
PFKP	Phosphofructokinase, Muscle	0,50501332	0,824149815
PGAM1	Phosphoglycerate Mutase 1	0,250807361	0,64997372
PGAM2	Phosphoglycerate Mutase 2	0,328717036	1,379190149
PGAM4	Phosphoglycerate Mutase Family Member 4	NA	NA
PGK1	Phosphoglycerate Kinase 1	0,513743119	1,254509508
PGK2	Phosphoglycerate Kinase 2	NA	NA
PGM1	Phosphoglucomutase 1	0,133241902	0,564135805
PGM2	Phosphoglucomutase 2	0,663724045	0,857083844
PKLR	Pyruvate Kinase L/R	NA	NA
PKM	Pyruvate Kinase M1/2	0,287013207	0,812611286
SLC2A2	Solute Carrier Family 2 Member 2	NA	NA
TPI1	Triosephosphate Isomerase 1	0,941820314	1,025247097

Gene expression with a significant fold-change (p-value < 0.05) are represented in bold

**Table 7: Modulation of apoptosis-related genes**

Gene Identifier	Name	p value	Fold-change
AIFM1	apoptosis-inducing factor, mitochondrion-associated, 1	0,199362328	0,639449303
AKT1	v-akt murine thymoma viral oncogene homolog 1	0,903326174	1,045649301
AKT2	v-akt murine thymoma viral oncogene homolog 2	0,112855784	1,821651378
AKT3	v-akt murine thymoma viral oncogene homolog 3	0,117804516	0,6797718
APAF1	apoptotic peptidase activating factor 1	0,769889443	0,902560231
ATM	ATM serine/threonine kinase	0,511187665	1,277774245
BAD	BCL2-associated agonist of cell death	0,286118299	0,689757837
BAX	BCL2-associated X protein	0,830420541	0,924623448
BCL2	B-cell CLL/lymphoma 2	0,839741507	0,941174
BCL2L1	BCL2-like 1	0,566707941	0,802924421
BID	BH3 interacting domain death agonist	0,384302331	0,829104049
BIRC2	baculoviral IAP repeat containing 2	0,437575137	0,744767992
BIRC3	baculoviral IAP repeat containing 3	0,400114784	0,768869328
CAPN1	calpain 1, (mu/I) large subunit	0,833694816	0,932174964
CAPN2	calpain 2, (m/II) large subunit	0,225296094	1,317461714
CASP10	caspase 10, apoptosis-related cysteine peptidase	0,971011931	0,992892787
CASP3	caspase 3, apoptosis-related cysteine peptidase	0,713346177	1,144247914
CASP6	caspase 6, apoptosis-related cysteine peptidase	0,661580935	1,162712082
CASP7	caspase 7, apoptosis-related cysteine peptidase	0,285262781	0,685107988
CASP8	caspase 8, apoptosis-related cysteine peptidase	0,668819826	1,149371757
CASP9	caspase 9, apoptosis-related cysteine peptidase	0,862854141	0,938388547
CFLAR	CASP8 and FADD-like apoptosis regulator	0,843239772	1,058259527
CHUK	conserved helix-loop-helix ubiquitous kinase	0,428063305	0,741342948
CSF2RB	colony stimulating factor 2 receptor, $\beta$ , low-affinity (granulocyte-macrophage)	0,776224604	1,109411913
CYCS	cytochrome c, somatic	0,719716149	0,88635877
DFFA	DNA fragmentation factor, 45kDa, $\alpha$ polypeptide	0,462816558	1,298604793
DFFB	DNA fragmentation factor, 40kDa, $\beta$ polypeptide (caspase-activated DNase)	0,701154449	1,131184516
ENDOG	endonuclease G	0,152357688	0,6478517
FADD	Fas (TNFRSF6)-associated via death domain	0,397813339	1,369562809
FAS	Fas cell surface death receptor	0,865235105	1,057710651
FASLG	Fas ligand (TNF superfamily, member 6)	0,973609065	0,994034229
IKBKB	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase $\beta$	0,095095502	0,530223183

IKBKG	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase $\gamma$	0,441894618	1,251377445
IL1A	interleukin 1, $\alpha$	0,711906163	1,127682815
IL1B	interleukin 1, $\beta$	0,495063417	0,808574454
IL1R1	interleukin 1 receptor, type I	0,970461484	0,992641039
IL1RAP	interleukin 1 receptor accessory protein	0,329273601	1,288384035
IL3	interleukin 3	0,973609065	0,994034229
IL3RA	interleukin 3 receptor, $\alpha$ (low affinity)	NA	NA
IRAK1	interleukin-1 receptor-associated kinase 1	0,837999596	0,924814635
IRAK2	interleukin-1 receptor-associated kinase 2	0,342780884	0,698998743
IRAK3	interleukin-1 receptor-associated kinase 3	0,159467763	1,631177653
IRAK4	interleukin-1 receptor-associated kinase 4	0,06213433	0,592806803
MAP3K14	mitogen-activated protein kinase kinase kinase 14	0,096472862	1,87260327
MYD88	myeloid differentiation primary response 88	0,681172738	1,16223696
NFKB1	nuclear factor of kappa light polypeptide gene enhancer in B-cells 1	0,343896474	1,436331212
NFKB2	nuclear factor of kappa light polypeptide gene enhancer in B-cells 2 (p49/p100)	0,501010256	0,791468983
<b>NFKBIA</b>	<b>nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, <math>\alpha</math></b>	<b>0,005436491</b>	<b>0,46509131</b>
NTRK1	neurotrophic tyrosine kinase, receptor, type 1	0,973609065	0,994034229
PIK3CA	phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit $\alpha$	0,67631771	0,90505911
PIK3CB	phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit $\beta$	0,385833499	0,719863814
<b>PIK3CD</b>	<b>phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit <math>\Delta</math></b>	<b>0,003686668</b>	<b>0,489329294</b>
PIK3CG	phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit $\gamma$	0,797625658	1,062819232
PIK3R1	phosphoinositide-3-kinase, regulatory subunit 1 ( $\alpha$ )	0,563889245	1,244823714
PIK3R2	phosphoinositide-3-kinase, regulatory subunit 2 ( $\beta$ )	0,929504446	1,034370431
<b>PIK3R3</b>	<b>phosphoinositide-3-kinase, regulatory subunit 3 (<math>\gamma</math>)</b>	<b>0,026179842</b>	<b>1,7</b>
PIK3R5	phosphoinositide-3-kinase, regulatory subunit 5	0,631900239	0,84227374
PPP3CA	protein phosphatase 3, catalytic subunit, $\alpha$ isozyme	0,419452095	1,362032333
PPP3CB	protein phosphatase 3, catalytic subunit, $\beta$ isozyme	0,824791962	0,9187095
PPP3CC	protein phosphatase 3, catalytic subunit, $\gamma$ isozyme	0,844255245	1,057924772
PPP3R1	protein phosphatase 3, regulatory subunit B, $\alpha$	0,596861336	0,895383811
PPP3R2	protein phosphatase 3, regulatory subunit B, $\beta$	0,598350247	1,194020511

PRKACA	protein kinase, cAMP-dependent, catalytic, $\alpha$	0,357147556	1,395403713
PRKACB	protein kinase, cAMP-dependent, catalytic, $\beta$	NA	NA
PRKACG	protein kinase, cAMP-dependent, catalytic, $\gamma$	NA	NA
PRKAR1A	protein kinase, cAMP-dependent, regulatory, type I, $\alpha$	0,844314868	1,064447347
PRKAR1B	protein kinase, cAMP-dependent, regulatory, type I, $\beta$	0,368720741	1,396584565
PRKAR2A	protein kinase, cAMP-dependent, regulatory, type II, $\alpha$	0,655175022	1,181994531
PRKAR2B	protein kinase, cAMP-dependent, regulatory, type II, $\beta$	0,19097457	0,621829009
RELA	v-rel avian reticuloendotheliosis viral oncogene homolog A	0,353293688	0,788962533
RIPK1	receptor (TNFRSF)-interacting serine-threonine kinase 1	0,09829423	0,533611587
TNF	tumor necrosis factor	0,937429777	0,977638217
TNFRSF10A	tumor necrosis factor receptor superfamily, member 10a	0,973609065	0,994034229
TNFRSF10B	tumor necrosis factor receptor superfamily, member 10b	0,7894674	0,961831425
TNFRSF10C	tumor necrosis factor receptor superfamily, member 10c	NA	NA
TNFRSF10D	tumor necrosis factor receptor superfamily, member 10d	0,375784319	1,394687092
TNFRSF1A	tumor necrosis factor receptor superfamily, member 1A	0,534048139	0,797867305
TNFSF10	tumor necrosis factor (ligand) superfamily, member 10	0,701968935	1,149350836
TP53	tumor protein p53	0,596464012	1,188908967
TRADD	TNFRSF1A-associated via death domain	0,778666734	1,103323869
TRAF2	TNF receptor-associated factor 2	0,878319658	1,056791317

Gene expression with a significant fold-change (p-value < 0.05) are represented in bold



# Material and Methods

## Strains

Strains	Genotype	Ref.
<b><i>Acanthamoeba castellanii</i></b>		
	Environmental isolate	
<b><i>Cell line</i></b>		
HeLa	Epithelial cells	INSERM U1111
THP-1	Monocytes	INSERM U1111
U937	Monocytes	INSERM U1111
HEK293T	Embryonic kidney cells	INSERM U1111
A549	Pneumocytes	INSERM U1111
<b><i>Saccharomyces cerevisiae</i></b>		
BY4742	<i>MAT<math>\alpha</math>; his3D1; leu2D0; lys2D0; ura3D0</i>	INSERM U1111
<b><i>Escherichia coli</i></b>		
DH5 $\alpha$	<i>fhuA2<math>\Delta</math>(argF-lacZ)U169 phoA glnV44 <math>\phi</math>80 <math>\Delta</math>(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17</i>	New England Biolabs (C2987H)
XL1-Blue	<i>endA1 gyrA96(nalR) thi-1 recA1 relA1 lac glnV44 F'[:Tn10 proAB+ lacIq <math>\Delta</math>(lacZ)M15] hsdR17 (rK - mK+)</i>	Stratagene
BL21 (DE3) (pREP4-groESL)	F- <i>ompT gal dcm lon hsdSB(rB- mb-) <math>\lambda</math>(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])</i>	[521]
<b><i>Legionella pneumophila</i></b>		
Paris	Clinical isolate	INSERM U1111
Paris $\Delta$ <i>dotA</i>	Paris strain without <i>dotA</i>	INSERM U1111
Paris $\Delta$ <i>legK4</i>	Paris strain without <i>lpp0267</i>	This study
Paris $\Delta$ <i>legK4</i> + <i>legK4</i>	Paris $\Delta$ <i>legK4</i> complemented with <i>legK4</i>	This study
Paris $\Delta$ <i>legK4</i> + <i>legK4</i> catalytic mutant	Paris $\Delta$ <i>legK4</i> complemented with catalytic mutant of LegK4 ( <i>legK4</i> <sub>D195N</sub> )	This study
Paris $\Delta$ <i>mavR</i>	Paris $\Delta$ <i>lpp0268</i>	INSERM U1111
Paris SYFP	Paris with <i>syfg</i> insertion in one tRNA gene	INSERM U1111
Paris dTomato	Paris with <i>dTomato</i> insertion in one tRNA gene	INSERM U1111
Paris $\Delta$ <i>legK4</i> $\Delta$ <i>mavR</i>	Paris without <i>legK4</i> and <i>mavR</i>	INSERM U1111
Paris $\Delta$ <i>legK4</i> $\Delta$ <i>mavR</i> + <i>legK4 mavR</i>	Paris $\Delta$ <i>legK4</i> $\Delta$ <i>mavR</i> complemented with <i>lpp0267</i> and <i>lpp0268</i>	This study
CIP 108286 Lens	Clinical isolate	[63]
Philadelphia-1	Clinical isolate	CIRI

Table 8: List of strains

## Plasmids

Plasmid	Description	Ref.
Donor vector		
pDONR <sup>TM</sup> 207	Gateway donor vector	Invitrogen
pDONR <sup>TM</sup> 207- <i>legK4</i>	Gateway donor vector with <i>legK4</i> insertion	This study
pDONR <sup>TM</sup> 207- <i>legK4</i> <sub>D195N</sub>	Gateway donor vector with <i>legK4</i> catalytic mutant gene insertion	This study
pDONR <sup>TM</sup> 207- <i>legK4</i> Cap	Gateway donor vector with insertion of the first 162 bp of <i>legK4</i> (Cap domain)	This study
pDONR <sup>TM</sup> 207- <i>legK4</i> lpp0267 ΔCap	Gateway donor vector with deletion of the first 162 bp of <i>legK4</i>	This study
pDONR <sup>TM</sup> 207- <i>legK4</i> <sub>NLS</sub>	Gateway donor vector with LegK4 NLS	This study
pENTR <sup>TM</sup> 1A	Donor vector	INSERM U1111
pENTR <sup>TM</sup> 1A- <i>legK4</i>	Gateway donor vector with LegK4	This study
pENTR <sup>TM</sup> 1A- <i>legK4</i> <sub>D195N</sub>	Gateway donor vector with <i>LegK4</i> <sub>D195N</sub>	This study
Expression vector		
peGFPc	Mammal cells expression vector for protein fusion with the eGFP (at C-terminal)	INSERM U1111
peGFP- <i>legK4</i>	peGFP with <i>legK4</i> insertion producing eGFP-LegK4	This study
peGFP- <i>legK4</i> <sub>D195N</sub>	peGFP with <i>legK4</i> insertion producing eGFP-LegK4 <sub>D195N</sub>	This study
peGFP- <i>legK4</i> <sub>1-445</sub>	Expression of eGFP-LegK4 <sub>1-445</sub> (from amino acid 1 to 445)	This study
peGFP- <i>legK4</i> <sub>1-54</sub>	Expression of eGFP-LegK4 <sub>1-54</sub>	This study
peGFP- <i>legK4</i> <sub>54-445</sub>	Expression of eGFP-LegK4 <sub>54-445</sub>	This study
peGFP- <i>NLS</i>	Expression of eGFP- KVLGNKGIRKLIK-SANGK	This study
peGFP- <i>legK4</i> Δ <i>NLS</i>	Expression of eGFP-LegK4 without NLS sequence	This study
peGFP- <i>NLSmut1</i>	Expression of eGFP-AVLGNKGIRKLIK-SANGK	This study
peGFP- <i>NLSmut2</i>	Expression of eGFP-KVLGNAGIRKLIK-SANGK	This study
peGFP- <i>NLSmut3</i>	Expression of eGFP-KVLGNKGIAKLIK-SANGK	This study



Plasmid	Description	Ref.
peGFP- <i>NLSmut4</i>	Expression of eGFP- KVLGNK-GIRALKIKSANGK	This study
peGFP- <i>NLSmut5</i>	Expression of eGFP- KVLGNK-GIRKLAIKSANGK	This study
peGFP- <i>NLSmut6</i>	Expression of eGFP- KVLGNGKIRK-LKIASANGK	This study
peGFP- <i>legK4<sub>NLSmut7</sub></i>	Expression of eGFP- KVLGNGKIRK-LKIKSANGA	This study
peGFP- <i>legK4<sub>NLSmut8</sub></i>	Expression of eGFP- <b>AVLGNAGIAALAIRASANGA</b>	This study
peGFP- <i>legK4</i> Philadelphia	Expression of eGFP-LegK4 Philadelphia	This study
peGFP- <i>legK4</i> $\Delta$ 1-58 Philadelphia	Expression of eGFP-LegK4 Philadelphia without the first 58 amino acids	This study
peGFP- <i>legK4</i> + 1-58 Philadelphia	Expression of eGFP-LegK4 Paris with the first 58 amino acids of LegK4 Philadelphia	This study
peGFP- <i>legK4</i> Lens	Expression of eGFP-LegK4 Lens	This study
pDEST27	Mammal cells expression vector for protein fusion with the GST (at C-terminal)	Invitrogen
pDEST27- <i>legK4</i>	Expression of GST-LegK4	INSERM U1111
pCl-Neo3Flag	Mammal cells expression vector for protein fusion with 3Flag Tag (at C-terminal)	INSERM U1111
pCl-Neo3Flag- <i>legK4</i>	Expression of 3Flag-LegK4	This study
pINDUCER21	Gateway expression vector for HA-tagged protein	INSERM U1111
pINDUCER21- <i>legK4</i>	Expression of HA-LegK4 under doxycycline induction	This study
pINDUCER21- <i>legK4<sub>D195N</sub></i>	Expression of HA- <i>LegK4<sub>D195N</sub></i> under doxycycline induction	This study

## Protein overexpression

*E. coli*

pQE30	Overexpression of 6His fused protein by IPTG induction	Invitrogen
pQE30- <i>legK4</i>	Overexpression of 6His-LegK4 of	This study
peGFPc- <i>legK4<sub>D195N</sub></i>	Overexpression of 6His-LegK4 <sub>D195N</sub>	This study
pGEX-6P-3	Overexpression of GST	GE Healthcare
pGEX-6P-3- <i>mavR</i>	Overexpression of GST-MavR	This study
pGEX-6P-1-H3.1	Overexpression of GST-H3.1	GE Healthcare
pGEX-6P-1-H3.1 <sub>T32A</sub>	Overexpression of GST-H3.1 <sub>T32A</sub>	This study

Plasmid	Description	Ref.
<b>Yeast</b>		
pBT3N	Two-hybrid screen yeast vector	INSERM U1111
pBT3N- <i>yGFPc</i>	Modified vector expressing yGFPc under galactose induction	This study
pBT3N- <i>yGFPc-legK4</i>	Modified vector expressing yGFP-LegK4 under galactose induction	This study
pMet-YC	Yeast expression vector	INSERM U1111
pMet-YC- <i>yGFPc</i>	Constitutive expression of yGFPc	This study
pMet-YC- <i>yGFPc-legK4</i>	Constitutive expression of yGFP-LegK4	This study
<b><i>L. pneumophila</i></b>		
pMMB207-4HA	Modified pXDC61 with 4HA tag	[196]
pMMB207-4HA- <i>legK4</i>	Expression of 4HA-LegK4 under IPTG induction	This study
pMMB207-4HA- <i>legK4<sub>D195N</sub></i>	Expression of 4HA-LegK4 <sub>D195N</sub> under IPTG induction	This study
pMMB207-4HA- <i>legK4 mavR</i>	Expression of 4HA-LegK4 and MavR under IPTG induction	This study
pXDC50	Expression of a protein of interest and mCherry fluorescent protein under IPTG induction	INSERM U1111
pXDC50- <i>legK4</i>	Expression of LegK4 and mCherry under IPTG induction	This study
pXDC50- <i>legK4</i>	Expression of LegK4 and mCherry under <i>legK4 promoter</i>	This study
pXDC50- <i>legK4 mavR</i>	Expression of LegK4, MavR and mCherry under <i>legK4 promoter</i>	This study

**Table 9: List of plasmids**

## Antibodies

Primary antibody	Ref.	Dilution for	
		Western-blot	Immunofluorescence
Mouse anti-C23 (D6)	SC-17827 (Santa Cruz)	-	1/100
Rabbit anti-BiP	Ab32618 (Abcam)	1/200	-
Rabbit anti-RBBP4	Ab79416 (Abcam)	1/50 000	-
Rabbit anti-PTEN (138G6)	9559 (Cell Signaling)	1/1 000	-
Mouse Anti-GST	13-6700 (Invitrogen)	1/3000	-
Mouse Anti-Flag	F1804 Sigma	1/1 000	1/100
Mouse Anti-phosphothreonine	P6623 (Sigma)	1/500	-
Rabbit Anti-phosphothreonine	9381S (Cell Signalling)	1/5 000	-
Rabbit Anti-HA	3724S (Cell Signalling)	1/1000	1/1 600
Rabbit Anti-Histone H3	4499S (Cell Signalling)	1/2 000	1/400
Rabbit Anti-GFP	INSERM U1111	1/5 000	-
Rabbit Anti-Histone H3 phospho T32 (H3T32)	ab4076 (Abcam)	1/1 000	-
Rabbit Anti-BiP	3177S (Cell signaling)	1/1 000	1/400
Rabbit Anti-CHOP	2895S (Cell signaling)	1/1 000	-
Rabbit Anti- $\beta$ actin	A2228 (Sigma)	1/5 000	

Secondary antibody	Ref.	Dilution for	
		Western-blot	Immunofluorescence
Peroxydase Goat anti-Mouse IgG	A0545 (Sigma)	1/5 000	-
Peroxydase Goat anti-Rabbit	A0163 (Sigma)	1/5 000	-
Goat anti-Rabbit IgG Dylight 594	115-115-003 (Jackson Immunoresearch)	-	1/500
Goat anti-IgG de Rabbit Dylight 594	115-515-003 (Jackson Immunoresearch)	-	1/250

**Table 10: List of antibodies**

## Primers

Name	Sequence
<i>legK4</i> -BamHI	ataggatccaattgcttcggtttcatgaatt
<i>legK4</i> -SalI	cgcgctgcagcttctggctaaatagcttgcg
<i>legK4</i> -SalI	cgcgctgcacctgccacatcaagtcacctc
<i>legK4</i>	ggttcatgattgcaacctctaattgctttctcgtccatgcc
5-KpnI- <i>legK4</i>	cggggctaccatgaaattgcttcggtttcatg
5-EcoRI- <i>legK4</i>	atagaattctatgaaattgcttcggtttcatg
3-SalI- <i>legK4</i>	atagtcgacttaatatggcaaatgatgacg
5- <i>legK4</i> -nt650	cagttctcggataaaagaacc
5- <i>legK4</i> -nt1300	cagctaaaatgatggcagcg
3-XhoI <i>legK4</i>	ccatactcgagtaatatggcaaatgatgacgttgc
$\alpha 3$ f	gttttctaagcctttgccaactttttgtacaaactgtttgagctcg
$\alpha 3$ r	caaaaaagtggcaaggcatagaaaactaataaaactgccaacggc
$\Delta \alpha 1$ f	aattgcttcggaatcattgccaggcatggacgagaaagc
$\Delta \alpha 1$ r	tggcaatgatttccgaagcaatttcatgccaaactttttgtac
$\Delta \alpha 2$ f	ttgccaggcatggacaacaaaggcattagaaaactaataaaactgcc
$\Delta \alpha 2$ r	tctaattgcctttgttgcctatcctggcaatgattttaattcatgaaaccg
$\Delta \alpha 3$ f	gttttaggcaactctgccaacggcaagcctatctcacgcgaaataatg
$\Delta \alpha 3$ r	tgccgttggcagagttgcctaaaaccttatcaacaattctaattgc
$\Delta \alpha 23$ f	ttgccaggcatggactctgccaacggcaagcctatctcacgcgaaataatg
$\Delta \alpha 23$ r	tgccgttggcagagtcctatcctggcaatgattttaattcatgaaaccg
$\alpha 2 3$ f	ttctaattgcctttctgccaactttttgtacaaactgtttgagctcg
$\alpha 2 3$ r	caaaaaagtggcgagaaagcattagaattgttgataaagggttttaggc
NLS f	caaaaaagtggcaagggttttaggcaacaaaccgtaattcttttc
NLS r	gcctaaaacctgccaactttttgtacaaactgtttgagctcg
$\Delta$ NLS f	gaattgttgatacctatctcacgcgaaataatgattcatgagtttgaacc
$\Delta$ NLS r	ttcgctgagataggatcaacaattctaattgctttctcgtccatgcc
R33G f	gttggcagattttattagttttccaatgcctttgttggc
R33G r	ggcaacaaaggcattggaaaactaataaaactgccacc
R33A f	gccgttggcagattttattagttttgcaatgcctttgttggc
R33A r	ggcaacaaaggcattgcaaaaactaataaaactgccaacggc
K25A f	gcctttgttgcctaaaaccgcttgcctaaattctaattgc
K25A r	gcattagaattgggcaacgcggttttaggcaacaaaggc
K30A f	tattagttttctaattgcctgcttgctaaaacctttatc
K30A r	gataaagggttttaggcaacgcaggcattagaaaactaata
K34A f	gttggcagattttattagtgctctaattgcctttgttgc
K34A r	gcaacaaaggcattagagcactaataaaactgcccaac
K37A f	gcttgccttggcagatgctattagttttctaattgcctttgttgc
K37A r	caacaaaggcattagaaaactaattagcatctgcaacggcaagc
K42A f	ctaataaaactgcccaacggcgccttatctcacgcgaaataatg
K42A r	cattatttgcgctgagataagcgcgccgttggctgattttattag
R33A K34A f	ggcagattttattagtgctccaatgcctttgttgcctaaaacc
R33A K34A r	ggtttttaggcaacaaaggcattggagcactaataaaactgcc
K25A K30A R33A K34A f	attagaattgttgatagcggtttaggcaacgcaggcattgcagcac
K25A K30A R33A K34A r	tgctgcaatgcctgcgttgcctaaaaccgctctcaacaattctaattg
NLS to Alanine f	agcactaatagcatctgccaacggcgcgcttatctcacgcg
NLS to Alanine r	ttcgctgagataggcgcgccgttggcagatgctattagtgctg
Delta 1-58 phila f	aaaaaagtggcatgaaattgcttcggtttcatgaattaaatc
Delta 1-58 phil ra	gaagcaatttcatgccaactttttgtacaaactgtttgagctgc
1-58 phila into Paris f	aaacaagttgtacaaaaagtggcttgtgcaatcattctctaattgagcccc
1-58 phila into Paris r	atgaaaccgaagcaatttcataaggaacctcacatatttac
P1-lpp0267	tgaccattcctgtcactctac
P2-lpp0267	ggcccaattcgcctatagtgagctgtaattgctttctcgtccatgcc
P3-lpp0267	gggtttgctcgggtcggtggcatatggatggttgc aaatcatgaacc
P4-lpp0267	tttactcacctgggtttctc
P5-lpp0267	ggttcatgattgcaacctctaattgctttctcgtccatgcc
P6-lpp0267	ggcatggacgagaaagcattagatggttgc aaatcatgaacc

Table 11: List of primers

## Culture condition and medium

### Bacteria

#### *Legionella pneumophila*

*L. pneumophila* Paris strain was provided by the Legionella Reference National Center (CNR), Lyon, France. *L. pneumophila* Philadelphia-1 strain was kindly provided by X. Charpentier, International Center for Infectology Research (CIRI), Lyon, France. Philadelphia-1 strain used in this study had five single nucleotide polymorphisms compared to sequenced genome (NC\_002942.5): G688274A, T816562G, GT820457/58CG, del1355972T.

*Legionella pneumophila* strains used for this study were stored in cryovial at  $-80^{\circ}\text{C}$  freezer (TS/80-MX). Frozen bacteria were then plated 3 days at  $37^{\circ}\text{C}$  on CYE (charcoal yeast extract agar) plates containing : ACES 10g/L (acid N-(2-acetamido)-2-aminoethanesulfonic) (Roth, 9138.2), yeast extract 10 g/L (Merck, 1.03753.0500), bacterial agar 15 g/L (Roth, 5210.4) and charcoal 2 g/L (Sigma, C9157) diluted in commercial water (Roth, 3478.4), pH was adjusted to 6.9 by potassium hydroxide addition (Roth, 0995.2) then the mix was sterilized by autoclave 20 minutes at  $120^{\circ}\text{C}$ .

Medium was supplemented with 0.4 g/L of L-cysteine (Roth, 1693.2) and 0.25 g/L of iron nitrate (Sigma, 216828). Finally, bacteria from the freezer were plated 3 days at  $37^{\circ}\text{C}$ . In liquid medium, bacteria were grown in AYE medium (ACES-buffered yeast extract) containing : 0.3 g/L of iron pyrophosphate (Sigma, P6526), 0.1 g of cysteine (Roth, 1693), 10 g of ACES (Roth, 9138.2) and 12 g of yeast extract (BD, 212750) diluted in commercial water, pH was adjusted to 6.9 by potassium hydroxide addition and sterilized by filtration ( $0.2\ \mu\text{m}$ ). Bacteria are cultivated at  $37^{\circ}\text{C}$  under stirring (200 rpm) in aerobic condition during the appropriated time. If necessary, 2 mM of isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) (Roth, 2316.1) or/and 5  $\mu\text{g}/\text{mL}$  and/or chloramphenicol (Roth, 3886.1) were added.

#### *E. coli*

*E. coli* strains used in this study are stored in  $-80^{\circ}\text{C}$  freezer in a mix of LB medium (Luria-Bertani) and 15 % glycerol. Bacteria from freezer are cultivated on LB plate containing 20 g/L of bacterial agar and then incubated overnight at  $37^{\circ}\text{C}$ .

Liquid culture were performed in aerobic condition at  $37^{\circ}\text{C}$  in LB medium, overnight at 200 rpm (rotation per minute)

### Eukaryotic cells

In this study, HeLa human epithelial cell line , U937 macrophages derived cell from human monocytes, *A. castellanii* amoebae and *S. cerevisiae* have been used (Table 8).

## Human HeLa epithelial cells, HEK293T, U937 and THP-1 monocytes

HeLa epithelial cells, HEK293T, U937 and THP-1 monocytes was nitrogen-frozen in FBS (Fetal Bovine Serum) (Gibco, 10270-106) containing 10 % DMSO (dimethyl sulfoxide) (AppliChem, A3672). HeLa epithelial and HEK293T cell lines maintenance is done in DMEM (Dulbecco's modified Eagle medium) (Gibco, 41966) supplemented with 10 % FBS and as well 100 units/mL penicillin and 100  $\mu\text{g}/\text{mL}$  streptomycin. The cells are incubated at 37°C in a 5 %  $\text{CO}_2$  containing atmosphere and used from passage 5 to 15. A passage is achieved when removing supernatant and cells are then washed with PBS (phosphate buffered saline) (Gibco, 14190). Cells were then detached using 0.05 % trypsin-EDTA (Gibco, 253004) at 37°C for 5 minutes. The cells are then placed in culture medium whose FBS which inhibits trypsin activity and centrifuged for 3 minutes at 300g. The cell pellet is diluted at the appropriate concentration in brand-new cell culture medium.

Maintenance of the U937 and THP-1 monocytes cell line was performed in RPMI 1640 (Roswell park memorial institute medium) (Gibco, 61870) supplemented with 10 %FBS, 100 units/mL of penicillin and 100 $\mu\text{g}/\text{mL}$  streptomycin. The cells are incubated at 37°C in a 5 %  $\text{CO}_2$  containing atmosphere and used from passage 5 to 15. The passage of these non-adherent cells is done by counting cells in a Malassez cell after a 1/2 dilution in trypan blue and with cell culture medium. Trypan blue penetrates only in dead cells with permeabilized membrane, appearing blue. Since the U937 are non-adherent, dead cells are not eliminated during cell passage, so it is essential to exclude them. To plate and differentiate these monocytes into macrophages, the cells are centrifuged for 3 minutes at 300g and then plated in RPMI supplemented with 10 %FBS and 50 ng/mL PMA (phorbol 12-myristate 13-acetate) (Sigma P8139) for 3 days at 37°C in an atmosphere containing 5 %  $\text{CO}_2$ .

## *Acanthamoeba castellanii* amoebae

*A. castellanii* are frozen at -80°C in PYG (peptone yeast glucose) containing: 20 g/L peptone proteose (Difco 211684), 1 g/L yeast extract, 16 mM  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$  (Roth, P027.2), 0.4 mM  $\text{CaCl}_2$  (Acros, 423525000), 3.4 mM sodium citrate (Roth, 3580.1), 0.05 mM  $(\text{NH}_4)_2\text{Fe}(\text{SO}_4)_2 \cdot 6\text{H}_2\text{O}$  (Sigma, F3754), 2.5 mM  $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$  (Euromedex, 1309), 2.5mM  $\text{KH}_2\text{PO}_4$  (Roth, 3904) and 0.1 M glucose (Euromedex UG3050) in water, adjusted 6.5 pH with potassium hydroxide and supplemented with 10 %DMSO, 100 units/mL penicillin and 100  $\mu\text{g}/\text{mL}$  streptomycin (Gibco, 15140122). Amoeba are maintained in PYG supplemented with 100 units/mL penicillin and 100  $\mu\text{g}/\text{mL}$  streptomycin at 30°C. Amoebae used for infections are cultured in the same medium for 4 days to induce nutrient deficiency. These cells are then detached by incubation for 10 minutes at 4°C and then centrifuged for 3 minutes at 300g. After being count with a Malassez cell, amoebae are plated at the necessary concentration in PY (PYG medium without glucose).

## *Saccharomyces cerevisiae*

BY4742 strain was cultured in YPD (Yeast extract, peptone, Dextrose) rich medium containing: 1% yeast extract, 2 % peptone, 2 % glucose and 0.004 % adenine or in Synthetic Dextrose (SD) medium containing 1.7 g/L of yeast nitrogen base, 5 g/L  $(\text{NH}_4)_2\text{SO}_4$ . SD medium require to add specific amino acids depending on the experiment. These two liquid mediums can be use as plate by addition of agar. Before use, YPD and SD were sterilized 20 minutes at 120°C and amino acid complement are filtrated.

## Molecular biology

### Genomic DNA Extraction of *L. pneumophila*

After 3 days at 37°C on CYE plate, approximately 10  $\mu\text{L}$  of bacteria are mixed in 1.5 mL NTE buffer made of: 100 mM NaCl (Roth, 3957.1), 20 mM Tris (trishydroxymethylaminomethane) (Roth, 5429.3) and 1 mM EDTA (Ethylenediaminetetraacetic acid) (Euromedex, EU0007-B) diluted in commercial water, pH was adjusted to 7.4 with hydrochloric acid solution (Roth, 9277.1). After a centrifugation step at 10 000g for 10 minutes, pellet containing bacteria was resuspended in 780  $\mu\text{L}$  of NTE buffer then 20 $\mu\text{L}$  of 10 % SDS (sodium dodecyl sulfate) (Euromedex 1012.D) and 2  $\mu\text{L}$  of proteinase K (20mg/mL) (Thermo EO0491) to lyse cells. The mix was incubated 1 hour at 37°C under stirring. To eliminate RNA, the mix was incubated 1 more hour with 1  $\mu\text{L}$  (100mg/mL) of ribonuclease (Qiagen, 19101). To eliminate proteins, 800  $\mu\text{L}$  of a phenol/chloroform/isoamyl alcohol mix (25/24/1) (Roth, A156.2) was added to the solution. After several minutes of stirring, the tubes were centrifuged 5 minutes at 10 000g until white pellet containing proteins is totally separated from supernatant. Supernatant which contained nucleic acids was kept and mixed with 700  $\mu\text{L}$  of chloroform (Roth 6340.1)/isoamyl alcohol (Prolabo 20798.295) (24/1) to eliminate remaining phenol and proteins. After stirring, the mix was centrifuged 5 minutes at 10 000g until pellet is well separated from supernatant. Aqueous phase containing DNA was mixed with 1.4 mL of ice-cold absolute ethanol (AnalaR NORMAPUR, 20821.310) and 72  $\mu\text{L}$  of a solution containing 5 M potassium acetate (Roth, T874.2) and 11.5 % of ice-cold acetic acid (Roth, 3738.1). Chromosomal DNA precipitate was picked with a platinum loop, dried and resuspended in 100  $\mu\text{L}$  of ultra-pure water.

### Plasmid extraction

Plasmidic DNA used for mammalian cells transfection or bacterial transformation have been extracted depending on sample size thank to Qiagen purification kit: Qiaprep Spin Miniprep Kit (27106) or Qiagen Plasmid Midi Kit (Q12145). These kits are based on DNA fixation on silica matrix avoiding in sample phenol contamination.

### Plasmid construct

#### *L. pneumophila* genes mutants generation

- Homologous recombination cassette construction

To obtain clean deletion mutant in *Legionella* chromosome, the first step was to make a Homologous Recombination (HR) cassette called HR1 which will be integrated in bacteria genome and then exchange by a second HR cassette (HR2) (Figure 32). HR1 cassette was made by joint PCR and contained a *mazF* toxic cassette squeezed between upstream and downstream region of the gene of interest, allowing homologous recombination. MazF cassette is made of *mazF* gene under the control of *lacI* promoter and allow kanamycine resistance (Figure 32). Unlike HR1, HR2 is only made of upstream and down region of the gene of interest. Each DNA fragment was amplified by PrimeSTAR<sup>®</sup> DNA

Polymerase. Upstream region was amplified from genomic DNA of *Legionella* Paris strain with P1/P2 primers for HR1 cassette and P1/P5 for HR2 cassette. On the other hand, downstream region was amplified with P3/P4 for RH1 cassette and P6/P4 for HR2 cassette.

MazF cassette was amplified from pGEM-*mazF* with MazFk7-F/MazFk7-R primers provided by X. Charpentier. Size of each amplicons were verified by electrophoresis on 1% agarose gel. Then, gels were cut and purified using Quiagen<sup>®</sup> kit, according to manufacturer's instructions. Finally, two other PCRs were necessary to joint each fragment together and made HR cassettes. Once again, the size of the cassette was determined and checked on agarose gel and purified.

- Natural transformation and Homologous recombination

After 3 days at 37°C on CYE plate, bacteria were spotted on a brand new CYE plate for 24 hours in the same condition. Bacteria were resuspended in AYE medium at OD<sub>600nm</sub> 0.0124, 0.025 or 0.05 for 16 hours at 37°C under stirring. When one of the three conditions reach an OD<sub>600nm</sub> between 1.1 and 1.5, tubes were centrifuged at 4500 rpm during 5 minutes at room temperature. Half of the supernatant was discarded and bacterial pellet was resuspended. 500 ng of the HR cassette was added to the bacteria and incubated without stirring at 30°C for 16 to 24 hours.



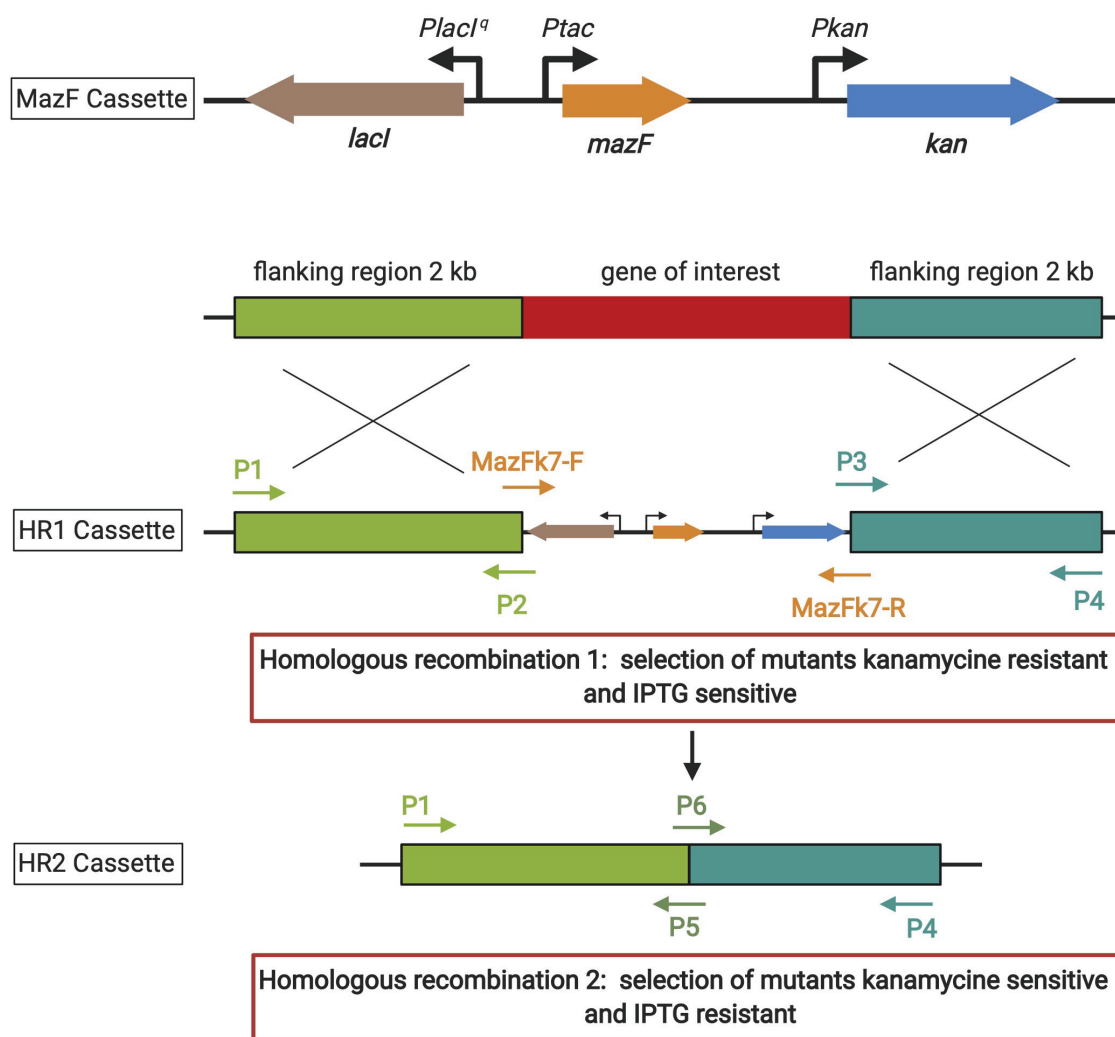


Figure 32: Creation of scar free deletion in *Legionella* genome

- Natural transformation and Homologous recombination

- Homologous recombination 1: Bacteria that have integrated HR1 cassette became kanamycin resistant and IPTG sensitive. Indeed, under IPTG induction *LacI* repressor activity is suppressed and *MazF* toxin is expressed and killed bacteria. These selections allow us to distinguish modified bacteria which contained HR1 cassette from wild type bacteria. Bacteria were plated on two CYE Kan15 (15  $\mu\text{g}/\text{mL}$  kanamycin) with respectively 100  $\mu\text{L}$  and the rest of the culture and were incubated 4 days at 37°C. Positive clones were patched on a new CYE Kan15 containing 0.5M of IPTG. Presence or absence of the HR cassette was verified by colony PCR using DreamTaq<sup>®</sup> DNA Polymerase.

- Homologous recombination 2: Clone from HR1 recombination were then transformed in the same condition but using HR2 cassette. Bacteria were plated on CYE IPTG for 4 days at 37°C and positive mutants were then plated

on a new CYE IPTG and on CYE Kan15. Mutants possessing HR2 are kanamycin sensitive because HR2 do not give kanamycin resistance and IPTG resistance. Absence (scar free mutation) was verified by colony PCR using DreamTaq<sup>®</sup> DNA Polymerase.

## Checking

To check if every strain created during this project do not have growth issue, bacteria were plated 3 days on CYE plate at 37°C, then 24 hours on a new CYE plate. Finally, bacteria were incubated at OD 0.1 in AYE medium in 96 well plate (Greiner bioone, 655090) and OD was measured using a plate reader (TECAN, Infinite 200pro). The plate was incubated at 37°C and OD was measured every 30 minutes right after 30 seconds stirring, for 30 hours.

# Transformation of bacteria and yeast cells

## *E.coli* transformation

### Chemically competent cell

*E.coli* competent cells have been prepared by calcium chloride ( $\text{CaCl}_2$ ) treatment. Among all the competent cell preparation protocols, it is the most efficient technique to increase bacterial cell's ability to incorporate plasmid DNA. Addition of  $\text{CaCl}_2$  to the cell suspension allows the binding of plasmid DNA to LPS. Then, negatively charged DNA backbone and LPS come together, and when heat shock is provided, plasmid DNA passes into the bacterial cell.

After an overnight culture, one loop of bacteria (DH5 $\alpha$  or BL21) has been diluted in new LB medium then incubated at 37°C until  $OD_{600nm} = 0.6$ . Then, culture were centrifuged at 7000 rpm during 10 minute and the bacterial pellet was resuspended in half the volume of the culture with ice-cold  $\text{CaCl}_2$  (50 mM). Samples were incubated on ice for 40 minutes and then centrifuged. Bacterial pellet was resuspended with ice-cold solution of 50 mM  $\text{CaCl}_2$  and 20 % glycerol. 150  $\mu\text{L}$  samples were alicoted and immediately store in -80°C freezer.

## Transformation

One tube from the last step was thawed on ice and mix with 15  $\mu\text{L}$  of ligation mix or 2  $\mu\text{L}$  of plasmide and incubated at least 20 minutes on ice. Then, samples was incubated in a 37°C water-bath to do the thermal shock for 3 minutes and put on ice for 5 more minutes. 1 mL of warm LB was added to all samples and were incubated at least 1 hour at 37°C. Finally, bacteria were plated on the appropriated medium allow discrimination between bacteria with or without gene of interest.

50  $\mu$  of ultracompetent cell XL1-Blue were thawed on ice and add to 1  $\mu\text{L}$  of Gateway mix on ice for 30 minutes. Then, tubes were put in a water-bath at 42°C for 30 secondes, put back on ice for 5 minutes and then 950  $\mu\text{L}$  of LB was

added. Tubes were incubated at 37°C for 1 hour and then plated on the appropriated medium.

## ***L.pneumophila* transformation**

### **Electrocompetent cells**

Three days prior the experiment, *Legionella* strains were plated on CYE plate and incubated at 37°C. From these plates, a bacterial suspension was made in 200 mL of ice-cold ultrapure water and centrifuged 20 min at 4500 rpm at 4°C. This washing step was done three times and finally the bacterial pellet was resuspended in 1 mL of ice-cold 30 % glycerol. Suspension was split in 100  $\mu$ L samples and were then store in -80°C freezer until electroporation.

### **Electroporation**

A tube of electrocompetent cells were slowly thawed on ice, then 2  $\mu$ L of plamid was added. The mix was put in an ice-cold electroporation chamber and subject to an electric field about 2.3 kV, 200  $\Omega$  and 25  $\mu$ F. Then, 1 mL of warm AYE was added and all this was incubated at least 1 hour at 37°C. Finally, bacterial suspension was plated on appropriated medium and incubated at 37°C for 5 days.

## ***S.cerevisiae* transformation**

### **Competent cell**

From an overnight culture of *S.cerevisiae*, yeast were seeded at  $OD_{600nm} = 0.2$  in 50 mL of YPD medium and incubated at 30°C until  $OD_{600nm} = 0.7$ . Cells were centrifuged three minutes at 1300g, washed with 10 mL of sterile water and centrifuged again. Yeast pellet was resuspended in 5 mL of LiAc/TE (100 mM LiAc, 1 mM EDTA, 10 mM Tris HCL pH 7.5) and then centrifuged 3 minutes at 1300g. Finally, the pellet was resuspended with 1 mL of 1 mM LiAc.

### **Transformation**

80  $\mu$ L of competent cells were mixed to 5  $\mu$ L of denatured fragmented salmon sperm DNA, 4  $\mu$ L of vector (1 to 2  $\mu$ g), 6  $\mu$ L of DMSO and 300  $\mu$ L of LiAc/TE/PEG (LiAc/Te, 40% PEG 4000). After mixing by slowly turning the tube over, tubes were incubated at 30°C for 30 minutes and at 42°C for 15 minutes. Finally, yeast were plated on SD medium supplemented with the appropriated amino acids and incubated at 30°C for 3 days.

## Mammalian cells transfection

### Plasmidic DNA

The day before transfection, HeLa cells were seeded at 150 000 cells in 6-well plate in 2 mL of DMEM (Gibco, 41966) supplemented with 10 % SVF and 100 units/mL penicillin and 100  $\mu\text{g}/\text{mL}$  streptomycin. The cells were then incubated for 24 hours at 37°C in a 5 % CO<sub>2</sub> atmosphere. The day of transfection, 2  $\mu\text{g}$  of plasmid and 200  $\mu\text{L}$  of jetPRIME<sup>®</sup> buffer (Polyplus, 114-07) were mixed, vortex for 10 seconds and spin down. Then 4  $\mu\text{L}$  jetPRIME<sup>®</sup> reagent (Polyplus, 114-07) was added to the mix, vortex 1 second, spin down and incubated for 10 minutes at room temperature. Finally all the mix was added to the cells, incubated for 4 hours at 37°C in a 5 % CO<sub>2</sub> atmosphere, then medium was changed with 2 mL of DMEM and put back for 20 more hours.

### siRNA

The day before transfection, HeLa cells were seeded at 150 000 cells in 6-well plate in 2mL of DMEM (Gibco, 41966) supplemented with 10 % SVF and 100 units/mL penicillin and 100  $\mu\text{g}/\text{mL}$  streptomycin. The cells were then incubated for 24 hours at 37°C in a 5 % CO<sub>2</sub> atmosphere. The day of transfection, 110 pmoles (1524 ng) of siRNA (BiP or control (Thermo, AM16708)) and 200  $\mu\text{L}$  of jetPRIME<sup>®</sup> buffer (Polyplus, 114-07) were mixed, vortex for 10 seconds and spin down. Then 4  $\mu\text{L}$  jetPRIME<sup>®</sup> reagent (Polyplus, 114-07) was added to the mix, vortex 1 second, spin down and incubated for 10 minutes at room temperature. Finally all the mix was added to the cells and incubated for 48 hours at 37°C in a 5 % CO<sub>2</sub> atmosphere. 48 hours after transfection, a portion of the cells is analysed by Western-Blot to check the absence/presence of the target protein.

## LegK4 Stable and inducible cell line

### Construction of expression vectors

Amplification of *legK4* and *legK4<sub>D195N</sub>* genes is performed by PCR using the 5-KpnI and 3-XhoI-*legK4* primers. The two amplicons are digested by XhoI and KpnI (Biolabs), purified and then ligated to obtain two Gateway<sup>®</sup> input vectors: pENTR<sup>™</sup>1A-*legK4* and pENTR<sup>™</sup>1A-*legK4<sub>D195N</sub>*. Both plasmids were amplified in the *CaCl<sub>2</sub>* chemiocompetent *E. coli* DH5 $\alpha$ . Presence/absence of the insert was verified by PCR, restriction digestion and sequencing. The transfer of the insert from the input vector to the destination vector, pINDUCER21 (donated by T. Henri, CIRI, Lyon) was done by recombination reaction 'LR' with Gateway<sup>®</sup> LR Clonase<sup>®</sup> Enzyme Mix or Gateway<sup>®</sup> LR Clonase<sup>™</sup> II Enzyme Mix according to the supplier's recommendations (Invitrogen). This reaction was then used to transform, by thermal shock, ultra-competent *E. coli* DH5 $\alpha$  cells (BioLabs C2987H). The expression vectors, pINDUCER21-*legK4* and pINDUCER21-*legK4<sub>D195N</sub>*, obtained are purified with the QIAfilter Plasmid Midi kit (Qiagen) and then verified by PCR and restriction

digestion. The expression vectors pINDUCER21-*legK4* and pINDUCER21-*legK4*<sub>D195N</sub> make it possible to obtain an HA tagged protein whose sequence contains the viral integration sequence  $\psi$ .

## Lentivirus production

HEK 293T cells were transfected with the three following plasmids: p8.9 (gag-pol expressor), pMDG (VSV-G expressor) and pINDUCER21-*legK4* or pINDUCER21-*legK4*<sub>D195N</sub> or pINDUCER21 in OptiMEM medium and polyethylenimine (PEI) for 4h and then cells were washed with 1X PBS. Virions produced were collected after 48 hours and 72 hours post-transfection by filtering the supernatant with a 0.45  $\mu$  m low binding protein filter (SLHP033RS, Merk). In ultracentrifugation tubes containing the filtered culture supernatant, a 20 % sucrose solution was gently added to the bottom of each tube to create a sucrose gradient and collect only virions after centrifugation at 28000 rpm for 1 hour and half. Finally, virions were resuspended in PBS 1 % glycerol during 2 hours at room temperature.

## Transduction

50  $\mu$ L of pure or diluted lentivirus solution at the 10<sup>th</sup>, 100<sup>th</sup> and 1000<sup>th</sup> and 8  $\mu$ M Polybrene<sup>®</sup> (ThermoFisher) were added to 10 000 HeLa and THP-1 cells in 24-well plates. Plates were then centrifuged at 1500 rpm for 1 hour 40 to do spinoculation. GFP positive cells were sorted in 1X PBS and 1 % SVF medium with the BD FACSAria II sorter to collect only transduced cells.

## Optimal doxycycline concentration for LegK4 induction

Doxycycline was added at 0, 0.02, 0.04, 0.08 and 0.1  $\mu$ g/mL for 16 hours, or 24 hours at 0, 0.006, 0.012, 0.025, 0.05, 0.1, 0.2, 0.39, 0.78, 1.56, 3.12, 6.25, 12.5, 25, 50 or 100  $\mu$ g / mL to 100 000 HeLa cells to induce the expression of more or less *LegK4*<sub>D195N</sub>-HA. The presence of LegK4-HA was verified by Western-Blot using an anti-HA and anti-actin antibody. The optimal concentration corresponds to the concentration allowing to obtain living cells which LegK4 is detected.

## Biochemistry

### Overexpression and purification of GST-tagged protein

#### Bacterial culture and Overexpression

BL21 (pREP4-*groESL*) bacteria that contain a plasmid derived from pGEX-6P-3 allow the overproduction of recombinant proteins containing a N-terminal GST tag. They also overproduce the chaperone proteins GroES and GroEL

helping to fold overproduced proteins.

A loop of bacteria were used to inoculate 5 mL of LB medium containing ampicillin (pGEX selection) and kanamycin (pREP4groESL selection). After an overnight incubation at 37°C under stirring, the preculture was used to inoculate 1/50 of 100 mL of LB medium with ampicillin and kanamycin. The culture was incubated at 37°C under stirring until bacterial exponential growth phase ( $OD_{600nm} = 0.7$ ). Around  $3 \cdot 10^8$  bacteria was taken (Uninduced protein extract), centrifuged at 12,000 rpm for 5 minutes and pellet was resuspended with 40  $\mu$ L of SDS-PAGE 2X deposition buffer (125 mM Tris pH 6.8, 2 % SDS, 20 % glycerol, 0.005 % bromophenol blue, 0.1 M  $\beta$ -mercaptoethanol) and stored at room temperature until further analysis by SDS-PAGE.

The overproduction of the target protein was induced by addition of 1 mM IPTG to the culture and the latter was incubated under agitation for 4 h at 20°C. After 4 hours,  $3 \cdot 10^8$  bacteria were collected and treated as before (Induced protein extract), remaining cells were centrifuged at 6000 g for 10 minutes at 4°C. Cell pellet was washed with 10 mL of ice-cold ultrapure water, resuspended in ice-cold PBS and finally lysed twice in a French Press. Samples were then centrifuged 30 minutes at 14 000 g at 4°C and the supernatant was kept. To evaluate the solubility of the recombinant protein, 10  $\mu$ L of supernatant (Lysate) can be taken and added to 10  $\mu$ L of SDS-PAGE 2X deposition buffer and store at room temperature until analysis by SDS-PAGE.

## GST purification

Purification was done incubating the protein extract for 4 hours at 4°C with a glutathione-agarose-4B (Macherey-Nagel, 745500.10) resin previously equilibrated with ice-cold PBS. Proteins which have not interact with the matrix were removed by three washes of PBS (kept for analysis). Finally beads were resuspended in PBS and put in a regular column.

## Elution, GST cleavage and check

Elution of the protein of interest was done by cleaved the GST tag using PreScission Protease (Merck, GE27-0843-01). 160 units of PreScission Protease and 920  $\mu$ L of PreScission cleavage buffer were added into the column and incubated overnight at 4°C. Elution will contain only the protein of interest and no GST which still bind to the matrix and PreScission Protease.

The different fractions of the purification (lysate, washes and elution) were analysed by SDS-PAGE (10 $\mu$ L of each fraction with 10  $\mu$ L of SDS-PAGE 2X deposition buffer), to evaluate quality of the purification. Fractions containing the protein of interest were then dialysed using a membrane (Spectra/por<sup>®</sup> Cellulose Ester cut-off: 5000-Spectrum) against 1000 volumes of dialysis buffer (50 mM Tris HCl pH 7.5, 150 mM NaCl, 10 % glycerol) overnight at 4°C.

## Overexpression and purification of 6His-tagged protein

BL21 (pREP4groESL) bacteria containing pQE30 allow the overproduction of recombinant proteins containing a N-terminal 6-histidine (6His). The protein and protein extract shall be prepared according to the protocol in section [III](#).

### 6His purification

The bacterial pellet was resuspended in A buffer A (50 mM Tris HCl pH 7.5, 150 mM NaCl, 1 mM  $\beta$ -mercaptoethanol, 1 mM EDTA, 10 % glycerol, 0.1 % NP40). Purification was done incubating the bacterial lysate for 4 hours with Nickel Magnetic Beads (PureProteome Nickel Magnetic Beads, Millipore, LSKMAGH02), previously equilibrated with A buffer. Proteins which have not interact with the matrix were removed by three washes of A Buffer supplemented with 10 mM imidazole (kept for analysis). Protein of interest was eluted three times using 150  $\mu$ L of A buffer supplemented with 150mM imidazole. The different fractions were analysed by SDS-PAGE. The eluted proteins are then dialysed against 1000 volumes of dialysis buffer, as described above.

## Cell lysate

### For *in vitro* interaction test

HeLa or U937 cells were lysed with RIPA buffer (50 mM Tris pH 7.5, 150 mM NaCl, 0.5 % sodium deoxycholate) supplemented with 250U of Benzonase<sup>®</sup> DNase (70746, Millipore), 2.5 mM of MgCl<sub>2</sub>, and protease inhibitor cocktail. Cells were then incubated for 1 hour at 4°C with stirring, followed by 15 min at room temperature. After centrifugation at maximum speed, the supernatant was kept, stored on ice and use immediately.

### For GFP-Trap<sup>®</sup> interaction test

A day after transfection, HeLa cells were washed twice with ice-cold PBS, scrapped and collected. After centrifugation, cells pellet was resuspended in 200  $\mu$ L of ice-cold lysis buffer made of 10 mM Tris HCl pH 7.5, 150 mM NaCl, 0.5 % NP40, 0.5 mM EDTA, supplemented with 1  $\mu$ L of Benzonase<sup>®</sup> DNase, 2.5 mM MgCl<sub>2</sub>, protease inhibitors cocktail and 1 mM PMSF (PMSF-RO, Roche). Tubes were then placed on ice for 1 hour and half with extensively pipetting every 15 min. After centrifugation at 20.000 g for 10 min at 4°C, the supernatant was kept, stored on ice and use immediately.

## *In vitro* Phosphorylation test

In this study, phosphorylation tests were based on Western-Blot analysis using an anti-phosphothreonine antibody to detect the presence of phosphorylated threonine residues.

Phosphorylation tests were performed in a phosphorylation medium containing 25 mM Tris HCl pH 7.0, 5 mM  $MnCl_2$ , 1 mM EDTA, 5 mM DTT, 50  $\mu$ M ATP, 2  $\mu$ g of overproduced protein kinase 6His-LegK4 or 6His-LegK4<sub>D195N</sub> and 10  $\mu$ g of the target protein. Samples were incubated for 30min at 37°C and reaction was stopped by addition of 2X SDS-PAGE buffer.

## Interaction test

### GST pull-down

Based on the capture of a GST labelled protein bait by a glutathione-agarose resin, GST-pull down is a technique for studying protein protein interactions from a cell lysate. Also called partners, proteins interacting with the bait can be then identified by SDS-PAGE and/or Western-blot.

100  $\mu$ g of GST-purified protein were mixed with 1 mL of total HeLa or U937 cell extract 1 hour at 4°C under gentle agitation, then 3 hours at 4°C after addition of glutathion beads. Then, beads were washed three times with PBS to remove all unbound proteins. Elution of the GST fusion protein and any proteins bound to it was done adding 50  $\mu$ L of 20 mM reduced glutathione in 50 mM Tris-HCl pH 8.0 to the beads.

### 6His Pull-down

100  $\mu$ g of 6His-purified protein (6His-LegK4 or 6His-LegK4<sub>D195N</sub>) were mixed with 1 mL of total HeLa or U937 cell extract 1 hour at 4°C under gentle agitation, then 3 hours at 4°C after addition of Nickel magnetic beads. The beads are then washed three times with A buffer containing 10 mM imidazole to remove all unbound proteins. Elution was performed by adding 100  $\mu$ L of 150 mM imidazole and collected in a new tube containing 25  $\mu$ L of 5X SDS-PAGE buffer.

### GFP-Trap<sup>®</sup>

Cell extract was diluted in 300  $\mu$ L of dilution buffer which composition is similar to lysis buffer without NP40 and EDTA. 50  $\mu$ L of diluted lysate was kept as our input control. In the meantime, 25  $\mu$ L of GFP-Trap<sup>®</sup> agarose beads were equilibrated three times with dilution buffer and then add to cell extract. Tubes tumble end-over-end vertically for 1 hour at 4°C then 30 minutes horizontally. Three washes were performed using 500  $\mu$ L of ice-cold dilution buffer. Interactants were eluted from the beads with 100  $\mu$ L of 2X SDS-PAGE buffer. SDS-PAGE and Western-Blot were performed using both cell lysate and elution.



## RNA extraction

Bacteria were prepared to macrophages infection at MOI 10 as written in [III](#). After 45 minutes post-infection, RNA were extracted according to the Direct-zol RNA Miniprep (R2051, Zymo<sup>®</sup>) protocol: The medium was removed and Trizol<sup>®</sup> was added to lysis the cells for 5 minutes at room temperature. Then, cells were centrifuged for 2 minutes at 10 000 rpm to remove particulate matter. Supernatant were pour into a 2 mL eppendorf and absolute ethanol was added. Mixture was added into a column placed in a collecting tube and centrifuge at maximum speed 30 secondes. Direct-zol RNA PreWash was added to the column and centrifuge at maximum speed 30 secondes twice. Then, RNA Wash Buffer was added into the column, centrifuge for 1 minutes and again for 2 minutes. Finally, 40  $\mu$ L of DNase/RNase-Free Water was put directly to the center of the column to elute RNAs. Genomic DNA was digested and eliminate thanks to DNase activity and RNA was purified other time following the same process.

## Protein analysis

Proteins concentration was determined using spectrophotometry (Nanodrop 2000 (Thermo) at OD<sub>280nm</sub>).

## Polyacrylamide gel electrophoresis SDS-PAGE

Under an electric field, the constant negative charges provided by the SDS and reduction of disulfide bridges in the presence of  $\beta$ -mercaptoethanol allow proteins to be electrophoretically mobile towards the anode and inverse function of their molecular weight. The polyacrylamide gel consists of a stacking gel 10 % acrylamide/disacrylamide (Roth,3029.1), 0.125 M Tris-HCl pH 6.8, 0.1 %, 0.1 % ammonium persulfate (Roth, 9592.3), 0.1 % TEMED (Roth, 2367.3) and a separation gel (identical composition except for the acrylamide/bisacrylamide (30/0.8) concentration from 8 to 12 % and the 0.375M Tris-HCl buffer at pH 8.8). 10  $\mu$ L of SDS-PAGE 2X deposition buffer (Laemmli, 1970) is added to 10  $\mu$ L of protein sample to be analyzed. The proteins are then denatured by heating for 5 minutes at 100°C and put in the electrophoresis gel. The molecular weight marker used is the PageRuler Prestained Protein Ladder (Fermentas, 26617). The electrophoresis is conducted in migration buffer (25 mM Tris, 192 mM Glycine, 0.1% SDS) at 35 mA until proteins pass through the separation gel and at 45 mA for 1 hour at room temperature.

## Coomassie Blue staining

Depending the experiment, the proteins can be stained on the gel by immersing it for 20 min under gentle agitation in a colouring solution (10 % acetic acid, 40 % ethanol, 0.04 % Coomassie Brilliant Blue R250). The gel is then destained 30 min to 1 hour under gentle agitation in a solution made of 10 % acetic acid and 5 % ethanol.

## Protein transfer on membrane

After SDS-PAGE, proteins can be transferred to a nitrocellulose membrane (Amersham™ Protran®<sup>®</sup>, GE10600118). The membrane as well as Whatman papers were soaked with Bjerrum and Schafer-Nielsen Transfer Buffer made of 40 mM Glycine, 48 mM Tris base, 130  $\mu$ M SDS and 20 % ethanol.

Then, the following sandwich was made with Whatman 3M paper:

Cathode (-)

———— Three pre-soaked Whatman sheets

———— SDS-PAGE electrophoresis gel

———— pre-soaked nitrocellulose membrane

———— Three pre-soaked Whatman sheets

Anode (+)

Proteins were transferred in a Trans-Blot® Turbo™ (BioRad) semi-dry chamber, at constant intensity for 30 min (25V-1.0A). Transferred proteins on nitrocellulose can then be revealed using specific antibodies.

## Red Ponceau staining

Red Ponceau staining is a quick and reversible staining method used to detect protein bands on nitrocellulose membranes. Red Ponceau is negatively charged allowing it to bind to the positively charged functional groups of the protein and non-polar regions of the protein.

After the electotransfert, the blot was put in plastic box and rinse it with water three times, five minutes each. Then, it was stained with Red Ponceau (Interchim, 050260) for 30 seconds to 1 minute. Finally, to reveal protein bands the blot was destained with several changes of water for 30 seconds to 1 minute each.

## Immunorevelation of proteins (Western Blot)

After the electotransfert, the blot was blocked for 1 hour under rotation in TBS (100 mM Tris pH 8, 150 mM NaCl) supplemented with 5 % BSA. After three 5 minutes washes in 0.1% Tween20 (Sigma, P9416) TBS, the blot was incubated at 4°C overnight under agitation in a 0.1% Tween20 0.5 % BSA TBS solution containing the primary antibody (Table 10). Then, the blot was washed again three times with 0.1% Tween20 TBS. The membrane was then incubated for 1 hour under agitation at room temperature in 0.1% Tween20 0.5 % BSA TBS in the presence of the secondary antibody conjugated to peroxidase and then rinsed three times with 0.1% Tween20 TBS. Revelation was performed using the SuperSignal West Pico Chemiluminescent Substrate kit (34579, Pierce). Signals were then detected by direct reading on ChemiStart 5000 (Fisher Bioblock Scientific).

## Characterization of LegK4 function

### Localization by immunofluorescence

Cells were seeded on glass coverslips and transfected. 24 hours post-transfection, Cells were washed twice in PBS, fixed for 20 minutes at room temperature with 4% paraformaldehyde diluted in PBS and washed again three times with PBS. A blocking step of the free aldehyde groups was performed by incubating coverslips twice for 10 minutes in PBS-0.1 % glycine and washed twice. Then, cells were permeabilized with PBS-0.1 % X100 Triton for 5 minutes at room temperature. After a washing step with PBS, the non-specific sites are blocked for 1 hour with PBS-1 % BSA. Coverslips were incubated with the primary antibodies in PBS-1 % BSA for 1 hour in a wet chamber, washed three times and incubated with secondary antibody coupled with the appropriate fluorochrome and 1/100 DAPI to stain the nuclei for 30 minutes. Observations were made on an inverted confocal microscope (Axiovert 200M; Zeiss, Thornwood, USA).

### Phenotype

- U937

After counting cell with a Mallasez cell, cell concentration was adjusted and plated 24 hours before infection depending on the experiment.

U937 cells, 2.5  $\mu$ L of PMA per 50 mL of cell suspension was added away from any light for 72 hours.

- HeLa or A549

HeLa or A549 have been plated 24 hours prior the infection.

3 days prior infection, bacterial strains were plated on CYE plates and again on new CYE supplemented with IPTG for 24 hours. From this plate, bacterial suspension were made in RPMI supplemented with or without IPTG or antibiotics, OD was measured to calibrate the experiment. Tubes were at 37°C without any agitation for 2 hours to stress bacteria.

Finally, 100  $\mu$ L of bacterial suspension were added in each well, plates were centrifuged at 1500 rpm for 5 minutes at 25°C and incubated at 37°C in  $CO_2$  incubator for the desired infection time.



# References

- [1] D. W. Fraser, T. R. Tsai, W. Orenstein, W. E. Parkin, H. J. Beecham, R. G. Sharrar, J. Harris, G. F. Mallison, S. M. Martin, J. E. Mcdade, C. C. Shepard, and P. S. Brachman, "Legionnaires' Disease: Description of an Epidemic of Pneumonia," *New England Journal of Medicine*, vol. 297, pp. 1189–1197, dec 1977.
- [2] J. E. Mcdade, C. C. Shepard, D. W. Fraser, T. R. Tsai, M. A. Redus, and W. R. Dowdle, "Legionnaires' Disease: Isolation of a Bacterium and Demonstration of Its Role in Other Respiratory Disease," *New England Journal of Medicine*, vol. 297, pp. 1197–1203, dec 1977.
- [3] A. Khodr, E. Kay, L. Gomez-Valero, C. Ginevra, P. Doublet, C. Buchrieser, and S. Jarraud, "Molecular epidemiology, phylogeny and evolution of Legionella," sep 2016.
- [4] T. H. Glick, M. B. Gregg, B. Berman, G. Mallison, W. W. Rhodes JR, and I. Kassanoff, "Pontiac fever: an epidemic of unknown etiology in a health department: I. clinical and epidemiologic aspects," *American Journal of Epidemiology*, vol. 107, no. 2, pp. 149–160, 1978.
- [5] CDC, "Legionnaires Disease Outbreaks — Legionella — CDC." <https://www.cdc.gov/legionella/outbreaks.html>.
- [6] U. Zähringer, Y. A. Knirel, B. Lindner, J. H. Helbig, A. Sonesson, R. Marre, and E. T. Rietschel, "The lipopolysaccharide of Legionella pneumophila serogroup 1 (strain Philadelphia 1): chemical structure and biological significance.," *Progress in clinical and biological research*, vol. 392, pp. 113–39, 1995.
- [7] V. L. Yu, J. F. Plouffe, M. C. Pastoris, J. E. Stout, M. Schousboe, A. Widmer, J. Summersgill, T. File, C. M. Heath, D. L. Paterson, and A. Cheresky, "Distribution of Legionella Species and Serogroups Isolated by Culture in Patients with Sporadic Community-Acquired Legionellosis: An International Collaborative Survey," *The Journal of Infectious Diseases*, vol. 186, pp. 127–128, jul 2002.
- [8] O. Duron, P. Doublet, F. Vavre, and D. Bouchon, "The importance of revisiting legionellales diversity," *Trends in parasitology*, vol. 34, no. 12, pp. 1027–1037, 2018.
- [9] W. G. Weisburg, M. E. Dobson, J. E. Samuel, G. A. Dasch, L. P. Mallavia, O. Baca, L. Mandelco, J. E. Sechrest, E. Weiss, and C. R. Woese, "Phylogenetic diversity of the Rickettsiae.," *Journal of bacteriology*, vol. 171, pp. 4202–6, aug 1989.

- [10] F. W. Chandler, J. A. Blackmon, M. D. Hicklin, R. M. Cole, and C. S. Callaway, "Ultrastructure of the Agent of Legionnaires' Disease in the Human Lung," *American Journal of Clinical Pathology*, vol. 71, pp. 43–50, jan 1979.
- [11] G. A. Hébert, C. S. Callaway, E. P. Ewing, and Jr, "Comparison of Legionella pneumophila, L. micdadei, L. bozemanii, and L. dumoffii by transmission electron microscopy.," *Journal of clinical microbiology*, vol. 19, pp. 116–121, feb 1984.
- [12] R. H. Doust, M. Mobarez, H. Bagheri, and N. Khoramabadi, "Interaction of legionellae and free-living amoebae within hospital water supplies," *Research Journal of Parasitology*, vol. 3, pp. 104–113, 03 2008.
- [13] M. Taylor, K. Ross, and R. Bentham, "Legionella, protozoa, and biofilms: interactions within complex microbial systems," *Microbial ecology*, vol. 58, no. 3, pp. 538–547, 2009.
- [14] V. Thomas, T. Bouchez, V. Nicolas, S. Robert, J. Loret, and Y. Levi, "Amoebae in domestic water systems: resistance to disinfection treatments and implication in legionella persistence," *Journal of applied microbiology*, vol. 97, no. 5, pp. 950–963, 2004.
- [15] S. Mondino, S. Schmidt, and C. Buchrieser, "Molecular mimicry: a paradigm of host-microbe coevolution illustrated by legionella," *Mbio*, vol. 11, no. 5, pp. e01201–20, 2020.
- [16] WHO, "Légionellose "<https://www.who.int/fr/news-room/fact-sheets/detail/legionellosis>"."
- [17] M. Shaheen, C. Scott, and N. J. Ashbolt, "Long-term persistence of infectious legionella with free-living amoebae in drinking water biofilms," *International journal of hygiene and environmental health*, vol. 222, no. 4, pp. 678–686, 2019.
- [18] B. A. Cunha, A. Burillo, and E. Bouza, "Legionnaires' disease," *The Lancet*, vol. 387, no. 10016, pp. 376–385, 2016.
- [19] L. Gavaldà, M. Garcia-Nuñez, S. Quero, C. Gutierrez-Milla, and M. Sabrià, "Role of hot water temperature and water system use on legionella control in a tertiary hospital: An 8-year longitudinal study," *Water research*, vol. 149, pp. 460–466, 2019.
- [20] D. A. Lytle, S. Pfaller, C. Muhlen, I. Struewing, S. Triantafyllidou, C. White, S. Hayes, D. King, and J. Lu, "A comprehensive evaluation of monochloramine disinfection on water quality, legionella and other important microorganisms in a hospital," *Water Research*, vol. 189, p. 116656, 2021.
- [21] D. Wüthrich, S. Gautsch, R. Spieler-Denz, O. Dubuis, V. Gaia, J. Moran-Gilad, V. Hinic, H. M. Seth-Smith, C. H. Nickel, S. Tschudin-Sutter, *et al.*, "Air-conditioner cooling towers as complex reservoirs and continuous source of legionella pneumophila infection evidenced by a genomic analysis study in 2017, switzerland," *Eurosurveillance*, vol. 24, no. 4, p. 1800192, 2019.
- [22] A. M. Correia, J. S. Ferreira, V. Borges, A. Nunes, B. Gomes, R. Capucho, J. Gonçalves, D. M. Antunes, S. Almeida, A. Mendes, M. Guerreiro, D. A. Sampaio, L. Vieira, J. Machado, M. J. Simões, P. Gonçalves, and J. P. Gomes,

- “Probable Person-to-Person Transmission of Legionnaires’ Disease,” *New England Journal of Medicine*, vol. 374, pp. 497–498, feb 2016.
- [23] A. Dagan, D. Epstein, A. Mahagneh, J. Nashashibi, Y. Geffen, A. Neuberger, and A. Miller, “Community-acquired versus nosocomial legionella pneumonia: factors associated with legionella-related mortality,” *European Journal of Clinical Microbiology & Infectious Diseases*, pp. 1–8, 2021.
- [24] D. M. Pierre, J. Baron, V. L. Yu, and J. E. Stout, “Diagnostic testing for Legionnaires’ disease.,” *Annals of clinical microbiology and antimicrobials*, vol. 16, p. 59, aug 2017.
- [25] C. Campèse, G. Descours, A. Lepoutre, L. Beraud, C. Maine, D. Che, and S. Jarraud, “Legionnaires’ disease in France,” *Médecine et Maladies Infectieuses*, vol. 45, pp. 65–71, mar 2015.
- [26] F. F. Graham, P. S. White, D. J. G. Harte, and S. P. Kingham, “Changing epidemiological trends of legionellosis in New Zealand, 1979-2009.,” *Epidemiology and infection*, vol. 140, pp. 1481–96, aug 2012.
- [27] O. A. T. Alli, S. Zink, N. K. von Lackum, and Y. Abu-Kwaik, “Comparative assessment of virulence traits in Legionella spp.,” *Microbiology*, vol. 149, pp. 631–641, mar 2003.
- [28] A. Doleans, H. Aurell, M. Reyrolle, G. Lina, J. Freney, F. Vandenesch, J. Etienne, and S. Jarraud, “Clinical and environmental distributions of legionella strains in france are different,” *Journal of Clinical Microbiology*, vol. 42, no. 1, pp. 458–460, 2004.
- [29] N. Phin, F. Parry-Ford, T. Harrison, H. R. Stagg, N. Zhang, K. Kumar, O. Lortholary, A. Zumla, and I. Abubakar, “Epidemiology and clinical management of Legionnaires’ disease,” *The Lancet Infectious Diseases*, vol. 14, pp. 1011–1021, oct 2014.
- [30] O. Pappa, D. Chochlakis, V. Sandalakis, C. Dioli, A. Psaroulaki, and A. Mavridou, “Antibiotic resistance of legionella pneumophila in clinical and water isolates—a systematic review,” *International journal of environmental research and public health*, vol. 17, no. 16, p. 5809, 2020.
- [31] C. Massip, G. Descours, C. Ginevra, P. Doublet, S. Jarraud, and C. Gilbert, “Macrolide resistance in Legionella pneumophila: The role of LpeAB efflux pump,” *Journal of Antimicrobial Chemotherapy*, vol. 72, pp. 1327–1333, jan 2017.
- [32] M. M. L. Maurin, D. Schneider, L. Shadoud, S. Jarraud, J. F. Timsit, and J. Etienne, “Method for the in vitro detection of strains of legionella pneumophila resistant to antibiotics,” Oct. 23 2018. US Patent 10,106,858.
- [33] S. P. France, “La légionellose, notre action.” <https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/legionellose/notre-action/#tab>.
- [34] S. P. France, “La légionellose : données.” <https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/legionellose/donnees/#tabs>.

- [35] L. Y. Gao, O. S. Harb, and Y. Abu Kwaik, "Utilization of similar mechanisms by *Legionella pneumophila* to parasitize two evolutionarily distant host cells, mammalian macrophages and protozoa.," *Infection and immunity*, vol. 65, pp. 4738–4746, nov 1997.
- [36] A. M. Copenhaver, C. N. Casson, H. T. Nguyen, T. C. Fung, M. M. Duda, C. R. Roy, and S. Shin, "Alveolar macrophages and neutrophils are the primary reservoirs for *Legionella pneumophila* and mediate cytosolic surveillance of type IV secretion.," *Infection and immunity*, vol. 82, pp. 4325–36, oct 2014.
- [37] B. M. THOMASON, F. W. CHANDLER, and D. G. HOLLIS, "Flagella on legionnaires' disease bacteria: an interim report.," *Annals of Internal Medicine*, vol. 91, no. 2, pp. 224–226, 1979.
- [38] N. R. Payne and M. A. Horwitz, "Phagocytosis of *legionella pneumophila* is mediated by human monocyte complement receptors.," *The Journal of experimental medicine*, vol. 166, no. 5, pp. 1377–1389, 1987.
- [39] C. Krinos, A. High, and F. Rodgers, "Role of the 25 kda major outer membrane protein of *legionella pneumophila* in attachment to u-937 cells and its potential as a virulence factor for chick embryos," *Journal of applied microbiology*, vol. 86, no. 2, pp. 237–244, 1999.
- [40] L. Vandersmissen, E. De Buck, V. Saels, D. A. Coil, and J. Anné, "A *legionella pneumophila* collagen-like protein encoded by a gene with a variable number of tandem repeats is involved in the adherence and invasion of host cells," *FEMS microbiology letters*, vol. 306, no. 2, pp. 168–176, 2010.
- [41] S. L. Cirillo, J. Lum, and J. D. Cirillo, "Identification of novel loci involved in entry by *legionella pneumophila* the genbank accession numbers for the *enh1* and *enh2* loci reported in this paper are af057703 and af057704, respectively.," *Microbiology*, vol. 146, no. 6, pp. 1345–1359, 2000.
- [42] H. J. Newton, F. M. Sansom, J. Dao, A. D. McAlister, J. Sloan, N. P. Cianciotto, and E. L. Hartland, "Sel1 repeat protein LpnE is a *Legionella pneumophila* virulence determinant that influences vacuolar trafficking," *Infection and Immunity*, vol. 75, pp. 5575–5585, dec 2007.
- [43] B. J. Stone and Y. A. 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," *Infection and immunity*, vol. 66, no. 4, pp. 1768–1775, 1998.
- [44] G. D'Auria, N. Jiménez-Hernández, F. Peris-Bondia, A. Moya, and A. Latorre, "*Legionella pneumophila* pangenome reveals strain-specific virulence factors," *BMC Genomics*, vol. 11, p. 181, mar 2010.
- [45] L. Gomez-Valero, C. Rusniok, C. Cazalet, and C. Buchrieser, "Comparative and functional genomics of *legionella* identified eukaryotic like proteins as key players in host–pathogen interactions," *Frontiers in microbiology*, vol. 2, p. 208, 2011.
- [46] M. Watarai, I. Derre, J. Kirby, J. D. Gowney, W. F. Dietrich, and R. R. Isberg, "*Legionella pneumophila* is internalized by a macropinocytotic uptake pathway controlled by the Dot/Icm system and the mouse *Lgn1* locus.," *The Journal of experimental medicine*, vol. 194, pp. 1081–1096, oct 2001.



- [47] S. Weber, M. Wagner, and H. Hilbi, "Live-cell imaging of phosphoinositide dynamics and membrane architecture during legionella infection," *MBio*, vol. 5, no. 1, 2014.
- [48] B. E. García-Pérez, J. J. De la Cruz-López, J. I. Castañeda-Sánchez, A. R. Muñoz-Duarte, A. D. Hernández-Pérez, H. Villegas-Castrejón, E. García-Latorre, A. Caamal-Ley, and J. Luna-Herrera, "Macropinocytosis is responsible for the uptake of pathogenic and non-pathogenic mycobacteria by B lymphocytes (Raji cells)," *BMC Microbiology*, vol. 12, p. 246, oct 2012.
- [49] K. Maruta, M. Ogawa, H. Miyamoto, K. Izu, and S.-I. Yoshida, "Entry and intracellular localization of legionella dumoffii in vero cells," *Microbial pathogenesis*, vol. 24, no. 2, pp. 65–73, 1998.
- [50] E. Veiga and P. Cossart, "The role of clathrin-dependent endocytosis in bacterial internalization," *Trends in cell biology*, vol. 16, no. 10, pp. 499–504, 2006.
- [51] M. S. Swanson and E. Fernandez-Moreira, "A microbial strategy to multiply in macrophages: the pregnant pause," *Traffic*, vol. 3, no. 3, pp. 170–177, 2002.
- [52] G. Oliva, 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*, vol. 8, p. 3, 2018.
- [53] M. V. Fonseca and M. S. Swanson, "Nutrient salvaging and metabolism by the intracellular pathogen legionella pneumophila," *Frontiers in cellular and infection microbiology*, vol. 4, p. 12, 2014.
- [54] T. Sahr, C. Rusniok, F. Impens, G. Oliva, O. Sismeiro, J.-Y. Coppée, and C. Buchrieser, "The legionella pneumophila genome evolved to accommodate multiple regulatory mechanisms controlled by the csra-system," *PLoS genetics*, vol. 13, no. 2, p. e1006629, 2017.
- [55] R. Hochstrasser and H. Hilbi, "Intra-species and inter-kingdom signaling of legionella pneumophila," *Frontiers in microbiology*, vol. 8, p. 79, 2017.
- [56] L. Gomez-Valero and C. Buchrieser, "Intracellular parasitism, the driving force of evolution of Legionella pneumophila and the genus Legionella," *Genes & Immunity*, vol. 20, pp. 394–402, may 2019.
- [57] K. H. Berger, J. J. Merriam, and R. R. Isberg, "Altered intracellular targeting properties associated with mutations in the Legionella pneumophila dotA gene," *Molecular Microbiology*, vol. 14, pp. 809–822, nov 1994.
- [58] M. Chien, I. Morozova, S. Shi, H. Sheng, J. Chen, S. M. Gomez, G. Asamani, K. Hill, J. Nuara, M. Feder, J. Rineer, J. J. Greenberg, V. Steshenko, S. H. Park, B. Zhao, E. Teplitskaya, J. R. Edwards, S. Pampou, A. Georghiou, I.-C. Chou, W. Iannuccilli, M. E. Ulz, D. H. Kim, A. Geringer-Sameth, C. Goldsberry, P. Morozov, S. G. Fischer, G. Segal, X. Qu, A. Rzhetsky, P. Zhang, E. Cayanis, P. J. De Jong, J. Ju, S. Kalachikov, H. A. Shuman, and J. J. Russo, "The Genomic Sequence of the Accidental Pathogen Legionella pneumophila," *Science*, vol. 305, pp. 1966–1968, sep 2004.

- [59] C. Lawrence, M. Reyrolle, S. Dubrou, F. Forey, B. Decludt, C. Goulvestre, P. Matsiota-Bernard, J. Etienne, and C. Nauciel, "Single clonal origin of a high proportion of *Legionella pneumophila* serogroup 1 isolates from patients and the environment in the area of Paris, France, over a 10-year period.," *Journal of clinical microbiology*, vol. 37, pp. 2652–5, aug 1999.
- [60] H. Aurell, J. Etienne, F. Forey, M. Reyrolle, P. Girardo, P. Farge, B. Decludt, C. Campese, F. Vandenesch, and S. Jarraud, "Legionella pneumophila Serogroup 1 Strain Paris: Endemic Distribution throughout France," *Journal of Clinical Microbiology*, vol. 41, pp. 3320–3322, jul 2003.
- [61] C. Cazalet, S. Jarraud, Y. Ghavi-Helm, F. Kunst, P. Glaser, J. Etienne, and C. Buchrieser, "Multigenome analysis identifies a worldwide distributed epidemic *Legionella pneumophila* clone that emerged within a highly diverse species.," *Genome research*, vol. 18, pp. 431–41, mar 2008.
- [62] T. M. Nhu Nguyen, D. Ilef, S. Jarraud, L. Rouil, C. Campese, D. Che, S. Haeghebaert, F. Ganiayre, F. Marcel, J. Etienne, and J. Desenclos, "A Community-Wide Outbreak of Legionnaires Disease Linked to Industrial Cooling Towers—How Far Can Contaminated Aerosols Spread?," *The Journal of Infectious Diseases*, vol. 193, pp. 102–111, jan 2006.
- [63] C. Cazalet, C. Rusniok, H. Brüggemann, N. Zidane, A. Magnier, L. Ma, M. Tichit, S. Jarraud, C. Bouchier, F. Vandenesch, *et al.*, "Evidence in the legionella pneumophila genome for exploitation of host cell functions and high genome plasticity," *Nature genetics*, vol. 36, no. 11, pp. 1165–1173, 2004.
- [64] M. Steinert, K. Heuner, C. Buchrieser, C. Albert-Weissenberger, and G. Glöckner, "Legionella pathogenicity: genome structure, regulatory networks and the host cell response.," *International journal of medical microbiology : IJMM*, vol. 297, pp. 577–87, nov 2007.
- [65] G. Glöckner, C. Albert-Weissenberger, E. Weinmann, S. Jacobi, E. Schunder, M. Steinert, J. Hacker, and K. Heuner, "Identification and characterization of a new conjugation/type IVA secretion system (trb/tra) of *Legionella pneumophila* Corby localized on two mobile genomic islands.," *International journal of medical microbiology : IJMM*, vol. 298, pp. 411–28, jul 2008.
- [66] G. N. Schroeder, N. K. Petty, A. Mousnier, C. R. Harding, A. J. Vogrin, B. Wee, N. K. Fry, T. G. Harrison, H. J. Newton, N. R. Thomson, *et al.*, "Legionella pneumophila strain 130b possesses a unique combination of type iv secretion systems and novel dot/icm secretion system effector proteins," *Journal of bacteriology*, vol. 192, no. 22, pp. 6001–6016, 2010.
- [67] C. Ginevra, F. Forey, C. Campèse, M. Reyrolle, D. Che, J. Etienne, and S. Jarraud, "Lorraine Strain of *Legionella pneumophila* Serogroup 1, France," *Emerging Infectious Diseases*, vol. 14, pp. 673–675, apr 2008.
- [68] L. Gomez-Valero, C. Rusniok, S. Jarraud, B. Vacherie, Z. Rouy, V. Barbe, C. Medigue, J. Etienne, and C. Buchrieser, "Extensive recombination events and horizontal gene transfer shaped the *Legionella pneumophila* genomes," *BMC Genomics*, vol. 12, p. 536, dec 2011.

- [69] L. Gomez-Valero and C. Buchrieser, "Genome dynamics in legionella: the basis of versatility and adaptation to intracellular replication," *Cold Spring Harbor Perspectives in Medicine*, vol. 3, no. 6, p. a009993, 2013.
- [70] D. E. Voth, L. J. Broederdorf, and J. G. Graham, "Bacterial type iv secretion systems: versatile virulence machines," *Future microbiology*, vol. 7, no. 2, pp. 241–257, 2012.
- [71] K. H. Berger and R. R. Isberg, "Two distinct defects in intracellular growth complemented by a single genetic locus in *Legionella pneumophila*," *Molecular Microbiology*, vol. 7, pp. 7–19, jan 1993.
- [72] A. Marra, S. J. Blander, M. A. Horwitz, and H. A. Shuman, "Identification of a *Legionella pneumophila* locus required for intracellular multiplication in human macrophages," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 89, no. 20, pp. 9607–9611, 1992.
- [73] S. D. Zink, L. Pedersen, N. P. Cianciotto, and Y. Abu-Kwaik, "The Dot/Icm type IV secretion system of *Legionella pneumophila* is essential for the induction of apoptosis in human macrophages.," *Infection and immunity*, vol. 70, pp. 1657–63, mar 2002.
- [74] K. C. Jeong, D. Ghosal, Y. W. Chang, G. J. Jensen, and J. P. Vogel, "Polar delivery of *Legionella* type IV secretion system substrates is essential for virulence," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 114, pp. 8077–8082, jul 2017.
- [75] D. Ghosal, Y.-W. Chang, K. C. Jeong, J. P. Vogel, and G. J. Jensen, "In situ structure of the *Legionella* Dot/Icm type IV secretion system by electron cryotomography.," *EMBO reports*, vol. 18, no. 5, pp. 726–732, 2017.
- [76] C. D. Vincent, J. R. Friedman, K. C. Jeong, E. C. Buford, J. L. Miller, and J. P. Vogel, "Identification of the core transmembrane complex of the *Legionella* Dot/Icm type IV secretion system," *Molecular Microbiology*, vol. 62, pp. 1278–1291, dec 2006.
- [77] K. V. Korotkov, J. R. Delarosa, and W. G. Hol, "A dodecameric ring-like structure of the N0 domain of the type II secretin from enterotoxigenic *Escherichia coli*," *Journal of Structural Biology*, vol. 183, pp. 354–362, sep 2013.
- [78] K. C. Jeong, J. Gyore, L. Teng, D. Ghosal, G. J. Jensen, and J. P. Vogel, "Polar targeting and assembly of the *Legionella* Dot/Icm type IV secretion system (T4SS) by T6SS-related components," *bioRxiv*, p. 315721, 2018.
- [79] N. Nakano, T. Kubori, M. Kinoshita, K. Imada, and H. Nagai, "Crystal Structure of *Legionella* DotD: Insights into the Relationship between Type IVB and Type II/III Secretion Systems," *PLoS Pathogens*, vol. 6, p. e1001129, oct 2010.
- [80] D. Ghosal, Y.-W. Chang, K. C. Jeong, J. P. Vogel, and G. J. Jensen, "Molecular architecture of the legionella dot/icm type iv secretion system," *BioRxiv*, p. 312009, 2018.
- [81] Z.-Q. Luo and R. R. Isberg, "Multiple substrates of the *Legionella pneumophila* Dot/Icm system identified by interbacterial protein transfer.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, pp. 841–6, jan 2004.

- [82] J. Allombert, F. Fuche, C. Michard, and P. Doublet, "Molecular mimicry and original biochemical strategies for the biogenesis of a *Legionella pneumophila* replicative niche in phagocytic cells," *Microbes and Infection*, vol. 15, no. 14-15, pp. 981–988, 2013.
- [83] M. C. Sutherland, K. A. Binder, P. Y. Cuaing, and J. P. Vogel, "Reassessing the role of dotf in the legionella pneumophila type iv secretion system," *PloS one*, vol. 8, no. 6, p. e65529, 2013.
- [84] P. J. Christie, "The mosaic type iv secretion systems," *EcoSal Plus*, vol. 7, no. 1, 2016.
- [85] C. D. Vincent, J. R. Friedman, K. C. Jeong, M. C. Sutherland, and J. P. Vogel, "Identification of the DotL coupling protein subcomplex of the Legionella Dot/Icm type IV secretion system," *Molecular Microbiology*, vol. 85, pp. 378–391, jul 2012.
- [86] M.-J. Kwak, J. D. Kim, H. Kim, C. Kim, J. W. Bowman, S. Kim, K. Joo, J. Lee, K. S. Jin, Y.-G. Kim, N. K. Lee, J. U. Jung, and B.-H. Oh, "Architecture of the type IV coupling protein complex of *Legionella pneumophila*," *Nature microbiology*, vol. 2, p. 17114, jul 2017.
- [87] B. A. Buscher, G. M. Conover, J. L. Miller, S. A. Vogel, S. N. Meyers, R. R. Isberg, and J. P. Vogel, "The DotL protein, a member of the TraG-coupling protein family, is essential for Viability of *Legionella pneumophila* strain Lp02.," *Journal of bacteriology*, vol. 187, pp. 2927–2938, may 2005.
- [88] K. Atmakuri, E. Cascales, and P. J. Christie, "Energetic components VirD4, VirB11 and VirB4 mediate early DNA transfer reactions required for bacterial type IV secretion.," *Molecular microbiology*, vol. 54, pp. 1199–1211, dec 2004.
- [89] M. C. Sutherland, T. L. Nguyen, V. Tseng, and J. P. Vogel, "The legionella icmsw complex directly interacts with dotl to mediate translocation of adaptor-dependent substrates," *PLoS Pathog*, vol. 8, no. 9, p. e1002910, 2012.
- [90] H. J. Yeo, S. N. Savvides, A. B. Herr, E. Lanka, and G. Waksman, "Crystal structure of the hexameric traffic ATPase of the *Helicobacter pylori* type IV secretion system.," *Molecular cell*, vol. 6, pp. 1461–72, dec 2000.
- [91] A. Meir, D. Chetrit, L. Liu, C. R. Roy, and G. Waksman, "Legionella DotM structure reveals a role in effector recruiting to the Type 4B secretion system.," *Nature communications*, vol. 9, no. 1, p. 507, 2018.
- [92] S. Ninio, D. M. Zuckman-Cholon, E. D. Cambronne, and C. R. Roy, "The Legionella IcmS-IcmW protein complex is important for Dot/Icm-mediated protein translocation," *Molecular Microbiology*, vol. 55, pp. 912–926, dec 2004.
- [93] E. D. Cambronne and C. R. Roy, "The Legionella pneumophila IcmSW complex interacts with multiple Dot/Icm effectors to facilitate type IV translocation.," *PLoS pathogens*, vol. 3, p. e188, dec 2007.
- [94] H. Nagai and T. Kubori, "Type IVB Secretion Systems of Legionella and Other Gram-Negative Bacteria.," *Frontiers in microbiology*, vol. 2, p. 136, 2011.

- [95] J. Xu, D. Xu, M. Wan, L. Yin, X. Wang, L. Wu, Y. Liu, X. Liu, Y. Zhou, and Y. Zhu, "Structural insights into the roles of the IcmS-IcmW complex in the type IVb secretion system of *Legionella pneumophila*.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 114, pp. 13543–13548, dec 2017.
- [96] J. Allombert, C. Jaboulay, C. Michard, C. Andréa, X. Charpentier, A. Vianney, and P. Doublet, "Orchestrated delivery of *Legionella* effectors by the Icm/Dot secretion system,” *bioRxiv*, p. 754762, sep 2019.
- [97] G. Duménil, T. P. Montminy, M. Tang, and R. R. Isberg, "IcmR-regulated Membrane Insertion and Efflux by the *Legionella pneumophila* IcmQ Protein,” *Journal of Biological Chemistry*, vol. 279, pp. 4686–4695, feb 2004.
- [98] S. Raychaudhury, J. D. Farelli, T. P. Montminy, M. Matthews, J.-F. Ménétret, G. Duménil, C. R. Roy, J. F. Head, R. R. Isberg, and C. W. Akey, "Structure and function of interacting IcmR-IcmQ domains from a type IVb secretion system in *Legionella pneumophila*.,” *Structure (London, England : 1993)*, vol. 17, pp. 590–601, apr 2009.
- [99] G. Dumenil and R. R. Isberg, "The *Legionella pneumophila* IcmR protein exhibits chaperone activity for IcmQ by preventing its participation in high-molecular-weight complexes,” *Molecular Microbiology*, vol. 40, pp. 1113–1127, jun 2001.
- [100] M. Matthews and C. R. Roy, "Identification and subcellular localization of the *Legionella pneumophila* IcmX protein: a factor essential for establishment of a replicative organelle in eukaryotic host cells.,” *Infection and immunity*, vol. 68, pp. 3971–3982, jul 2000.
- [101] G. Segal, M. Feldman, and T. Zusman, "The icm/dot type-iv secretion systems of *legionella pneumophila* and *coxiella burnetii*,” *FEMS microbiology reviews*, vol. 29, no. 1, pp. 65–81, 2005.
- [102] R. K. Laguna, E. A. Creasey, Z. Li, N. Valtz, and R. R. Isberg, "A *legionella pneumophila*-translocated substrate that is required for growth within macrophages and protection from host cell death,” *Proceedings of the National Academy of Sciences*, vol. 103, no. 49, pp. 18745–18750, 2006.
- [103] Z. Lifshitz, D. Burstein, M. Peeri, T. Zusman, K. Schwartz, H. A. Shuman, T. Pupko, and G. Segal, "Computational modeling and experimental validation of the *Legionella* and *Coxiella* virulence-related type-IVB secretion signal,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, pp. E707–15, feb 2013.
- [104] J. P. Bardill, J. L. Miller, and J. P. Vogel, "IcmS-dependent translocation of SdeA into macrophages by the *Legionella pneumophila* type IV secretion system,” *Molecular Microbiology*, vol. 56, pp. 90–103, feb 2005.
- [105] E. M. Campodonico, L. Chesnel, and C. R. Roy, "A yeast genetic system for the identification and characterization of substrate proteins transferred into host cells by the *Legionella pneumophila* Dot/Icm system.,” *Molecular microbiology*, vol. 56, pp. 918–33, may 2005.
- [106] M. Heidtman, E. J. Chen, M. Y. Moy, and R. R. Isberg, "Large-scale identification of *Legionella pneumophila* Dot/Icm substrates that modulate host cell vesicle trafficking pathways,” *Cellular Microbiology*, vol. 11, no. 2, pp. 230–248, 2009.

- [107] D. Burstein, T. Zusman, E. Degtyar, R. Viner, G. Segal, and T. Pupko, "Genome-scale identification of Legionella pneumophila effectors using a machine learning approach.," *PLoS pathogens*, vol. 5, pp. e1000508–e1000508, jul 2009.
- [108] L. Gomez-Valero, C. Rusniok, M. Rolando, M. Neou, D. Dervins-Ravault, J. Demirtas, Z. Rouy, R. J. Moore, H. Chen, N. K. Petty, S. Jarraud, J. Etienne, M. Steinert, K. Heuner, S. Gribaldo, C. Médigue, G. Glöckner, E. L. Hartland, and C. Buchrieser, "Comparative analyses of Legionella species identifies genetic features of strains causing Legionnaires' disease.," *Genome biology*, vol. 15, p. 505, jan 2014.
- [109] K. M. O'Brien, E. L. Lindsay, and V. J. Starai, "The Legionella pneumophila Effector Protein, LegC7, Alters Yeast Endosomal Trafficking," *PLOS ONE*, vol. 10, pp. e0116824–e0116824, feb 2015.
- [110] M. P. Machner, A.-M. Krachler, M. L. Vasil, T. J. O'connor, and S. Ghosh, "Beyond Paralogs: The Multiple Layers of Redundancy in Bacterial Pathogenesis," *Frontiers in Cellular and Infection Microbiology* — *www.frontiersin.org*, vol. 7, p. 467, 2017.
- [111] H. Nagai, E. D. Cambronne, J. C. Kagan, J. C. Amor, R. A. Kahn, and C. R. Roy, "A C-terminal translocation signal required for Dot/Icm-dependent delivery of the Legionella RalF protein to host cells," *Proceedings of the National Academy of Sciences*, vol. 102, pp. 826–831, jan 2005.
- [112] J. Qiu and Z.-Q. Luo, "Effector translocation by the legionella dot/icm type iv secretion system," in *Molecular Mechanisms in Legionella Pathogenesis*, pp. 103–115, Springer, 2013.
- [113] L. Huang, D. Boyd, W. M. Amyot, A. D. Hempstead, Z.-Q. Luo, T. J. O'Connor, C. Chen, M. Machner, T. Montminy, and R. R. Isberg, "The E Block motif is associated with Legionella pneumophila translocated substrates.," *Cellular microbiology*, vol. 13, pp. 227–245, feb 2011.
- [114] D. Xu, A. Farmer, and Y. M. Chook, "Recognition of nuclear targeting signals by Karyopherin- $\beta$  proteins.," *Current opinion in structural biology*, vol. 20, pp. 782–790, dec 2010.
- [115] M. A. Horwitz and F. R. Maxfield, "Legionella pneumophila inhibits acidification of its phagosome in human monocytes," *The Journal of Cell Biology*, vol. 99, pp. 1936–1943, dec 1984.
- [116] R. Behnia and S. Munro, "Organelle identity and the signposts for membrane traffic," *Nature*, vol. 438, no. 7068, pp. 597–604, 2005.
- [117] M. A. D. Matteis and A. Godi, "PI-loting membrane traffic," *Nature Cell Biology*, vol. 6, pp. 487–492, jun 2004.
- [118] M. Molmeret, O. T. Alli, S. Zink, A. Fliieger, N. P. Cianciotto, and Y. A. Kwai, "icmt is essential for pore formation-mediated egress of legionella pneumophila from mammalian and protozoan cells," *Infection and immunity*, vol. 70, no. 1, pp. 69–78, 2002.
- [119] B. Steiner, S. Weber, and H. Hilbi, "Formation of the Legionella-containing vacuole: phosphoinositide conversion, GTPase modulation and ER dynamics," *International Journal of Medical Microbiology*, vol. 308, no. 1, pp. 49–57, 2018.

- [120] N. Dong, M. Niu, L. Hu, Q. Yao, R. Zhou, and F. Shao, "Modulation of membrane phosphoinositide dynamics by the phosphatidylinositol 4-kinase activity of the Legionella LepB effector.," *Nature microbiology*, vol. 2, p. 16236, dec 2016.
- [121] Q. Yu, L. Hu, Q. Yao, Y. Zhu, N. Dong, D.-C. Wang, and F. Shao, "Structural analyses of legionella lepb reveal a new gap fold that catalytically mimics eukaryotic rasgap," *Cell Research*, vol. 23, no. 6, pp. 775–787, 2013.
- [122] S. S. Weber, C. Ragaz, K. Reus, Y. Nyfeler, and H. Hilbi, "Legionella pneumophila exploits PI(4)P to anchor secreted effector proteins to the replicative vacuole.," *PLoS pathogens*, vol. 2, p. e46, may 2006.
- [123] L. Toulabi, X. Wu, Y. Cheng, and Y. Mao, "Identification and structural characterization of a Legionella phosphoinositide phosphatase.," *The Journal of biological chemistry*, vol. 288, pp. 24518–27, aug 2013.
- [124] A. V. F. Nielsen, I. Tetens, and A. S. Meyer, "Potential of Phytase-Mediated Iron Release from Cereal-Based Foods: A Quantitative View," *Nutrients*, vol. 5, no. 8, p. 3074, 2013.
- [125] S. Weber, B. Steiner, A. Welin, and H. Hilbi, "Legionella-Containing Vacuoles Capture PtdIns(4)P-Rich Vesicles Derived from the Golgi Apparatus.," *mBio*, vol. 9, pp. e02420–18, dec 2018.
- [126] K. G. Wooldridge and P. H. Williams, "Iron uptake mechanisms of pathogenic bacteria," *FEMS Microbiology Reviews*, vol. 12, pp. 325–348, nov 1993.
- [127] B. R. Wilson, A. R. Bogdan, M. Miyazawa, K. Hashimoto, and Y. Tsuji, "Siderophores in Iron Metabolism: From Mechanism to Therapy Potential.," *Trends in molecular medicine*, vol. 22, no. 12, pp. 1077–1090, 2016.
- [128] M. Folly-Klan, E. Alix, D. Stalder, P. Ray, L. V. Duarte, A. Delprato, M. Zeghouf, B. Antonny, V. Campanacci, C. R. Roy, and J. Cherfils, "A Novel Membrane Sensor Controls the Localization and ArfGEF Activity of Bacterial RalF," *PLoS Pathogens*, vol. 9, no. 11, 2013.
- [129] A. Godi, P. Pertile, R. Meyers, P. Marra, G. Di Tullio, C. Iurisci, A. Luini, D. Corda, and M. A. De Matteis, "ARF mediates recruitment of PtdIns-4-OH kinase-beta and stimulates synthesis of PtdIns(4,5)P2 on the Golgi complex.," *Nature cell biology*, vol. 1, pp. 280–7, sep 1999.
- [130] A. Skippen, D. H. Jones, C. P. Morgan, M. Li, and S. Cockcroft, "Mechanism of ARF-stimulated phosphatidylinositol 4,5-bisphosphate synthesis in HL60 cells," *J Biol Chem*, vol. 14, p. 14, 2001.
- [131] J. Li and V. W. Hsu, "An ACAP1 coat complex acting in endocytic recycling," *Methods in Cell Biology*, vol. 130, pp. 81–99, jan 2015.
- [132] A. Spang, "Arf1 regulatory factors and copi vesicle formation," *Current opinion in cell biology*, vol. 14, no. 4, pp. 423–427, 2002.
- [133] R. K. Draper, R. T. Hudson, and T. Hu, "Use of Aminoglycoside Antibiotics and Related Compounds to Study ADP-Ribosylation Factor (ARF) /Coatomeer Function in Golgi Traffic," *Methods in Enzymology*, vol. 329, pp. 372–379, jan 2001.

- [134] L. P. Haynes, M. W. Sherwood, N. J. Dolman, and R. D. Burgoyne, "Specificity, Promiscuity and Localization of ARF Protein Interactions with NCS-1 and Phosphatidylinositol-4 Kinase-III $\beta$ ," *Traffic*, vol. 8, pp. 1080–1092, aug 2007.
- [135] E. L. Clayton, S. Minogue, and M. G. Waugh, "Mammalian phosphatidylinositol 4-kinases as modulators of membrane trafficking and lipid signaling networks.," *Progress in lipid research*, vol. 52, pp. 294–304, jul 2013.
- [136] R. Viner, D. Chetrit, M. Ehrlich, and G. Segal, "Identification of Two Legionella pneumophila Effectors that Manipulate Host Phospholipids Biosynthesis," *PLoS Pathogens*, vol. 8, p. e1002988, nov 2012.
- [137] G. N. Schroeder, P. Aurass, C. V. Oates, E. W. Tate, E. L. Hartland, A. Flieger, and G. Frankel, "Legionella pneumophila Effector LpdA Is a Palmitoylated Phospholipase D Virulence Factor," *Infection and Immunity*, vol. 83, pp. 3989–4002, oct 2015.
- [138] M. Amadio, F. Battaini, and A. Pascale, "The different facets of protein kinases C: old and new players in neuronal signal transduction pathways," *Pharmacological Research*, vol. 54, pp. 317–325, nov 2006.
- [139] S. S. Weber, 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.," *Cellular microbiology*, vol. 11, pp. 442–60, mar 2009.
- [140] N. Hyvola, A. Diao, E. McKenzie, A. Skippen, S. Cockcroft, and M. Lowe, "Membrane targeting and activation of the Lowe syndrome protein OCRL1 by rab GTPases.," *The EMBO journal*, vol. 25, pp. 3750–3761, aug 2006.
- [141] U. Lichter-Konecki, L. Farber, J. Cronin, S. Suchy, and R. Nussbaum, "The effect of missense mutations in the rhogap-homology domain on ocr11 function," *Molecular genetics and metabolism*, vol. 89, no. 1-2, pp. 121–128, 2006.
- [142] A. Welin, S. Weber, and H. Hilbi, "Quantitative Imaging Flow Cytometry of Legionella-Infected Dictyostelium Amoebae Reveals the Impact of Retrograde Trafficking on Pathogen Vacuole Composition.," *Applied and environmental microbiology*, vol. 84, no. 11, 2018.
- [143] M. Lucas, A. H. Gaspar, C. Pallara, A. L. Rojas, J. Fernández-Recio, M. P. Machner, and A. Hierro, "Structural basis for the recruitment and activation of the Legionella phospholipase VipD by the host GTPase Rab5.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, pp. E3514–23, aug 2014.
- [144] A. H. Gaspar and M. P. Machner, "VipD is a Rab5-activated phospholipase A1 that protects Legionella pneumophila from endosomal fusion.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, pp. 4560–5, mar 2014.
- [145] S. M. VanRheenen, Z.-Q. Luo, T. O'Connor, and R. R. Isberg, "Members of a Legionella pneumophila family of proteins with ExoU (phospholipase A) active sites are translocated to target cells.," *Infection and immunity*, vol. 74, pp. 3597–606, jun 2006.



- [146] C. R. Harding, C. Mattheis, A. Mousnier, C. V. Oates, E. L. Hartland, G. Frankel, and G. N. Schroeder, "Ltpd is a novel legionella pneumophila effector that binds phosphatidylinositol 3-phosphate and inositol monophosphatase impa1," *Infection and immunity*, vol. 81, no. 11, pp. 4261–4270, 2013.
- [147] S. Toma-Fukai and T. Shimizu, "Structural insights into the regulation mechanism of small gtpases by gefs," *Molecules*, vol. 24, no. 18, p. 3308, 2019.
- [148] E. Brombacher, S. Urwyler, C. Ragaz, S. S. Weber, K. K. Kami, M. Overduin, and H. Hilbi, "Rab1 guanine nucleotide exchange factor SidM is a major phosphatidylinositol 4-phosphate-binding effector protein of Legionella pneumophila.," *The Journal of biological chemistry*, vol. 284, pp. 4846–56, feb 2009.
- [149] E. C. So, G. N. Schroeder, D. Carson, C. Mattheis, A. A. Mousnier, M. Broncel, E. W. Tate, and G. Frankel, "The Rab-binding profiles of bacterial virulence factors during infection," *Journal of Biological Chemistry*, vol. 291, p. jbc.M115.700930, mar 2016.
- [150] M. P. Müller, H. Peters, J. Blümer, W. Blankenfeldt, R. S. Goody, and A. Itzen, "The legionella effector protein drra ampylates the membrane traffic regulator rab1b," *Science*, vol. 329, no. 5994, pp. 946–949, 2010.
- [151] L. K. Oesterlin, R. S. Goody, and A. Itzen, "Posttranslational modifications of Rab proteins cause effective displacement of GDP dissociation inhibitor.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, pp. 5621–6, apr 2012.
- [152] C. A. Hardiman and C. R. Roy, "AMPylation is critical for Rab1 localization to vacuoles containing Legionella pneumophila.," *mBio*, vol. 5, pp. e01035–13, feb 2014.
- [153] G. M. Conover, I. Derré, J. P. Vogel, and R. R. Isberg, "The Legionella pneumophila LidA protein: a translocated substrate of the Dot/Icm system associated with maintenance of bacterial integrity," *Molecular Microbiology*, vol. 48, pp. 305–321, apr 2003.
- [154] W. Cheng, K. Yin, D. Lu, B. Li, D. Zhu, Y. Chen, H. Zhang, S. Xu, J. Chai, and L. Gu, "Structural Insights into a Unique Legionella pneumophila Effector LidA Recognizing Both GDP and GTP Bound Rab1 in Their Active State," *PLoS Pathogens*, vol. 8, p. e1002528, mar 2012.
- [155] M. R. Neunuebel, S. Mohammadi, M. Jarnik, and M. P. Machner, "Legionella pneumophila LidA Affects Nucleotide Binding and Activity of the Host GTPase Rab1," *Journal of Bacteriology*, vol. 194, pp. 1389–1400, mar 2012.
- [156] M. P. Machner and R. R. Isberg, "Targeting of host Rab GTPase function by the intravacuolar pathogen Legionella pneumophila.," *Developmental cell*, vol. 11, pp. 47–56, jul 2006.
- [157] A. H. Hutagalung and P. J. Novick, "Role of Rab GTPases in Membrane Traffic and Cell Physiology," *Physiological Reviews*, vol. 91, pp. 119–149, jan 2011.
- [158] S. Urwyler, Y. Nyfeler, C. Ragaz, H. Lee, L. N. Mueller, R. Aebersold, and H. Hilbi, "Proteome analysis of legionella vacuoles purified by magnetic immunoseparation reveals secretory and endosomal gtpases," *Traffic*, vol. 10, no. 1, pp. 76–87, 2009.

- [159] X. Hou, N. Hagemann, S. Schoebel, W. Blankenfeldt, R. S. Goody, K. S. Erdmann, and A. Itzen, “A structural basis for Lowe syndrome caused by mutations in the Rab-binding domain of OCRL1,” *The EMBO Journal*, vol. 30, pp. 1659–1670, apr 2011.
- [160] I. Derré and R. R. Isberg, “Lida, a translocated substrate of the legionella pneumophila type iv secretion system, interferes with the early secretory pathway,” *Infection and immunity*, vol. 73, no. 7, pp. 4370–4380, 2005.
- [161] K. Arasaki, D. K. Toomre, and C. R. Roy, “The Legionella pneumophila effector DrrA is sufficient to stimulate SNARE-dependent membrane fusion.,” *Cell host & microbe*, vol. 11, pp. 46–57, jan 2012.
- [162] S. Mukherjee, X. Liu, K. Arasaki, J. McDonough, J. E. Galán, and C. R. Roy, “Modulation of Rab GTPase function by a protein phosphocholine transferase.,” *Nature*, vol. 477, pp. 103–6, sep 2011.
- [163] P. R. Goody, K. Heller, L. K. Oesterlin, M. P. Müller, A. Itzen, and R. S. Goody, “Reversible phosphocholination of Rab proteins by Legionella pneumophila effector proteins,” *EMBO Journal*, vol. 31, pp. 1774–1784, apr 2012.
- [164] A. Ingmundson, A. Delprato, D. G. Lambright, and C. R. Roy, “Legionella pneumophila proteins that regulate Rab1 membrane cycling.,” *Nature*, vol. 450, pp. 365–9, nov 2007.
- [165] Y. Tan, R. J. Arnold, and Z.-Q. Luo, “Legionella pneumophila regulates the small GTPase Rab1 activity by reversible phosphorylcholation.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, pp. 21212–7, dec 2011.
- [166] E. Mihai Gazdag, A. Streller, I. Haneburger, H. Hilbi, I. R. Vetter, R. S. Goody, and A. Itzen, “Mechanism of Rab1b deactivation by the Legionella pneumophila GAP LepB,” *EMBO reports*, vol. 14, pp. 199–205, jan 2013.
- [167] M. R. Neunuebel, Y. Chen, A. H. Gaspar, P. S. Backlund, A. Yergey, and M. P. Machner, “De-AMPylation of the small GTPase Rab1 by the pathogen Legionella pneumophila.,” *Science (New York, N.Y.)*, vol. 333, pp. 453–6, jul 2011.
- [168] A. Hubber, K. Arasaki, F. Nakatsu, C. Hardiman, D. Lambright, P. De Camilli, H. Nagai, and C. R. Roy, “The machinery at endoplasmic reticulum-plasma membrane contact sites contributes to spatial regulation of multiple legionella effector proteins,” *PLoS Pathog*, vol. 10, no. 7, p. e1004222, 2014.
- [169] K. C. Jeong, J. A. Sexton, and J. P. Vogel, “Spatiotemporal regulation of a Legionella pneumophila T4SS substrate by the metaeffector SidJ.,” *PLoS pathogens*, vol. 11, p. e1004695, mar 2015.
- [170] N. Shohdy, J. A. Efe, S. D. Emr, and H. A. Shuman, “Pathogen effector protein screening in yeast identifies Legionella factors that interfere with membrane trafficking.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, pp. 4866–71, mar 2005.
- [171] Y.-S. Sohn, H.-C. Shin, W. S. Park, J. Ge, C.-H. Kim, B. L. Lee, W. Do Heo, J. U. Jung, D. J. Rigden, and B.-H. Oh, “Lpg0393 of Legionella pneumophila Is a Guanine-Nucleotide Exchange Factor for Rab5, Rab21 and Rab22,” *PLOS ONE*, vol. 10, p. e0118683, mar 2015.

- [172] Y. Belyi, T. Jank, and K. Aktories, "Effector Glycosyltransferases in Legionella," *Frontiers in Microbiology*, vol. 2, p. 76, apr 2011.
- [173] T. Jank, K. E. Böhmer, T. Tzivelekidis, C. Schwan, Y. Belyi, and K. Aktories, "Domain organization of Legionella effector SetA," *Cellular Microbiology*, vol. 14, pp. 852–868, jun 2012.
- [174] C.-W. Park and K.-Y. Ryu, "Cellular ubiquitin pool dynamics and homeostasis," *BMB reports*, vol. 47, pp. 475–482, sep 2014.
- [175] P. Ebner, G. A. Versteeg, and F. Ikeda, "Ubiquitin enzymes in the regulation of immune responses," *Critical reviews in biochemistry and molecular biology*, vol. 52, no. 4, pp. 425–460, 2017.
- [176] S. Gilberto and M. Peter, "Dynamic ubiquitin signaling in cell cycle regulation," *Journal of Cell Biology*, vol. 216, no. 8, pp. 2259–2271, 2017.
- [177] P. Schwertman, S. Bekker-Jensen, and N. Mailand, "Regulation of dna double-strand break repair by ubiquitin and ubiquitin-like modifiers," *Nature reviews Molecular cell biology*, vol. 17, no. 6, p. 379, 2016.
- [178] M. J. Clague and S. Urbé, "Integration of cellular ubiquitin and membrane traffic systems: focus on deubiquitylases," *The FEBS journal*, vol. 284, no. 12, pp. 1753–1766, 2017.
- [179] W. M. Bruckert, C. T. Price, and Y. Abu Kwaik, "Rapid nutritional remodeling of the host cell upon attachment of Legionella pneumophila," *Infection and immunity*, vol. 82, pp. 72–82, jan 2014.
- [180] W. M. Bruckert and Y. Abu Kwaik, "Lysine-11 Linked Polyubiquitination of the AnkB F-Box Effector of Legionella pneumophila," *Infection and Immunity*, vol. 84, pp. 99–107, jan 2016.
- [181] M. Lomma, D. Dervins-Ravault, M. Rolando, T. Nora, H. J. Newton, F. M. Sansom, T. Sahr, L. Gomez-Valero, M. Jules, E. L. Hartland, and C. Buchrieser, "The Legionella pneumophila F-box protein Lpp2082 (AnkB) modulates ubiquitination of the host protein parvin B and promotes intracellular replication.," *Cellular microbiology*, vol. 12, pp. 1272–91, sep 2010.
- [182] A. W. Ensminger and R. R. Isberg, "E3 ubiquitin ligase activity and targeting of BAT3 by multiple Legionella pneumophila translocated substrates.," *Infection and immunity*, vol. 78, pp. 3905–19, sep 2010.
- [183] K. Wong, J. D. Perpich, G. Kozlov, M. Cygler, Y. Abu Kwaik, and K. Gehring, "Structural Mimicry by a Bacterial F Box Effector Hijacks the Host Ubiquitin-Proteasome System," *Structure*, vol. 25, pp. 376–383, feb 2017.
- [184] T. Kubori, X. T. Bui, A. Hubber, and H. Nagai, "Legionella RavZ Plays a Role in Preventing Ubiquitin Recruitment to Bacteria-Containing Vacuoles," *Frontiers in Cellular and Infection Microbiology*, vol. 7, p. 384, aug 2017.
- [185] S. Bhogaraju, S. Kalayil, Y. Liu, F. Bonn, T. Colby, I. Matic, and I. Dikic, "Phosphoribosylation of Ubiquitin Promotes Serine Ubiquitination and Impairs Conventional Ubiquitination.," *Cell*, vol. 167, pp. 1636–1649.e13, dec 2016.

- [186] A. Akturk, D. J. Wasilko, X. Wu, Y. Liu, Y. Zhang, J. Qiu, Z.-Q. Luo, K. H. Reiter, P. S. Brzovic, R. E. Klevit, *et al.*, “Mechanism of phosphoribosyl-ubiquitination mediated by a single legionella effector,” *Nature*, vol. 557, no. 7707, pp. 729–733, 2018.
- [187] L. Kim, D. H. Kwon, B. H. Kim, J. Kim, M. R. Park, Z.-Y. Park, and H. K. Song, “Structural and biochemical study of the mono-*adp*-ribosyltransferase domain of *sdea*, a ubiquitylating/deubiquitylating enzyme from legionella pneumophila,” *Journal of molecular biology*, vol. 430, no. 17, pp. 2843–2856, 2018.
- [188] F. A. Horenkamp, S. Mukherjee, E. Alix, C. M. Schauder, A. M. Hubber, C. R. Roy, and K. M. Reinisch, “Legionella pneumophila subversion of host vesicular transport by SidC effector proteins.,” *Traffic (Copenhagen, Denmark)*, vol. 15, pp. 488–99, may 2014.
- [189] C. Ragaz, H. Pietsch, S. Urwyler, A. Tiaden, S. S. Weber, and H. Hilbi, “The Legionella pneumophila phosphatidylinositol-4 phosphate-binding type IV substrate SidC recruits endoplasmic reticulum vesicles to a replication-permissive vacuole.,” *Cellular microbiology*, vol. 10, pp. 2416–33, dec 2008.
- [190] F. Hsu, X. Luo, J. Qiu, Y.-B. Teng, J. Jin, M. B. Smolka, Z.-Q. Luo, and Y. Mao, “The Legionella effector SidC defines a unique family of ubiquitin ligases important for bacterial phagosomal remodeling.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, pp. 10538–43, jul 2014.
- [191] J. Qiu and Z.-Q. Luo, “Hijacking of the Host Ubiquitin Network by Legionella pneumophila,” *Frontiers in Cellular and Infection Microbiology*, vol. 7, p. 487, dec 2017.
- [192] M. S. Prevost, N. Pinotsis, M. Dumoux, R. D. Hayward, and G. Waksman, “The Legionella effector WipB is a translocated Ser/Thr phosphatase that targets the host lysosomal nutrient sensing machinery,” *Scientific Reports*, vol. 7, p. 9450, dec 2017.
- [193] E. P. Skaar, “The battle for iron between bacterial pathogens and their vertebrate hosts,” *PLoS Pathog*, vol. 6, no. 8, p. e1000949, 2010.
- [194] D. T. Isaac, R. K. Laguna, N. Valtz, and R. R. Isberg, “Mavn is a legionella pneumophila vacuole-associated protein required for efficient iron acquisition during intracellular growth,” *Proceedings of the National Academy of Sciences*, vol. 112, no. 37, pp. E5208–E5217, 2015.
- [195] E. T. Christenson, D. T. Isaac, K. Yoshida, E. Lipo, J.-S. Kim, R. Ghirlando, R. R. Isberg, and A. Banerjee, “The iron-regulated vacuolar Legionella pneumophila MavN protein is a transition-metal transporter.,” *Proceedings of the National Academy of Sciences of the United States of America*, p. 201902806, aug 2019.
- [196] P. Dolezal, M. Aili, J. Tong, J.-H. Jiang, C. M. Marobbio, S. fung Lee, R. Schuelein, S. Belluzzo, E. Binova, A. Mousnier, G. Frankel, G. Giannuzzi, F. Palmieri, K. Gabriel, T. Naderer, E. L. Hartland, and T. Lithgow, “Legionella pneumophila Secretes a Mitochondrial Carrier Protein during Infection,” *PLoS Pathogens*, vol. 8, pp. e1002459–e1002459, jan 2012.
- [197] R. Dominguez and K. C. Holmes, “Actin structure and function,” *Annual Review of Biophysics*, 2011.

- [198] I. S. Franco, N. Shohdy, and H. A. Shuman, "The Legionella pneumophila effector VipA is an actin nucleator that alters host cell organelle trafficking.," *PLoS pathogens*, vol. 8, p. e1002546, feb 2012.
- [199] J. N. Bugalhão, L. J. Mota, and I. S. Franco, "Identification of regions within the Legionella pneumophila VipA effector protein involved in actin binding and polymerization and in interference with eukaryotic organelle trafficking.," *MicrobiologyOpen*, vol. 5, pp. 118–133, feb 2016.
- [200] Z. Guo, R. Stephenson, J. Qiu, S. Zheng, and Z.-Q. Luo, "A Legionella effector modulates host cytoskeletal structure by inhibiting actin polymerization.," *Microbes and infection*, vol. 16, pp. 225–236, mar 2014.
- [201] C. Michard, D. Sperandio, N. Baïlo, J. Pizarro-Cerdá, L. LeClaire, E. Chadeau-Argaud, I. Pombo-Grégoire, E. Hervet, A. Vianney, C. Gilbert, M. Faure, P. Cossart, and P. Doublet, "The Legionella Kinase LegK2 Targets the ARP2/3 Complex To Inhibit Actin Nucleation on Phagosomes and Allow Bacterial Evasion of the Late Endocytic Pathway.," *mBio*, vol. 6, pp. e00354–15, jan 2015.
- [202] N. Pinotsis and G. Waksman, "Structure of the WipA protein reveals a novel tyrosine protein phosphatase effector from Legionella pneumophila.," *The Journal of biological chemistry*, vol. 292, pp. 9240–9251, jun 2017.
- [203] L. He, Y. Lin, Z.-h. Ge, S.-y. He, B.-b. Zhao, D. Shen, J.-g. He, and Y.-j. Lu, "The legionella pneumophila effector wipa disrupts host f-actin polymerisation by hijacking phosphotyrosine signalling," *Cellular microbiology*, vol. 21, no. 6, p. e13014, 2019.
- [204] Y. Liu, W. Zhu, Y. Tan, E. S. Nakayasu, C. J. Staiger, and Z.-Q. Luo, "A Legionella Effector Disrupts Host Cytoskeletal Structure by Cleaving Actin.," *PLoS pathogens*, vol. 13, no. 1, pp. e1006186–e1006186, 2017.
- [205] G. J. Brouhard and L. M. Rice, "Microtubule dynamics: An interplay of biochemistry and mechanics," *Nature Reviews Molecular Cell Biology*, vol. 19, no. 7, pp. 451–463, 2018.
- [206] M. S. Moore, "Ran and nuclear transport.," *The Journal of biological chemistry*, vol. 273, pp. 22857–22860, sep 1998.
- [207] R. E. Carazo-Salas, O. J. Gruss, I. W. Mattaj, and E. Karsenti, "Ran-GTP coordinates regulation of microtubule nucleation and dynamics during mitotic-spindle assembly," *Nature Cell Biology*, vol. 3, pp. 228–234, mar 2001.
- [208] E. Rothmeier, G. Pfaffinger, C. Hoffmann, C. F. Harrison, H. Grabmayr, U. Repnik, M. Hannemann, S. Wölke, A. Bausch, G. Griffiths, *et al.*, "Activation of ran gtpase by a legionella effector promotes microtubule polymerization, pathogen vacuole motility and infection," *PLoS Pathog*, vol. 9, no. 9, p. e1003598, 2013.
- [209] S. S. Ivanov, G. Charron, H. C. Hang, and C. R. Roy, "Lipidation by the host prenyltransferase machinery facilitates membrane localization of Legionella pneumophila effector proteins," *Journal of Biological Chemistry*, vol. 285, no. 45, pp. 34686–34698, 2010.
- [210] H. Hilbi, E. Rothmeier, C. Hoffmann, and C. F. Harrison, "Beyond Rab GTPases Legionella activates the small GTPase Ran to promote microtubule polymerization, pathogen vacuole motility, and infection.," *Small GTPases*, vol. 5, no. 3, pp. 1–6, 2014.

- [211] J.-J. Liu, “Retromer-mediated protein sorting and vesicular trafficking,” *Journal of Genetics and Genomics*, vol. 43, no. 4, pp. 165–177, 2016.
- [212] K. Bärlocher, C. A. J. Hutter, A. L. Swart, B. Steiner, A. Welin, M. Hohl, F. Letourneur, M. A. Seeger, and H. Hilbi, “Structural insights into Legionella RidL-Vps29 retromer subunit interaction reveal displacement of the regulator TBC1D5,” *Nature Communications*, vol. 8, p. 1543, dec 2017.
- [213] M. Lucas, D. C. Gershlick, A. Vidaurrazaga, A. L. Rojas, J. S. Bonifacino, and A. Hierro, “Structural mechanism for cargo recognition by the retromer complex,” *Cell*, vol. 167, no. 6, pp. 1623–1635, 2016.
- [214] J. Yao, F. Yang, X. Sun, S. Wang, N. Gan, Q. Liu, D. Liu, X. Zhang, D. Niu, Y. Wei, C. Ma, Z.-Q. Luo, Q. Sun, and D. Jia, “Mechanism of inhibition of retromer transport by the bacterial effector RidL.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 115, pp. E1446–E1454, feb 2018.
- [215] N. Levanova, C. Mattheis, D. Carson, K.-N. To, T. Jank, G. Frankel, K. Aktories, and G. N. Schroeder, “The Legionella effector LtpM is a new type of phosphoinositide-activated glucosyltransferase,” *The Journal of biological chemistry*, vol. 294, no. 8, pp. 2862–2879, 2019.
- [216] T. L. Bennett, S. M. Kraft, B. J. Reaves, J. Mima, K. M. O’Brien, and V. J. Starai, “LegC3, an Effector Protein from Legionella pneumophila, Inhibits Homotypic Yeast Vacuole Fusion In Vivo and In Vitro,” *PLoS ONE*, vol. 8, pp. e56798–e56798, feb 2013.
- [217] N. P. King, P. Newton, R. Schuelein, D. L. Brown, M. Petru, V. Zarsky, P. Dolezal, L. Luo, A. Bugarcic, A. C. Stanley, R. Z. Murray, B. M. Collins, R. D. Teasdale, E. L. Hartland, and J. L. Stow, “Soluble NSF attachment protein receptor molecular mimicry by a Legionella pneumophilaDot/Icm effector,” *Cellular Microbiology*, vol. 17, pp. 767–784, jun 2015.
- [218] C. M. Pike, R. Boyer-Andersen, L. N. Kinch, J. L. Caplan, and M. R. Neunuebel, “The Legionella effector RavD binds phosphatidylinositol-3-phosphate and helps suppress endolysosomal maturation of the Legionella-containing vacuole,” *Journal of Biological Chemistry*, vol. 294, no. 16, pp. 6405–6415, 2019.
- [219] E. Haenssler, V. Ramabhadran, C. S. Murphy, M. I. Heidtman, and R. R. Isberg, “Endoplasmic reticulum tubule protein reticulon 4 associates with the Legionella pneumophila vacuole and with translocated substrate Ceg9,” *Infection and Immunity*, vol. 83, no. 9, pp. 3479–3489, 2015.
- [220] B. H. Young, T. A. Caldwell, A. M. McKenzie, O. Kokhan, and C. E. Berndsen, “Characterization of the structure and catalytic activity of Legionella pneumophila VipF,” *Proteins: Structure, Function, and Bioinformatics*, vol. 84, pp. 1422–1430, oct 2016.
- [221] F. Dyda, D. C. Klein, and A. B. Hickman, “GCN5-Related N-Acetyltransferases: A Structural Overview,” *Annual Review of Biophysics and Biomolecular Structure*, vol. 29, pp. 81–103, jun 2000.
- [222] M. d. S. Siqueira, R. d. M. Ribeiro, and L. H. Travassos, “Autophagy and its interaction with intracellular bacterial pathogens,” *Frontiers in immunology*, vol. 9, p. 935, 2018.

- [223] A. Choy, J. Dancourt, B. Mugo, T. J. O'Connor, R. R. Isberg, T. J. Melia, and C. R. Roy, "The legionella effector ravz inhibits host autophagy through irreversible atg8 deconjugation," *Science*, vol. 338, no. 6110, pp. 1072–1076, 2012.
- [224] F. A. Horenkamp, K. J. Kauffman, L. J. Kohler, R. K. Sherwood, K. P. Krueger, V. Shteyn, C. R. Roy, T. J. Melia, and K. M. Reinisch, "The Legionella Anti-autophagy Effector RavZ Targets the Autophagosome via PI3P- and Curvature-Sensing Motifs," *Developmental Cell*, vol. 34, pp. 569–576, sep 2015.
- [225] A. Yang, S. Pantoom, and Y.-W. Wu, "Elucidation of the anti-autophagy mechanism of the Legionella effector RavZ using semisynthetic LC3 proteins.," *eLife*, vol. 6, apr 2017.
- [226] S. Pantoom, A. Yang, and Y.-W. Wu, "Lift and cut: Anti-host autophagy mechanism of legionella pneumophila," *Autophagy*, vol. 13, no. 8, pp. 1467–1469, 2017.
- [227] K. Arasaki, Y. Mikami, S. R. Shames, H. Inoue, Y. Wakana, and M. Tagaya, "Legionella effector Lpg1137 shuts down ER-mitochondria communication through cleavage of syntaxin 17," *Nature Communications*, vol. 8, p. 15406, aug 2017.
- [228] M. Gradowski and K. Pawlowski, "The Legionella pneumophila effector Lpg1137 is a homologue of mitochondrial SLC25 carrier proteins, not of known serine proteases," *PeerJ*, vol. 5, pp. e3849–e3849, sep 2017.
- [229] M. Rolando, P. Escoll, T. Nora, J. Botti, V. Boitez, C. Bedia, C. Daniels, G. Abraham, P. J. Stogios, T. Skarina, C. Christophe, D. Dervins-Ravault, C. Cazalet, H. Hilbi, T. W. T. Rupasinghe, D. Tull, M. J. McConville, S. Y. Ong, E. L. Hartland, P. Codogno, T. Levade, T. Naderer, A. Savchenko, and C. Buchrieser, "Legionella pneumophila S1P-lyase targets host sphingolipid metabolism and restrains autophagy.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 113, pp. 1901–6, feb 2016.
- [230] W. Zhu, L. A. Hammad, F. Hsu, Y. Mao, and Z. Q. Luo, "Induction of caspase 3 activation by multiple Legionella pneumophilaDot/Icm substrates," *Cellular Microbiology*, vol. 15, pp. 1783–1795, nov 2013.
- [231] S. Treacy-Abarca and S. Mukherjee, "Legionella suppresses the host unfolded protein response via multiple mechanisms," *Nature Communications*, vol. 6, p. 7887, jul 2015.
- [232] A. O. Amer, "Modulation of caspases and their non-apoptotic functions by Legionella pneumophila," *Cellular Microbiology*, vol. 12, pp. 140–147, feb 2010.
- [233] D. Bertheloot, E. Latz, and B. S. Franklin, "Necroptosis, pyroptosis and apoptosis: an intricate game of cell death," *Cellular & Molecular Immunology*, vol. 18, no. 5, pp. 1106–1121, 2021.
- [234] M. Y. Park, S. E. Ha, P. Vetrivel, H. H. Kim, P. B. Bhosale, A. Abusaliya, and G. S. Kim, "Differences of key proteins between apoptosis and necroptosis," *BioMed Research International*, vol. 2021, 2021.
- [235] J. P. Santavanond, S. F. Rutter, G. K. Atkin-Smith, and I. K. Poon, "Apoptotic bodies: Mechanism of formation, isolation and functional relevance," *SUBCELLULAR BIOCHEMISTRY (.) : Plenum Publishers*, vol. 97, pp. 61–88, 2021.

- [236] A. Abu-Zant, M. Santic, M. Molmeret, S. Jones, J. Helbig, and Y. Abu Kwaik, "Incomplete Activation of Macrophage Apoptosis during Intracellular Replication of *Legionella pneumophila*," *Infection and Immunity*, vol. 73, pp. 5339–5349, sep 2005.
- [237] S. Banga, P. Gao, X. Shen, V. Fiscus, W.-X. Zong, L. Chen, and Z.-Q. Luo, "Legionella pneumophila inhibits macrophage apoptosis by targeting pro-death members of the bcl2 protein family," *Proceedings of the National Academy of Sciences*, vol. 104, no. 12, pp. 5121–5126, 2007.
- [238] F. Hsu, W. Zhu, L. Brennan, L. Tao, Z.-Q. Luo, and Y. Mao, "Structural basis for substrate recognition by a unique legionella phosphoinositide phosphatase," *Proceedings of the National Academy of Sciences*, vol. 109, no. 34, pp. 13567–13572, 2012.
- [239] Z.-Q. Luo, "Striking a balance: modulation of host cell death pathways by legionella pneumophila," *Frontiers in microbiology*, vol. 2, p. 36, 2011.
- [240] V. P. Losick, E. Haenssler, M. Y. Moy, and R. R. Isberg, "LnaB: a *Legionella pneumophila* activator of NF- $\gamma$ B," *Cellular Microbiology*, vol. 12, pp. 1083–1097, aug 2010.
- [241] J. Ge, H. Xu, T. Li, Y. Zhou, Z. Zhang, S. Li, L. Liu, and F. Shao, "A *Legionella* type IV effector activates the NF-kappaB pathway by phosphorylating the I-kappaB family of inhibitors.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, pp. 13725–30, aug 2009.
- [242] M. F. Fontana, S. Banga, K. C. Barry, X. Shen, Y. Tan, Z.-Q. Luo, and R. E. Vance, "Secreted bacterial effectors that inhibit host protein synthesis are critical for induction of the innate immune response to virulent legionella pneumophila," *PLoS Pathog*, vol. 7, no. 2, p. e1001289, 2011.
- [243] X. Shen, S. Banga, Y. Liu, L. Xu, P. Gao, I. Shamovsky, E. Nudler, and Z.-Q. Luo, "Targeting eef1a by a legionella pneumophila effector leads to inhibition of protein synthesis and induction of host stress response," *Cellular microbiology*, vol. 11, no. 6, pp. 911–926, 2009.
- [244] Y. Belyi, R. Niggeweg, B. Opitz, M. Vogelsgesang, S. Hippenstiel, M. Wilm, and K. Aktories, "Legionella pneumophila glucosyltransferase inhibits host elongation factor 1A.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, pp. 16953–8, nov 2006.
- [245] T. Li, Q. Lu, G. Wang, H. Xu, H. Huang, T. Cai, B. Kan, J. Ge, and F. Shao, "SET-domain bacterial effectors target heterochromatin protein 1 to activate host rDNA transcription.," *EMBO reports*, vol. 14, pp. 733–40, aug 2013.
- [246] M. Rolando, S. Sanulli, C. Rusniok, L. Gomez-Valero, C. Bertholet, T. Sahr, R. Margueron, and C. Buchrieser, "Legionella pneumophila effector RomA uniquely modifies host chromatin to repress gene expression and promote intracellular bacterial replication.," *Cell host & microbe*, vol. 13, pp. 395–405, apr 2013.
- [247] W. Choi, E. A. Creasey, M. P. Lowe, and R. R. Isberg, "SdhA blocks disruption of the legionella-containing vacuole by hijacking the ocr1 phosphatase," *bioRxiv*, 2020.



- [248] X. Qu, X. Song, N. Zhang, J. Ma, and H. Ge, “The phospholipase a effector plaa from legionella pneumophila: expression, purification and crystallization,” *Acta Crystallographica Section F: Structural Biology Communications*, vol. 76, no. 3, pp. 138–144, 2020.
- [249] J. Ge, Y.-N. Gong, Y. Xu, and F. Shao, “Preventing bacterial dna release and absent in melanoma 2 inflammasome activation by a legionella effector functioning in membrane trafficking,” *Proceedings of the National Academy of Sciences*, vol. 109, no. 16, pp. 6193–6198, 2012.
- [250] M. Molmeret, S. Jones, M. Santic, F. Habyarimana, M. T. Garcia Esteban, and Y. Abu Kwaik, “Temporal and spatial trigger of post-exponential virulence-associated regulatory cascades by legionella pneumophila after bacterial escape into the host cell cytosol,” *Environmental microbiology*, vol. 12, no. 3, pp. 704–715, 2010.
- [251] M. Molmeret, S. D. Zink, L. Han, A. Abu-Zant, R. Asari, D. M. Bitar, and Y. Abu Kwaik, “Activation of caspase-3 by the dot/icm virulence system is essential for arrested biogenesis of the legionella-containing phagosome,” *Cellular microbiology*, vol. 6, no. 1, pp. 33–48, 2004.
- [252] A. L. Swart and H. Hilbi, “Phosphoinositides and the fate of legionella in phagocytes,” *Frontiers in immunology*, vol. 11, p. 25, 2020.
- [253] D. A. de Jesús-Díaz, C. Murphy, A. Sol, M. Dorer, and R. R. Isberg, “Host cell s phase restricts legionella pneumophila intracellular replication by destabilizing the membrane-bound replication compartment,” *MBio*, vol. 8, no. 4, 2017.
- [254] A. Sol, E. Lipo, D. A. de Jesús-Díaz, C. Murphy, M. Devereux, and R. R. Isberg, “Legionella pneumophila translocated translation inhibitors are required for bacterial-induced host cell cycle arrest.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 116, no. 8, pp. 3221–3228, 2019.
- [255] O. T. Alli, L.-Y. Gao, L. L. Pedersen, S. Zink, M. Radulic, M. Doric, and Y. Abu Kwaik, “Temporal pore formation-mediated egress from macrophages and alveolar epithelial cells by legionella pneumophila,” *Infection and immunity*, vol. 68, no. 11, pp. 6431–6440, 2000.
- [256] C. Hoffmann, I. Finsel, A. Otto, G. Pfaffinger, E. Rothmeier, M. Hecker, D. Becher, and H. Hilbi, “Functional analysis of novel Rab GTPases identified in the proteome of purified Legionella-containing vacuoles from macrophages,” *Cellular Microbiology*, vol. 16, pp. 1034–1052, feb 2014.
- [257] J. Chen, K. S. de Felipe, M. Clarke, H. Lu, O. R. Anderson, G. Segal, and H. A. Shuman, “Legionella effectors that promote nonlytic release from protozoa.,” *Science (New York, N.Y.)*, vol. 303, pp. 1358–1361, feb 2004.
- [258] M. K. Schuhmacher, M. Rolando, A. Bröhm, S. Weirich, S. Kudithipudi, C. Buchrieser, and A. Jeltsch, “The legionella pneumophila methyltransferase roma methylates also non-histone proteins during infection,” *Journal of molecular biology*, vol. 430, no. 13, pp. 1912–1925, 2018.
- [259] J. R. P. Knight, A. E. Willis, and J. Milner, “Active regulator of SIRT1 is required for ribosome biogenesis and function,” *Nucleic Acids Research*, vol. 41, pp. 4185–4197, apr 2013.

- [260] J. Son, C. H. Jo, R. N. Murugan, J. K. Bang, K. Y. Hwang, and W. C. Lee, "Crystal structure of legionella pneumophila type iv secretion system effector legas4," *Biochemical and Biophysical Research Communications*, vol. 465, no. 4, pp. 817–824, 2015.
- [261] C. T. D. Price and Y. A. Kwaik, "One bacterial effector with two distinct catalytic activities by different strains," *EMBO reports*, vol. 14, pp. 753–754, sep 2013.
- [262] P.-C. Lee, K. Beyrakhova, C. Xu, M. T. Boniecki, M. H. Lee, C. J. Onu, A. M. Grishin, M. P. Machner, and M. Cygler, "The Legionella kinase LegK7 exploits the Hippo pathway scaffold protein MOB1A for allostery and substrate phosphorylation," *Proceedings of the National Academy of Sciences of the United States of America*, jun 2020.
- [263] R. Schuelein, H. Spencer, L. F. Dagley, P. fei Li, L. Luo, J. L. Stow, G. Abraham, T. Naderer, L. Gomez-Valero, C. Buchrieser, C. Sugimoto, J. Yamagishi, A. I. Webb, S. Pasricha, and E. L. Hartland, "Targeting of RNA Polymerase II by a nuclear *Legionella pneumophila* Dot/Icm effector SnpL," *Cellular Microbiology*, vol. 20, p. e12852, sep 2018.
- [264] Y. Belyi, I. Tabakova, M. Stahl, and K. Aktories, "Lgt: a family of cytotoxic glucosyltransferases produced by legionella pneumophila," *Journal of bacteriology*, vol. 190, no. 8, pp. 3026–3035, 2008.
- [265] R. Hurtado-Guerrero, T. Zusman, S. Pathak, A. F. M. Ibrahim, S. Shepherd, A. Prescott, G. Segal, and D. M. F. Van Aalten, "Molecular mechanism of elongation factor 1A inhibition by a Legionella pneumophila glycosyltransferase," *Biochemical Journal*, vol. 426, pp. 281–292, mar 2010.
- [266] M. F. Fontana, S. Shin, and R. E. Vance, "Activation of host mitogen-activated protein kinases by secreted Legionella pneumophila effectors that inhibit host protein translation.," *Infection and immunity*, vol. 80, pp. 3570–5, oct 2012.
- [267] A. Flayhan, C. Bergé, N. Baïlo, P. Doublet, R. Bayliss, and L. Terradot, "The structure of Legionella pneumophila LegK4 type four secretion system (T4SS) effector reveals a novel dimeric eukaryotic-like kinase.," *Scientific reports*, vol. 5, p. 14602, jan 2015.
- [268] E. Hervet, X. Charpentier, A. Vianney, J.-C. Lazzaroni, C. Gilbert, D. Atlan, and P. Doublet, "Protein kinase LegK2 is a type IV secretion system effector involved in endoplasmic reticulum recruitment and intracellular replication of Legionella pneumophila.," *Infection and immunity*, vol. 79, pp. 1936–1950, may 2011.
- [269] S. M. Moss, I. R. Taylor, D. Ruggero, J. E. Gestwicki, K. M. Shokat, and S. Mukherjee, "A Legionella pneumophila Kinase Phosphorylates the Hsp70 Chaperone Family to Inhibit Eukaryotic Protein Synthesis," *Cell Host and Microbe*, vol. 25, no. 3, pp. 454–462.e6, 2019.
- [270] M. L. Urbanus, A. T. Quaile, P. J. Stogios, M. Morar, C. Rao, R. Di Leo, E. Evdokimova, M. Lam, C. Oatway, M. E. Cuff, J. Osipiuk, K. Michalska, B. P. Nocek, M. Taipale, A. Savchenko, and A. W. Ensminger, "Diverse mechanisms of metaeffector activity in an intracellular bacterial pathogen, Legionella pneumophila ," *Molecular Systems Biology*, vol. 12, p. 893, dec 2016.

- [271] A. M. Joseph and S. R. Shames, “Affecting the effectors: Regulation of legionella pneumophila effector function by metaeffectors,” *Pathogens*, vol. 10, no. 2, p. 108, 2021.
- [272] A. T. Quaile, M. L. Urbanus, P. J. Stogios, B. Nocek, T. Skarina, A. W. Ensminger, and A. Savchenko, “Molecular Characterization of LubX: Functional Divergence of the U-Box Fold by Legionella pneumophila.,” *Structure (London, England : 1993)*, vol. 23, pp. 1459–1469, aug 2015.
- [273] T. Kubori, N. Shinzawa, H. Kanuka, and H. Nagai, “Legionella metaeffector exploits host proteasome to temporally regulate cognate effector,” *PLoS Pathogens*, vol. 6, p. e1001216, dec 2010.
- [274] T. Kubori, A. Hyakutake, and H. Nagai, “Legionella translocates an e3 ubiquitin ligase that has multiple u-boxes with distinct functions,” *Molecular microbiology*, vol. 67, no. 6, pp. 1307–1319, 2008.
- [275] Y.-H. Lin, A. G. Doms, E. Cheng, B. Kim, T. R. Evans, and M. P. Machner, “Host Cell-catalyzed S-Palmitoylation Mediates Golgi Targeting of the Legionella Ubiquitin Ligase GobX.,” *The Journal of biological chemistry*, vol. 290, pp. 25766–81, oct 2015.
- [276] Y. Liu and Z.-Q. Luo, “The Legionella pneumophila effector SidJ is required for efficient recruitment of endoplasmic reticulum proteins to the bacterial phagosome.,” *Infection and immunity*, vol. 75, pp. 592–603, feb 2007.
- [277] A. M. Joseph, A. E. Pohl, T. J. Ball, T. G. Abram, D. K. Johnson, B. V. Geisbrecht, and S. R. Shames, “The Legionella pneumophila metaeffector Lpg2505 (SusF) regulates SidI-mediated translation inhibition and GDP-dependent glycosyltransferase activity,” *bioRxiv*, p. 845313, nov 2019.
- [278] A. M. Joseph, A. E. Pohl, T. J. Ball, T. G. Abram, D. K. Johnson, B. V. Geisbrecht, and S. R. Shames, “The Legionella pneumophila Metaeffector Lpg2505 (MesI) Regulates SidI-Mediated Translation Inhibition and Novel Glycosyl Hydrolase Activity ,” *Infection and Immunity*, vol. 88, mar 2020.
- [279] D. A. Machtens, J. M. Willerding, S. Eschenburg, and T. F. Reubold, “Crystal structure of the metaeffector MesI (Lpg2505) from Legionella pneumophila,” *Biochemical and Biophysical Research Communications*, may 2020.
- [280] D. S. Schwarz and M. D. Blower, “The endoplasmic reticulum: structure, function and response to cellular signaling,” *Cellular and Molecular Life Sciences*, vol. 73, no. 1, pp. 79–94, 2016.
- [281] K. Araki and K. Nagata, “Protein folding and quality control in the er,” *Cold Spring Harbor perspectives in biology*, vol. 3, no. 11, p. a007526, 2011.
- [282] X. Wu and T. A. Rapoport, “Mechanistic insights into ER-associated protein degradation,” aug 2018.
- [283] S. A. Houck, H. Y. Ren, V. J. Madden, J. N. Bonner, M. P. Conlin, J. A. Janovick, P. M. Conn, and D. M. Cyr, “Quality control autophagy degrades soluble erad-resistant conformers of the misfolded membrane protein gnhr,” *Molecular cell*, vol. 54, no. 1, pp. 166–179, 2014.
- [284] J. Hwang and L. Qi, “Quality Control in the Endoplasmic Reticulum: Crosstalk between ERAD and UPR pathways,” *Trends in Biochemical Sciences*, vol. 43, pp. 593–605, aug 2018.

- [285] G. Thibault and D. T. Ng, "The endoplasmic reticulum-associated degradation pathways of budding yeast," *Cold Spring Harbor perspectives in biology*, vol. 4, no. 12, p. a013193, 2012.
- [286] A. Raffaello, C. Mammucari, G. Gherardi, and R. Rizzuto, "Calcium at the center of cell signaling: interplay between endoplasmic reticulum, mitochondria, and lysosomes," *Trends in biochemical sciences*, vol. 41, no. 12, pp. 1035–1049, 2016.
- [287] D. Whitley, S. P. Goldberg, and W. D. Jordan, "Heat shock proteins: a review of the molecular chaperones," *Journal of vascular surgery*, vol. 29, no. 4, pp. 748–751, 1999.
- [288] A. Dubey, K. Prajapati, M. Swamy, and V. Pachauri, "Heat shock proteins: a therapeutic target worth to consider," *Veterinary world*, vol. 8, no. 1, p. 46, 2015.
- [289] W. R. Boorstein, T. Ziegelhoffer, and E. A. Craig, "Molecular evolution of the hsp70 multigene family," *Journal of molecular evolution*, vol. 38, no. 1, pp. 1–17, 1994.
- [290] S. Krenek, M. Schlegel, and T. U. Berendonk, "Convergent evolution of heat-inducibility during subfunctionalization of the hsp70 gene family," *BMC evolutionary biology*, vol. 13, no. 1, p. 49, 2013.
- [291] M. E. Murphy, "The HSP70 family and cancer," *Carcinogenesis*, vol. 34, pp. 1181–1188, jun 2013.
- [292] A. A. Asea and P. Kaur, *HSP70 in human diseases and disorders*, vol. 14. Springer, 2018.
- [293] F. Stricher, C. Macri, M. Ruff, and S. Muller, "HSPA8/HSC70 chaperone protein: Structure, function, and chemical targeting," dec 2013.
- [294] S. Akakura, M. Yoshida, Y. Yoneda, and S. Horinouchi, "A role for Hsc70 in regulating nucleocytoplasmic transport of a temperature-sensitive p53 (p53Val-135).," *The Journal of biological chemistry*, vol. 276, pp. 14649–14657, may 2001.
- [295] M. R. Fernández-Fernández, M. Gragera, L. Ochoa-Ibarrola, L. Quintana-Gallardo, and J. M. Valpuesta, "Hsp70—a master regulator in protein degradation," *FEBS letters*, vol. 591, no. 17, pp. 2648–2660, 2017.
- [296] R. Sousa and E. M. Lafer, "The role of molecular chaperones in clathrin mediated vesicular trafficking," *Frontiers in molecular biosciences*, vol. 2, p. 26, 2015.
- [297] S. Gupta, A. Deepti, S. Deegan, F. Lisbona, C. Hetz, and A. Samali, "HSP72 Protects Cells from ER Stress-induced Apoptosis via Enhancement of IRE1 $\alpha$ -XBP1 Signaling through a Physical Interaction," *PLoS Biology*, vol. 8, pp. e1000410–e1000410, jul 2010.
- [298] A. Saleh, S. M. Srinivasula, L. Balkir, P. D. Robbins, and E. S. Alnemri, "Negative regulation of the apaf-1 apoptosome by hsp70," *Nature cell biology*, vol. 2, no. 8, pp. 476–483, 2000.
- [299] A. Asea, "Hsp70: a chaperokine," in *Novartis Foundation symposium*, vol. 291, p. 173, Wiley Online Library, 2008.

- [300] N. Somensi, P. O. Brum, V. de Miranda Ramos, J. Gasparotto, A. Zanotto-Filho, D. C. Rostirolla, M. da Silva Morone, J. C. F. Moreira, and D. P. Gelain, "Extracellular hsp70 activates erk1/2, nf-kb and pro-inflammatory gene transcription through binding with rage in a549 human lung cancer cells," *Cellular Physiology and Biochemistry*, vol. 42, no. 6, pp. 2507–2522, 2017.
- [301] R. M. Vabulas, P. Ahmad-Nejad, S. Ghose, C. J. Kirschning, R. D. Issels, and H. Wagner, "Hsp70 as endogenous stimulus of the toll/interleukin-1 receptor signal pathway," *Journal of Biological Chemistry*, vol. 277, no. 17, pp. 15107–15112, 2002.
- [302] R. Wadhwa, K. Taira, and S. C. Kaul, "An hsp70 family chaperone, mortalin/mthsp70/pbp74/grp75: what, when, and where?," *Cell stress & chaperones*, vol. 7, no. 3, p. 309, 2002.
- [303] Z. Flachbartová and B. Kovacech, "Mortalin—a multipotent chaperone regulating cellular processes ranging from viral infection to neurodegeneration," *Acta virologica*, vol. 57, pp. 3–15, 2013.
- [304] P. R. Dores-Silva, L. R. S. Barbosa, C. H. I. Ramos, and J. C. Borges, "Human Mitochondrial Hsp70 (Mortalin): Shedding Light on ATPase Activity, Interaction with Adenosine Nucleotides, Solution Structure and Domain Organization," *PLOS ONE*, vol. 10, pp. e0117170–e0117170, jan 2015.
- [305] A. S. Lee, "The ER chaperone and signaling regulator GRP78/BiP as a monitor of endoplasmic reticulum stress.," *Methods (San Diego, Calif.)*, vol. 35, pp. 373–381, apr 2005.
- [306] M. Sun, J. L. Kotler, S. Liu, and T. O. Street, "The endoplasmic reticulum (er) chaperones bip and grp94 selectively associate when bip is in the adp conformation," *Journal of Biological Chemistry*, vol. 294, no. 16, pp. 6387–6396, 2019.
- [307] M. E. Rubio and R. J. Wenthold, "Calnexin and the immunoglobulin binding protein (bip) coimmunoprecipitate with ampa receptors," *Journal of neurochemistry*, vol. 73, no. 3, pp. 942–948, 1999.
- [308] J. Behnke, M. J. Mann, F.-L. Scruggs, M. J. Feige, L. M. Hendershot Correspondence, M. F. De, and L. M. Hendershot, "Members of the Hsp70 Family Recognize Distinct Types of Sequences to Execute ER Quality Control Article Members of the Hsp70 Family Recognize Distinct Types of Sequences to Execute ER Quality Control," *Molecular Cell*, vol. 63, pp. 739–752, 2016.
- [309] M. Wang, S. Wey, Y. Zhang, R. Ye, and A. S. Lee, "Role of the unfolded protein response regulator grp78/bip in development, cancer, and neurological disorders," *Antioxidants & redox signaling*, vol. 11, no. 9, pp. 2307–2316, 2009.
- [310] B. E. Ledford and G. H. Leno, "ADP-ribosylation of the molecular chaperone GRP78/BiP.," *Molecular and cellular biochemistry*, vol. 138, pp. 141–8, sep 1994.
- [311] A. Nakai, T. Kawatani, S. Ohi, H. Kawasaki, T. Yoshimori, Y. Tashiro, Y. Miyata, I. Yahara, M. Satoh, and K. Nagata, "Expression and phosphorylation of bip/grp78, a molecular chaperone in the endoplasmic reticulum, during the differentiation of a mouse myeloblastic cell line," *Cell structure and function*, vol. 20, no. 1, pp. 33–39, 1995.

- [312] S. Díaz-Troya, M. E. Pérez-Pérez, M. Pérez-Martín, S. Moes, P. Jenó, F. J. Florencio, and J. L. Crespo, "Inhibition of protein synthesis by tor inactivation revealed a conserved regulatory mechanism of the bip chaperone in chlamydomonas," *Plant physiology*, vol. 157, no. 2, pp. 730–741, 2011.
- [313] I. Braakman and D. N. Hebert, "Protein folding in the endoplasmic reticulum," *Cold Spring Harbor perspectives in biology*, vol. 5, no. 5, p. a013201, 2013.
- [314] P. S. Pathinayake, A. C.-Y. Hsu, D. W. Waters, P. M. Hansbro, L. G. Wood, and P. A. Wark, "Understanding the unfolded protein response in the pathogenesis of asthma," *Frontiers in Immunology*, vol. 9, p. 175, 2018.
- [315] A. Chakrabarti, A. W. Chen, and J. D. Varner, "A review of the mammalian unfolded protein response," *Biotechnology and bioengineering*, vol. 108, no. 12, pp. 2777–2793, 2011.
- [316] M. C. Kopp, N. Larburu, V. Durairaj, C. J. Adams, and M. M. Ali, "Upr proteins ire1 and perk switch bip from chaperone to er stress sensor," *Nature structural & molecular biology*, vol. 26, no. 11, pp. 1053–1062, 2019.
- [317] D. Ron and P. Walter, "Signal integration in the endoplasmic reticulum unfolded protein response," *Nature Reviews Molecular Cell Biology*, vol. 8, pp. 519–529, jul 2007.
- [318] H. P. Harding, Y. Zhang, A. Bertolotti, H. Zeng, and D. Ron, "Perk is essential for translational regulation and cell survival during the unfolded protein response," *Molecular cell*, vol. 5, no. 5, pp. 897–904, 2000.
- [319] W. Rozpedek, D. Pytel, B. Mucha, H. Leszczynska, J. A. Diehl, and I. Majsterek, "The role of the perk/eif2 $\alpha$ /atf4/chop signaling pathway in tumor progression during endoplasmic reticulum stress," *Current molecular medicine*, vol. 16, no. 6, pp. 533–544, 2016.
- [320] P. Walter and D. Ron, "The unfolded protein response: from stress pathway to homeostatic regulation," *Science*, vol. 334, no. 6059, pp. 1081–1086, 2011.
- [321] K. Zhang and R. J. Kaufman, "From endoplasmic-reticulum stress to the inflammatory response.," *Nature*, vol. 454, pp. 455–462, jul 2008.
- [322] S. Huang, Y. Xing, and Y. Liu, "Emerging roles for the er stress sensor ire1 $\alpha$  in metabolic regulation and disease," *Journal of Biological Chemistry*, vol. 294, no. 49, pp. 18726–18741, 2019.
- [323] J. Celli and R. M. Tsois, "Bacteria, the ER and the Unfolded Protein Response: Friends or Foes?," *Nat Rev Microbiol*, vol. 13, no. 2, pp. 71–82, 2015.
- [324] A. Chakrabarti, A. W. Chen, and J. D. Varner, "A review of the mammalian unfolded protein response.," *Biotechnology and bioengineering*, vol. 108, pp. 2777–2793, dec 2011.
- [325] M. Keynes and T. M. Cox, "William Bateson, the rediscoverer of Mendel.," *Journal of the Royal Society of Medicine*, vol. 101, p. 104, mar 2008.
- [326] C. H. Waddington, "The Epigenotype," *International Journal of Epidemiology*, vol. 41, pp. 10–13, feb 2012.

- [327] R. Holliday, "DNA Methylation and Epigenetic Inheritance," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 326, pp. 329–338, jan 1990.
- [328] R. Mao and L.-S. Chou, "Methylation Analysis by Restriction Endonuclease Digestion and Real-Time PCR," *Clinical Chemistry*, vol. 56, pp. 1050–1052, jul 2010.
- [329] N. Jansz, "Dna methylation dynamics at transposable elements in mammals," *Essays in Biochemistry*, vol. 63, no. 6, pp. 677–689, 2019.
- [330] Y. Li and T. O. Tollefsbol, "Dna methylation detection: bisulfite genomic sequencing analysis," in *Epigenetics Protocols*, pp. 11–21, Springer, 2011.
- [331] M. J. Olszewska, D. Gernand, and T. Sakowicz, "Methylation-sensitive restriction endonuclease digestion patterns revealed in *Vicia faba* L. chromosomes by in situ nick-translation.," *Folia histochemica et cytobiologica*, vol. 37, no. 4, pp. 267–74, 1999.
- [332] G. Altun, J. F. Loring, and L. C. Laurent, "Dna methylation in embryonic stem cells," *Journal of cellular biochemistry*, vol. 109, no. 1, pp. 1–6, 2010.
- [333] E. Prokhortchouk and P.-A. Defossez, "The cell biology of DNA methylation in mammals," *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, vol. 1783, pp. 2167–2173, nov 2008.
- [334] R. Métivier, R. Gallais, C. Tiffoche, C. Le Péron, R. Z. Jurkowska, R. P. Carmouche, D. Ibberson, P. Barath, F. Demay, G. Reid, V. Benes, A. Jeltsch, F. Gannon, and G. Salbert, "Cyclical DNA methylation of a transcriptionally active promoter," *Nature*, vol. 452, pp. 45–50, mar 2008.
- [335] G. Reid, R. Gallais, and R. Métivier, "Marking time: The dynamic role of chromatin and covalent modification in transcription," *The International Journal of Biochemistry & Cell Biology*, vol. 41, pp. 155–163, jan 2009.
- [336] S. Kangaspeska, B. Stride, R. Métivier, M. Polycarpou-Schwarz, D. Ibberson, R. P. Carmouche, V. Benes, F. Gannon, and G. Reid, "Transient cyclical methylation of promoter DNA," *Nature*, vol. 452, pp. 112–115, mar 2008.
- [337] M. Okano, D. W. Bell, D. A. Haber, and E. Li, "Dna methyltransferases dnmt3a and dnmt3b are essential for de novo methylation and mammalian development," *Cell*, vol. 99, no. 3, pp. 247–257, 1999.
- [338] Y.-M. Sun and Y.-Q. Chen, "Principles and innovative technologies for decrypting noncoding rnas: from discovery and functional prediction to clinical application," *Journal of Hematology & Oncology*, vol. 13, no. 1, pp. 1–27, 2020.
- [339] M. Kazimierczyk, M. K. Kasprowicz, M. E. Kasprzyk, and J. Wrzesinski, "Human long noncoding rna interactome: detection, characterization and function," *International journal of molecular sciences*, vol. 21, no. 3, p. 1027, 2020.
- [340] A. Loda and E. Heard, "Xist rna in action: Past, present, and future," *PLoS genetics*, vol. 15, no. 9, p. e1008333, 2019.

- [341] M. Yun, J. Wu, J. L. Workman, and B. Li, "Readers of histone modifications," *Cell research*, vol. 21, no. 4, pp. 564–578, 2011.
- [342] T. Jenuwein and C. D. Allis, "Translating the histone code," *Science*, vol. 293, no. 5532, pp. 1074–1080, 2001.
- [343] R. C. Allshire and K. Ekwall, "Epigenetic regulation of chromatin states in *Schizosaccharomyces pombe*," *Cold Spring Harbor Perspectives in Biology*, vol. 7, pp. 1–25, jul 2015.
- [344] H. Santos-Rosa, R. Schneider, A. J. Bannister, J. Sherriff, B. E. Bernstein, N. T. Emre, S. L. Schreiber, J. Mellor, and T. Kouzarides, "Active genes are tri-methylated at k4 of histone h3," *Nature*, vol. 419, no. 6905, pp. 407–411, 2002.
- [345] K. Ahmad and S. Henikoff, "The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly," *Molecular Cell*, vol. 9, pp. 1191–1200, jun 2002.
- [346] M. M. Valdivia, K. Hamdouch, M. Ortiz, and A. Astola, "Cenpa a genomic marker for centromere activity and human diseases," *Current Genomics*, vol. 10, no. 5, pp. 326–335, 2009.
- [347] E. R. Blanco, L. Y. Kadyrova, and F. A. Kadyrov, "DNA mismatch repair interacts with CAF-1- and ASF1A-H3-H4-dependent histone (H3-H4)<sub>2</sub> tetramer deposition," *Journal of Biological Chemistry*, vol. 291, pp. 9203–9217, apr 2016.
- [348] J. P. Wilson, A. S. Raghavan, Y.-Y. Yang, G. Charron, and H. C. Hang, "Proteomic analysis of fatty-acylated proteins in mammalian cells with chemical reporters reveals S-acylation of histone H3 variants.," *Molecular & cellular proteomics : MCP*, vol. 10, pp. M110.001198–M110.001198, mar 2011.
- [349] H. Tagami, D. Ray-Gallet, G. Almouzni, and Y. Nakatani, "Histone h3. 1 and h3. 3 complexes mediate nucleosome assembly pathways dependent or independent of dna synthesis," *Cell*, vol. 116, no. 1, pp. 51–61, 2004.
- [350] F. Xu, K. Zhang, and M. Grunstein, "Acetylation in histone h3 globular domain regulates gene expression in yeast," *Cell*, vol. 121, no. 3, pp. 375–385, 2005.
- [351] "UPR (1)."
- [352] P. J. Robinson, W. An, A. Routh, F. Martino, L. Chapman, R. G. Roeder, and D. Rhodes, "30 nm Chromatin Fibre Decompaction Requires both H4-K16 Acetylation and Linker Histone Eviction," *Journal of Molecular Biology*, vol. 381, pp. 816–825, sep 2008.
- [353] C. Yan and D. D. Boyd, "Histone H3 Acetylation and H3 K4 Methylation Define Distinct Chromatin Regions Permissive for Transgene Expression," *Molecular and Cellular Biology*, vol. 26, pp. 6357–6371, sep 2006.
- [354] S. Mehrotra, L. Galdieri, T. Zhang, M. Zhang, L. F. Pemberton, and A. Vancura, "Histone hypoacetylation-activated genes are repressed by acetyl-coa-and chromatin-mediated mechanism," *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, vol. 1839, no. 9, pp. 751–763, 2014.



- [355] M. D. Stewart, J. Li, and J. Wong, "Relationship between Histone H3 Lysine 9 Methylation, Transcription Repression, and Heterochromatin Protein 1 Recruitment," *Molecular and Cellular Biology*, vol. 25, pp. 2525–2538, apr 2005.
- [356] M. D. Litt, M. Simpson, M. Gaszner, C. D. Allis, and G. Felsenfeld, "Correlation between histone lysine methylation and developmental changes at the chicken  $\beta$ -globin locus," *Science*, vol. 293, pp. 2453–2455, sep 2001.
- [357] Y. Zhang and D. Reinberg, "Transcription regulation by histone methylation: Interplay between different covalent modifications of the core histone tails," sep 2001.
- [358] L. Vandel and D. Trouche, "Physical association between the histone acetyl transferase CBP and a histone methyl transferase," *EMBO Reports*, vol. 2, no. 1, p. 21, 2001.
- [359] R. J. Sims and D. Reinberg, "Histone H3 Lys 4 methylation: Caught in a bind?," oct 2006.
- [360] C. L. Peterson and M. A. Laniel, "Histones and histone modifications.," jul 2004.
- [361] H. Wang, L. Zhai, J. Xu, H.-Y. Joo, S. Jackson, H. Erdjument-Bromage, P. Tempst, Y. Xiong, and Y. Zhang, "Histone h3 and h4 ubiquitylation by the cul4-ddb-roc1 ubiquitin ligase facilitates cellular response to dna damage," *Molecular cell*, vol. 22, no. 3, pp. 383–394, 2006.
- [362] R. Meas and P. Mao, "Histone ubiquitylation and its roles in transcription and dna damage response," *DNA repair*, vol. 36, pp. 36–42, 2015.
- [363] X. Zhang, B. Li, A. H. Rezaeian, X. Xu, P.-C. Chou, G. Jin, F. Han, B.-S. Pan, C.-Y. Wang, J. Long, *et al.*, "H3 ubiquitination by nedd4 regulates h3 acetylation and tumorigenesis," *Nature communications*, vol. 8, no. 1, pp. 1–15, 2017.
- [364] Y. Xia, W. Yang, M. Fa, X. Li, Y. Wang, Y. Jiang, Y. Zheng, J.-H. Lee, J. Li, and Z. Lu, "Rnf8 mediates histone h3 ubiquitylation and promotes glycolysis and tumorigenesis," *Journal of Experimental Medicine*, vol. 214, no. 6, pp. 1843–1855, 2017.
- [365] H. Kikuchi, B. Yuan, E. Yuhara, N. Takagi, and H. Toyoda, "Involvement of histone H3 phosphorylation through p38 MAPK pathway activation in casticin-induced cytotoxic effects against the human promyelocytic cell line HL-60," *International Journal of Oncology*, vol. 43, pp. 2046–2056, dec 2013.
- [366] I. S. Strelkov and J. R. Davie, "Ser-10 phosphorylation of histone H3 and immediate early gene expression in oncogene-transformed mouse fibroblasts.," *Cancer research*, vol. 62, pp. 75–8, jan 2002.
- [367] G. P. Vicent, C. Ballaré, A. S. Nacht, J. Clausell, A. Subtil-Rodríguez, I. Quiles, A. Jordan, and M. Beato, "Convergence on chromatin of non-genomic and genomic pathways of hormone signaling," *Journal of Steroid Biochemistry and Molecular Biology*, vol. 109, pp. 344–349, apr 2008.
- [368] H. Bierne and R. Pourpre, "Bacterial factors targeting the nucleus: the growing family of nucleomodulins," *Toxins*, vol. 12, no. 4, p. 220, 2020.

- [369] H. E. Hanford, J. Von Dwingelo, and Y. Abu Kwaik, “Bacterial nucleomodulins: A coevolutionary adaptation to the eukaryotic command center,” *PLoS Pathogens*, vol. 17, no. 1, p. e1009184, 2021.
- [370] K. Maeshima, H. Iino, S. Hihara, and N. Imamoto, “Nuclear size, nuclear pore number and cell cycle,” *Nucleus*, vol. 2, no. 2, pp. 1065–1071, 2011.
- [371] I. V. Aramburu and E. A. Lemke, “Floppy but not sloppy: Interaction mechanism of fg-nucleoporins and nuclear transport receptors,” in *Seminars in cell & developmental biology*, vol. 68, pp. 34–41, Elsevier, 2017.
- [372] R. Wang and M. G. Brattain, “The maximal size of protein to diffuse through the nuclear pore is larger than 60 kda,” *FEBS letters*, vol. 581, no. 17, pp. 3164–3170, 2007.
- [373] N. Freitas and C. Cunha, “Mechanisms and Signals for the Nuclear Import of Proteins,” *Current Genomics*, vol. 10, pp. 550–557, 2009.
- [374] M. R. Hodel, A. H. Corbett, and A. E. Hodel, “Dissection of a nuclear localization signal,” *Journal of Biological Chemistry*, vol. 276, no. 2, pp. 1317–1325, 2001.
- [375] G. Cingolani, J. Bednenko, M. T. Gillespie, and L. Gerace, “Molecular basis for the recognition of a nonclassical nuclear localization signal by importin  $\beta$ ,” *Molecular cell*, vol. 10, no. 6, pp. 1345–1353, 2002.
- [376] A. Lange, R. E. Mills, C. J. Lange, M. Stewart, S. E. Devine, and A. H. Corbett, “Classical nuclear localization signals: definition, function, and interaction with importin  $\alpha$ ,” *Journal of Biological Chemistry*, vol. 282, no. 8, pp. 5101–5105, 2007.
- [377] R. Nakada, H. Hirano, and Y. Matsuura, “Structure of importin- $\alpha$  bound to a non-classical nuclear localization signal of the influenza a virus nucleoprotein,” *Scientific reports*, vol. 5, no. 1, pp. 1–9, 2015.
- [378] M. Iijima, M. Suzuki, A. Tanabe, A. Nishimura, and M. Yamada, “Two motifs essential for nuclear import of the hnmp a1 nucleocytoplasmic shuttling sequence m9 core,” *FEBS letters*, vol. 580, no. 5, pp. 1365–1370, 2006.
- [379] U. Fischer, V. Sumpster, M. Sekine, T. Satoh, and R. Lührmann, “Nucleo-cytoplasmic transport of U snRNPs: definition of a nuclear location signal in the Sm core domain that binds a transport receptor independently of the m3G cap,” *The EMBO journal*, vol. 12, pp. 573–83, feb 1993.
- [380] R. Contreras, P. Kallemini, M. P. González-García, A. Lazarova, J. J. Sánchez-Serrano, M. Sanmartín, and E. Rojo, “Identification of domains and factors involved in miniyo nuclear import,” *Frontiers in plant science*, vol. 10, p. 1044, 2019.
- [381] B. J. Lee, A. E. Cansizoglu, K. E. Süel, T. H. Louis, Z. Zhang, and Y. M. Chook, “Rules for nuclear localization sequence recognition by karyopherin beta 2,” *Cell*, vol. 126, pp. 543–558, aug 2006.
- [382] A. E. Cansizoglu, B. J. Lee, Z. C. Zhang, B. M. A. Fontoura, and Y. M. Chook, “Structure-based design of a pathway-specific nuclear import inhibitor,” *Nature structural & molecular biology*, vol. 14, pp. 452–4, may 2007.

- [383] P. Shaw and E. Jordan, "The nucleolus," *Annual review of cell and developmental biology*, vol. 11, no. 1, pp. 93–121, 1995.
- [384] M. Li, S. Wang, M. Cai, and C. Zheng, "Identification of nuclear and nucleolar localization signals of pseudorabies virus (prv) early protein ul54 reveals that its nuclear targeting is required for efficient production of prv," *Journal of virology*, vol. 85, no. 19, pp. 10239–10251, 2011.
- [385] K. Mekhail, L. Rivero-Lopez, A. Al-Masri, C. Brandon, M. Khacho, and S. Lee, "Identification of a common subnuclear localization signal," *Molecular biology of the cell*, vol. 18, no. 10, pp. 3966–3977, 2007.
- [386] K. Mekhail, M. Khacho, A. Carrigan, R. R. Hache, L. Gunaratnam, and S. Lee, "Regulation of ubiquitin ligase dynamics by the nucleolus," *The Journal of cell biology*, vol. 170, no. 5, pp. 733–744, 2005.
- [387] M. Wang, M. Bokros, P. R. Theodoridis, and S. Lee, "Nucleolar sequestration: remodeling nucleoli into amyloid bodies," *Frontiers in genetics*, vol. 10, p. 1179, 2019.
- [388] T. E. Audas, M. D. Jacob, and S. Lee, "Immobilization of proteins in the nucleolus by ribosomal intergenic spacer noncoding rna," *Molecular cell*, vol. 45, no. 2, pp. 147–157, 2012.
- [389] S. R. Gunawardena, B. L. Ruis, J. A. Meyer, M. Kapoor, and K. F. Conklin, "Nom1 targets protein phosphatase i to the nucleolus," *Journal of Biological Chemistry*, vol. 283, no. 1, pp. 398–404, 2008.
- [390] S. L. Brown, D. J. Garrison, and J. P. May, "Phase separation of a plant virus movement protein and cellular factors support virus-host interactions," *PLoS pathogens*, vol. 17, no. 9, p. e1009622, 2021.
- [391] H. Bierne and P. Cossart, "When bacteria target the nucleus: the emerging family of nucleomodulins," *Cellular microbiology*, vol. 14, pp. 622–633, may 2012.
- [392] D. T. Isaac and R. Isberg, "No Title," 2014.
- [393] A. N. N. Ba, A. Pogoutse, N. Provar, and A. M. Moses, "Nlstradamus: a simple hidden markov model for nuclear localization signal prediction," *BMC bioinformatics*, vol. 10, no. 1, p. 202, 2009.
- [394] X. Charpentier and E. Oswald, "Identification of the secretion and translocation domain of the enteropathogenic and enterohemorrhagic escherichia coli effector cif, using tem-1  $\beta$ -lactamase as a new fluorescence-based reporter," *Journal of bacteriology*, vol. 186, no. 16, pp. 5486–5495, 2004.
- [395] K. C. Jeong, D. Ghosal, Y.-W. Chang, G. J. Jensen, and J. P. Vogel, "Polar delivery of legionella type iv secretion system substrates is essential for virulence," *Proceedings of the National Academy of Sciences*, vol. 114, no. 30, pp. 8077–8082, 2017.
- [396] J.-H. Lee, J. You, E. Dobrota, and D. G. Skalnik, "Identification and characterization of a novel human pp1 phosphatase complex," *Journal of Biological Chemistry*, vol. 285, no. 32, pp. 24466–24476, 2010.

- [397] D. B. Bloch, M. Suzanne, P. Guigaouri, A. Filippov, and K. D. Bloch, "Identification and characterization of a leukocyte-specific component of the nuclear body," *Journal of Biological Chemistry*, vol. 271, no. 46, pp. 29198–29204, 1996.
- [398] M. Abbey, V. Trush, E. Gibson, and M. Vedadi, "Targeting human retinoblastoma binding protein 4 (rbbp4) and 7 (rbbp7)," *bioRxiv*, p. 303537, 2018.
- [399] B. D. Hopkins, C. Hodakoski, D. Barrows, S. M. Mense, and R. E. Parsons, "Pten function: the long and the short of it," *Trends in biochemical sciences*, vol. 39, no. 4, pp. 183–190, 2014.
- [400] C. Puri, M. M. Manni, M. Vicinanza, C. Hilcenko, Y. Zhu, G. Runwal, E. Stamatakou, F. M. Menzies, K. Mamchaoui, M. Bitoun, *et al.*, "A dnm2 centronuclear myopathy mutation reveals a link between recycling endosome scission and autophagy," *Developmental Cell*, vol. 53, no. 2, pp. 154–168, 2020.
- [401] T. N. M. Dao, S.-H. Kang, A. Bak, and S. Y. Folimonova, "A non-conserved p33 protein of citrus tristeza virus interacts with multiple viral partners," *Molecular Plant-Microbe Interactions*, vol. 33, no. 6, pp. 859–870, 2020.
- [402] M. Weber, B. Beyene, N. Nagler, J. Herfert, S. Schempp, M. Klecker, and S. Clemens, "A mutation in the essential and widely conserved damaged dna binding1-cullin4 associated factor gene *ozs3* causes hypersensitivity to zinc excess, cold and uv stress in *arabidopsis thaliana*," *The Plant Journal*, 2020.
- [403] M. Ni, Y. Zhang, and A. S. Lee, "Beyond the endoplasmic reticulum: atypical GRP78 in cell viability, signaling and therapeutic targeting," *Biochemical Journal*, vol. 434, no. 2, p. 181, 2011.
- [404] P. Rodriguez, B. Mitton, and E. G. Kranias, "Phosphorylation of glutathione-s-transferase by protein kinase c- $\alpha$  implications for affinity-tag purification," *Biotechnology letters*, vol. 27, no. 23-24, pp. 1869–1873, 2005.
- [405] A. D. Caperta, M. Rosa, M. Delgado, R. Karimi, D. Demidov, W. Viegas, and A. Houben, "Distribution patterns of phosphorylated Thr 3 and Thr 32 of histone H3 in plant mitosis and meiosis," *Cytogenetic and Genome Research*, vol. 122, no. 1, pp. 73–79, 2008.
- [406] L. Zhang, D. Song, B. Zhu, and X. Wang, "The role of nuclear matrix protein hnrnpu in maintaining the architecture of 3d genome," in *Seminars in cell & developmental biology*, vol. 90, pp. 161–167, Elsevier, 2019.
- [407] M.-E. Lalonde, N. Avvakumov, K. C. Glass, F.-H. Joncas, N. Saksouk, M. Holliday, E. Paquet, K. Yan, Q. Tong, B. J. Klein, *et al.*, "Exchange of associated factors directs a switch in hbo1 acetyltransferase histone tail specificity," *Genes & development*, vol. 27, no. 18, pp. 2009–2024, 2013.
- [408] Y. Doyon, C. Cayrou, M. Ullah, A.-J. Landry, V. Côté, W. Selleck, W. S. Lane, S. Tan, X.-J. Yang, and J. Côté, "Ing tumor suppressor proteins are critical regulators of chromatin acetylation required for genome expression and perpetuation," *Molecular cell*, vol. 21, no. 1, pp. 51–64, 2006.
- [409] Y. Liang, Y. Su, C. Xu, N. Zhang, D. Liu, G. Li, T. Tong, and J. Chen, "Protein kinase d1 phosphorylation of *kat7* enhances its protein stability and promotes replication licensing and cell proliferation," *Cell death discovery*, vol. 6, no. 1, pp. 1–14, 2020.

- [410] R. Tuteja and N. Tuteja, "Nucleolin: a multifunctional major nucleolar phosphoprotein," *Critical reviews in biochemistry and molecular biology*, vol. 33, no. 6, pp. 407–436, 1998.
- [411] A. Krause and I. Hoffmann, "Polo-like kinase 2-dependent phosphorylation of npm/b23 on serine 4 triggers centriole duplication," *PloS one*, vol. 5, no. 3, p. e9849, 2010.
- [412] C. R. Roy, K. H. Berger, and R. R. Isberg, "Legionella pneumophila dota protein is required for early phagosome trafficking decisions that occur within minutes of bacterial uptake," *Molecular microbiology*, vol. 28, no. 3, pp. 663–674, 1998.
- [413] N. Tandon, K. N. Thakkar, E. L. LaGory, Y. Liu, and A. J. Giaccia, "Generation of stable expression mammalian cell lines using lentivirus," *Bio-protocol*, vol. 8, no. 21, 2018.
- [414] R. J. Abbott, I. Spendlove, P. Roversi, H. Fitzgibbon, V. Knott, P. Teriete, J. M. McDonnell, P. A. Handford, and S. M. Lea, "Structural and functional characterization of a novel t cell receptor co-regulatory protein complex, cd97-cd55," *Journal of Biological Chemistry*, vol. 282, no. 30, pp. 22023–22032, 2007.
- [415] M. Daifuku, K. Nishi, T. Okamoto, and T. Sugahara, "Activation of j774. 1 murine macrophages by lactate dehydrogenase," *Cytotechnology*, vol. 66, no. 6, pp. 937–943, 2014.
- [416] A. N. Nguyen Ba, A. Pogoutse, N. Provar, and A. M. Moses, "NLStradamus: A simple Hidden Markov Model for nuclear localization signal prediction," *BMC Bioinformatics*, vol. 10, jun 2009.
- [417] M. S. Scott, P. V. Troshin, and G. J. Barton, "Nod: a nucleolar localization sequence detector for eukaryotic and viral proteins," *BMC bioinformatics*, vol. 12, no. 1, pp. 1–7, 2011.
- [418] T. E. Audas, M. D. Jacob, and S. Lee, "The nucleolar detention pathway: A cellular strategy for regulating molecular networks," *Cell Cycle*, vol. 11, no. 11, pp. 2059–2062, 2012.
- [419] A. Cassany and L. Gerace, "Reconstitution of Nuclear Import in Permeabilized Cells," in *Methods in Molecular Biology (Clifton, N.J.)*, vol. 464, pp. 181–205, NIH Public Access, 2008.
- [420] P. Goyal, D. Pandey, and W. Siess, "Phosphorylation-dependent regulation of unique nuclear and nucleolar localization signals of lim kinase 2 in endothelial cells," *Journal of Biological Chemistry*, vol. 281, no. 35, pp. 25223–25230, 2006.
- [421] C. Hoffmann, I. Finsel, and H. Hilbi, "Purification of pathogen vacuoles from legionella-infected phagocytes," *JoVE (Journal of Visualized Experiments)*, no. 64, p. e41118, 2012.
- [422] H. Tamada, N. V. Thuan, P. Reed, D. Nelson, N. Katoku-Kikyo, J. Wudel, T. Wakayama, N. Kikyo, N. Van Thuan, P. Reed, D. Nelson, N. Katoku-Kikyo, J. Wudel, T. Wakayama, and N. Kikyo, "Chromatin Decondensation and Nuclear Reprogramming by Nucleoplasmin," *Molecular and Cellular Biology*, vol. 26, pp. 1259–1271, feb 2006.
- [423] H.-O. Rashid, R. K. Yadav, H.-R. Kim, and H.-J. Chae, "Er stress: Autophagy induction, inhibition and selection," *Autophagy*, vol. 11, no. 11, pp. 1956–1977, 2015.

- [424] G. Oliva, 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*, vol. 8, p. 3, 2018.
- [425] A. T. Quaile, P. J. Stogios, O. Egorova, E. Evdokimova, D. Valleau, B. Nocek, P. S. Kompella, S. Peisajovich, A. F. Yakunin, A. W. Ensminger, and A. Savchenko, "The *Legionella pneumophila* effector Ceg4 is a phosphotyrosine phosphatase that attenuates activation of eukaryotic MAPK pathways," *Journal of Biological Chemistry*, vol. 293, pp. 3307–3320, mar 2018.
- [426] P. Aurass, T. Gerlach, D. Becher, B. Voigt, S. Karste, J. Bernhardt, K. Riedel, M. Hecker, and A. Flieger, "Life stage-specific proteomes of *legionella pneumophila* reveal a highly differential abundance of virulence-associated dot/icm effectors," *Molecular & Cellular Proteomics*, vol. 15, no. 1, pp. 177–200, 2016.
- [427] P. Aurass, M. Schlegel, O. Metwally, C. R. Harding, G. N. Schroeder, G. Frankel, and A. Flieger, "The *Legionella pneumophila* Dot/Icm-secreted effector PlcC/CegC1 together with PlcA and PlcB promotes virulence and belongs to a novel zinc metallophospholipase C family present in bacteria and fungi," *Journal of Biological Chemistry*, vol. 288, pp. 11080–11092, apr 2013.
- [428] A. A. Khweek, K. Caution, A. Akhter, B. A. Abdulrahman, M. Tazi, H. Hassan, N. Majumdar, A. Doran, E. Guirado, L. S. Schlesinger, H. Shuman, and A. O. Amer, "A bacterial protein promotes the recognition of the *Legionella pneumophila* vacuole by autophagy," *European Journal of Immunology*, vol. 43, pp. 1333–1344, apr 2013.
- [429] N. K. Glueck, K. M. O'Brien, D. C. Seguin, and V. J. Starai, "*Legionella pneumophila* legc7 effector protein drives aberrant endoplasmic reticulum: endosome contacts in yeast," *Traffic*, vol. 22, no. 8, pp. 284–302, 2021.
- [430] J. Wang, X. Li, A. Chen, and Y. Lu, "[*Legionella pneumophila* eukaryotic-like effector LegK3 inhibits growth of *Saccharomyces cerevisiae* and modulates its vesicle trafficking pathway].," *Wei sheng wu xue bao = Acta microbiologica Sinica*, vol. 54, pp. 417–423, apr 2014.
- [431] P.-C. Lee, M. P. Machner, and G. Abstract, "The *Legionella* Effector Kinase LegK7 Hijacks the Host Hippo Pathway to Promote Infection," *Cell host & microbe*, vol. 24, pp. 429–438.e6, sep 2018.
- [432] K. Beyrakhova, L. Li, C. Xu, A. Gagarinova, and M. Cygler, "*Legionella pneumophila* effector Lem4 is a membrane-associated protein tyrosine phosphatase," *Journal of Biological Chemistry*, vol. 293, pp. 13044–13058, aug 2018.
- [433] N. Nachmias, T. Zusman, and G. Segal, "Study of *legionella* effector domains revealed novel and prevalent phosphatidylinositol 3-phosphate binding domains," *Infection and immunity*, vol. 87, no. 6, 2019.
- [434] T. Kubori, T. Kitao, H. Ando, and H. Nagai, "LotA, a *Legionella* deubiquitinase, has dual catalytic activity and contributes to intracellular growth," *Cellular Microbiology*, vol. 20, jul 2018.
- [435] J. Chen, M. Reyes, M. Clarke, and H. A. Shuman, "Host cell-dependent secretion and translocation of the lepa and lepb effectors of *legionella pneumophila*," *Cellular microbiology*, vol. 9, no. 7, pp. 1660–1671, 2007.

- [436] A. Sreelatha, C. Nolan, B. C. Park, K. Pawłowski, D. R. Tomchick, and V. S. Tagliabracci, “A legionella effector kinase is activated by host inositol hexakisphosphate,” *Journal of Biological Chemistry*, vol. 295, no. 18, pp. 6214–6224, 2020.
- [437] N. Gan, H. Guan, Y. Huang, T. Yu, J. Fu, E. S. Nakayasu, K. Puvar, C. Das, D. Wang, S. Ouyang, and Z. Luo, “Legionella pneumophila regulates the activity of UBE2N by deamidase-mediated deubiquitination,” *The EMBO Journal*, vol. 39, feb 2020.
- [438] A. Mousnier, G. N. Schroeder, C. A. Stoneham, E. C. So, J. A. Garnett, L. Yu, S. J. Matthews, J. S. Choudhary, E. L. Hartland, and G. Frankel, “A new method to determine in vivo interactomes reveals binding of the Legionella pneumophila effector PieE to multiple rab GTPases,” *mBio*, vol. 5, pp. e01148–14, aug 2014.
- [439] Y.-H. Lin, M. Lucas, T. R. Evans, G. Abascal-Palacios, A. G. Doms, N. A. Beauchene, A. L. Rojas, A. Hierro, and M. P. Machner, “Ravn is a member of a previously unrecognized group of legionella pneumophila e3 ubiquitin ligases,” *PLoS pathogens*, vol. 14, no. 2, p. e1006897, 2018.
- [440] C. R. Harding, C. A. Stoneham, R. Schuelein, H. Newton, C. V. Oates, E. L. Hartland, G. N. Schroeder, and G. Frankel, “The Dot/Icm Effector SdhA Is Necessary for Virulence of Legionella pneumophila in Galleria mellonella and A/J Mice,” *Infection and Immunity*, vol. 81, no. 7, pp. 2598–2605, 2013.
- [441] E. A. Creasey and R. R. Isberg, “The protein sdha maintains the integrity of the legionella-containing vacuole,” *Proceedings of the National Academy of Sciences*, vol. 109, no. 9, pp. 3481–3486, 2012.
- [442] Y. Chen, I. Tascón, M. R. Neunuebel, C. Pallara, J. Brady, L. N. Kinch, J. Fernández-Recio, A. L. Rojas, M. P. Machner, and A. Hierro, “Structural basis for rab1 de-amylation by the legionella pneumophila effector sidd,” *PLoS Pathog*, vol. 9, no. 5, p. e1003382, 2013.
- [443] J. Qiu, M. J. Sheedlo, K. Yu, Y. Tan, E. S. Nakayasu, C. Das, X. Liu, and Z.-Q. Luo, “Ubiquitination independent of E1 and E2 enzymes by bacterial effectors,” *Nature*, vol. 533, pp. 120–4, may 2016.
- [444] A. Prashar, M. E. Ortiz, S. Lucarelli, E. Barker, Z. Tabatabeiyazdi, F. Shamoun, D. Raju, C. Antonescu, C. Guyard, and M. R. Terebiznik, “Small Rho GTPases and the Effector VipA Mediate the Invasion of Epithelial Cells by Filamentous Legionella pneumophila,” *Frontiers in Cellular and Infection Microbiology*, vol. 8, p. 133, may 2018.
- [445] Q. Mou and P. H. M. Leung, “Differential expression of virulence genes in Legionella pneumophila growing in Acanthamoeba and human monocytes,” *Virulence*, vol. 9, no. 1, pp. 185–196, 2018.
- [446] S. Shorbagi and I. R. Brown, “Dynamics of the association of heat shock protein HSPA6 (Hsp70B’) and HSPA1A (Hsp70-1) with stress-sensitive cytoplasmic and nuclear structures in differentiated human neuronal cells,” *Cell Stress and Chaperones*, vol. 21, pp. 993–1003, nov 2016.
- [447] C. A. S. Deane and I. R. Brown, “Differential Targeting of Hsp70 Heat Shock Proteins HSPA6 and HSPA1A with Components of a Protein Disaggregation/Refolding Machine in Differentiated Human Neuronal Cells following Thermal Stress,” *Frontiers in Neuroscience*, vol. 11, p. 227, apr 2017.

- [448] A. D. Bilog, L. Smulders, R. Oliverio, C. Labanieh, J. Zapanta, R. V. Stahelin, and N. Nikolaidis, "Membrane localization of hspa1a, a stress inducible 70-kda heat-shock protein, depends on its interaction with intracellular phosphatidylserine," *Biomolecules*, vol. 9, no. 4, p. 152, 2019.
- [449] J. Lee, Y. Han, Y. Yoon, C. Yun, S. Yun, S. Kim, H. Kwon, D. Jeong, M. Baek, H. Lee, *et al.*, "Role of hspa11 as a cellular prion protein stabilizer in tumor progression via hif-1 $\alpha$ /gp78 axis," *Oncogene*, vol. 36, no. 47, pp. 6555–6567, 2017.
- [450] K. A. Redgrove, B. Nixon, M. A. Baker, L. Hetherington, G. Baker, D.-Y. Liu, and R. J. Aitken, "The Molecular Chaperone HSPA2 Plays a Key Role in Regulating the Expression of Sperm Surface Receptors That Mediate Sperm-Egg Recognition," *PLoS ONE*, vol. 7, pp. e50851–e50851, nov 2012.
- [451] D. R. Sojka, A. Gogler-Pigłowska, N. Vydra, A. J. Cortez, P. T. Filipczak, Z. Krawczyk, and D. Sciegłńska, "Functional redundancy of hspa1, hspa2 and other hspa proteins in non-small cell lung carcinoma (nscLc); an implication for nscLc treatment," *Scientific reports*, vol. 9, no. 1, pp. 1–15, 2019.
- [452] B. A. Mohamed, A. Z. Barakat, W. H. Zimmermann, R. E. Bittner, C. Mühlfeld, M. Hünlich, W. Engel, L. S. Maier, and I. M. Adham, "Targeted disruption of Hspa4 gene leads to cardiac hypertrophy and fibrosis," *Journal of Molecular and Cellular Cardiology*, vol. 53, pp. 459–468, oct 2012.
- [453] T. Held, A. Z. Barakat, B. A. Mohamed, I. Paprotta, A. Meinhardt, W. Engel, and I. M. Adham, "Heat-shock protein hspa4 is required for progression of spermatogenesis," *Reproduction*, vol. 142, no. 1, p. 133, 2011.
- [454] H. Takahashi, T. Furukawa, T. Yano, N. Sato, J. Takizawa, T. Kurasaki, T. Abe, M. Narita, M. Masuko, S. Koyama, K. Toba, M. Takahashi, and Y. Aizawa, "Identification of an overexpressed gene, HSPA4L, the product of which can provoke prevalent humoral immune responses in leukemia patients," *Experimental Hematology*, vol. 35, pp. 1091–1099, jul 2007.
- [455] A. L. Laitusis, M. A. Brostrom, and C. O. Brostrom, "The dynamic role of GRP78/BiP in the coordination of mRNA translation with protein processing," *Journal of Biological Chemistry*, vol. 274, pp. 486–493, jan 1999.
- [456] M. Wang, S. Wey, Y. Zhang, R. Ye, and A. S. Lee, "Role of the unfolded protein response regulator GRP78/BiP in development, cancer, and neurological disorders.," *Antioxidants & redox signaling*, vol. 11, pp. 2307–2316, sep 2009.
- [457] Y. Fu, S. Wey, M. Wang, R. Ye, C.-P. Liao, P. Roy-Burman, and A. S. Lee, "Pten null prostate tumorigenesis and AKT activation are blocked by targeted knockout of ER chaperone GRP78/BiP in prostate epithelium.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, pp. 19444–19449, dec 2008.
- [458] P. Baumeister, D. Dong, Y. Fu, and A. S. Lee, "Transcriptional induction of GRP78/BiP by histone deacetylase inhibitors and resistance to histone deacetylase inhibitor-induced apoptosis," *Molecular Cancer Therapeutics*, vol. 8, no. 5, pp. 1086–1094, 2009.



- [459] T. K. Leung, C. Hall, M. Rajendran, N. K. Spurr, and L. Lim, "The human heat-shock genes hspa6 and hspa7 are both expressed and localize to chromosome 1," *Genomics*, vol. 12, no. 1, pp. 74–79, 1992.
- [460] J. Wang, A. Fedoseienko, B. Chen, E. Burstein, D. Jia, D. D. Billadeau, and R. J. Carver, "Endosomal Receptor Trafficking: Retromer and Beyond Graphical Abstract HHS Public Access," *Traffic*, vol. 19, no. 8, pp. 578–590, 2018.
- [461] J. Grizenkova, S. Akhtar, H. Hummerich, A. Tomlinson, E. A. Asante, A. Wenborn, J. Fizet, M. Poulter, F. K. Wiseman, E. M. C. Fisher, V. L. J. Tybulewicz, S. Brandner, J. Collinge, and S. E. Lloyd, "Overexpression of the Hspa13 (Stch) gene reduces prion disease incubation time in mice," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, pp. 13722–13727, aug 2012.
- [462] C. H. Yun, Y. C. Kim, and B. H. Jeong, "The First Report on Single Nucleotide Polymorphisms of the HSPA13 Gene in Koreans," *Russian Journal of Genetics*, vol. 54, pp. 353–357, mar 2018.
- [463] Q. Meng, B. X. Li, and X. Xiao, "Toward developing chemical modulators of hsp60 as potential therapeutics," *Frontiers in Molecular Biosciences*, vol. 5, p. 35, 2018.
- [464] T. Okamoto, H. Yamamoto, I. Kudo, K. Matsumoto, M. Odaka, E. Grave, and H. Itoh, "Hsp60 possesses a gtpase activity and mediates protein folding with hsp10," *Scientific reports*, vol. 7, no. 1, pp. 1–11, 2017.
- [465] R. Lixin, A. Efthymiadis, B. Henderson, and D. A. Jans, "Novel properties of the nucleolar targeting signal of human angiogenin," *Biochemical and Biophysical Research Communications*, vol. 284, pp. 185–193, jun 2001.
- [466] H. Kim, D. J. Chang, J. A. Lee, Y. S. Lee, and B. K. Kaang, "Identification of nuclear/nucleolar localization signal in Aplysia learning associated protein of slug with a molecular mass of 18 kDa homologous protein," *Neuroscience Letters*, vol. 343, pp. 134–138, jun 2003.
- [467] Y. Zhang and Y. Xiong, "Mutations in human ARF exon 2 disrupt its nucleolar localization and impair its ability to block nuclear export of MDM2 and p53," *Molecular Cell*, vol. 3, pp. 579–591, may 1999.
- [468] Y. Nishimura, T. Ohkubo, Y. Furuichi, and H. Umekawa, "Tryptophans 286 and 288 in the C-terminal Region of Protein B23.1 are Important for Its Nucleolar Localization," *Bioscience, Biotechnology and Biochemistry*, vol. 66, pp. 2239–2242, jan 2002.
- [469] A. H. Stegh, O. Schickling, A. Ehret, C. Scaffidi, C. Peterhänsel, T. G. Hofmann, I. Grummt, P. H. Kramer, and M. E. Peter, "Dedd, a novel death effector domain-containing protein, targeted to the nucleolus," *The EMBO journal*, vol. 17, no. 20, pp. 5974–5986, 1998.
- [470] Y. Y. Mo, C. Wang, and W. T. Beck, "A novel nuclear localization signal in human DNA topoisomerase I," *Journal of Biological Chemistry*, vol. 275, pp. 41107–41113, dec 2000.
- [471] J. Liu, X. Du, and Y. Ke, "Mapping nucleolar localization sequences of 1A6/DRIM," *FEBS Letters*, vol. 580, pp. 1405–1410, feb 2006.

- [472] Z. Sheng, J. A. Lewis, and W. J. Chirico, "Nuclear and nucleolar localization of 18-kDa fibroblast growth factor-2 is controlled by C-terminal signals," *Journal of Biological Chemistry*, vol. 279, pp. 40153–40160, sep 2004.
- [473] M. Antoine, K. Reimers, C. Dickson, and P. Kiefer, "Fibroblast growth factor 3, a protein with dual subcellular localization, is targeted to the nucleus and nucleolus by the concerted action of two nuclear localization signals and a nucleolar retention signal," *Journal of Biological Chemistry*, vol. 272, pp. 29475–29481, nov 1997.
- [474] F. Tamanini, L. L. Kirkpatrick, J. Schonkeren, L. v. Unen, C. Bontekoe, C. Bakker, D. L. Nelson, H. Galjaard, B. A. Oostra, and A. T. Hoogeveen, "The fragile x-related proteins fxr1p and fxr2p contain a functional nucleolar-targeting signal equivalent to the hiv-1 regulatory proteins," *Human molecular genetics*, vol. 9, no. 10, pp. 1487–1493, 2000.
- [475] W. S. Brooks, S. Banerjee, and D. F. Crawford, "G2E3 is a nucleo-cytoplasmic shuttling protein with DNA damage responsive localization," *Experimental Cell Research*, vol. 313, pp. 665–676, feb 2007.
- [476] Y. X. Guo, K. Dallmann, and J. Kwang, "Identification of nucleolus localization signal of betanodavirus GGNNV protein  $\alpha$ ," *Virology*, vol. 306, pp. 225–235, feb 2003.
- [477] M. Scott, F.-M. Boisvert, D. Vieyra, R. N. Johnston, D. P. Bazett-Jones, and K. Riabowol, "Uv induces nucleolar translocation of ing1 through two distinct nucleolar targeting sequences," *Nucleic Acids Research*, vol. 29, no. 10, pp. 2052–2058, 2001.
- [478] M. R. Rao, G. Kumari, D. Balasundaram, R. Sankaranarayanan, and S. Mahalingam, "A Novel Lysine-rich Domain and GTP Binding Motifs Regulate the Nucleolar Retention of Human Guanine Nucleotide Binding Protein, GNL3L," *Journal of Molecular Biology*, vol. 364, pp. 637–654, dec 2006.
- [479] S. Thébault, J. Basbous, B. Gay, C. Devaux, and J. M. Mesnard, "Sequence requirement for the nucleolar localization of human I-mfa domain-containing protein (HIC p40)," *European Journal of Cell Biology*, vol. 79, no. 11, pp. 834–838, 2000.
- [480] A. W. Cochrane, A. Perkins, and C. A. Rosen, "Identification of sequences important in the nucleolar localization of human immunodeficiency virus Rev: relevance of nucleolar localization to function.," *Journal of virology*, vol. 64, pp. 881–885, feb 1990.
- [481] H. Siomi, H. Shida, M. Maki, and M. Hatanaka, "Effects of a highly basic region of human immunodeficiency virus Tat protein on nucleolar localization.," *Journal of virology*, vol. 64, pp. 1803–7, apr 1990.
- [482] S. Horke, K. Reumann, M. Schweizer, H. Will, and T. Heise, "Nuclear Trafficking of La Protein Depends on a Newly Identified Nucleolar Localization Signal and the Ability to Bind RNA," *THE JOURNAL OF BIOLOGICAL CHEMISTRY*, vol. 279, no. 25, pp. 26563–26570, 2004.
- [483] C. V. Dang and W. Lee, "Nuclear and nucleolar targeting sequences of c-erb-a, c-myb, n-myc, p53, hsp70, and hiv tat proteins.," *Journal of Biological Chemistry*, vol. 264, no. 30, pp. 18019–18023, 1989.

- [484] J. R. Boyne and A. Whitehouse, "Nucleolar trafficking is essential for nuclear export of intronless herpesvirus mRNA," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, pp. 15190–15195, oct 2006.
- [485] Y. Yang, Y. Chen, C. Zhang, H. Huang, and S. M. Weissman, "Nucleolar localization of hTERT protein is associated with telomerase function," *Experimental Cell Research*, vol. 277, no. 2, pp. 201–209, 2002.
- [486] L. C. Goatley, M. B. Marron, S. C. Jacobs, J. M. Hammond, J. E. Miskin, C. C. Abrams, G. L. Smith, and L. K. Dixon, "Nuclear and nucleolar localization of an African swine fever virus protein, I14L, that is similar to the herpes simplex virus-encoded virulence factor ICP34.5," *Journal of General Virology*, vol. 80, no. 3, pp. 525–535, 1999.
- [487] M. L. Reed, B. K. Dove, R. M. Jackson, R. Collins, G. Brooks, and J. A. Hiscox, "Delineation and modelling of a nucleolar retention signal in the coronavirus nucleocapsid protein," *Traffic*, vol. 7, pp. 833–848, jul 2006.
- [488] H. S. Huang, P. Pozarowski, Y. Gao, Z. Darzynkiewicz, and E. Y. Lee, "Protein phosphatase-1 inhibitor-3 is co-localized to the nucleoli and centrosomes with PP1 $\gamma$ 1 and PP1 $\alpha$ , respectively," *Archives of Biochemistry and Biophysics*, vol. 443, pp. 33–44, nov 2005.
- [489] J. L. Goodier, E. M. Ostertag, K. A. Engleka, M. C. Seleme, and H. H. Kazazian, "A potential role for the nucleus in L1 retrotransposition," *Human Molecular Genetics*, vol. 13, no. 10, pp. 1041–1048, 2004.
- [490] H. J. Thirkettle, J. Girling, A. Y. Warren, I. G. Mills, K. Sahadevan, H. Leung, F. Hamdy, H. C. Whitaker, and D. E. Neal, "Lyric/aeg-1 is targeted to different subcellular compartments by ubiquitinylation and intrinsic nuclear localization signals," *Clinical Cancer Research*, vol. 15, no. 9, pp. 3003–3013, 2009.
- [491] M. A. Lohrum, M. Ashcroft, M. H. Kubbutat, and K. H. Vousden, "Identification of a cryptic nucleolar-localization signal in MDM2," *Nature Cell Biology*, vol. 2, no. 3, pp. 179–181, 2000.
- [492] J. L. Liu, L. F. Lee, Y. Ye, Z. Qian, and H. J. Kung, "Nucleolar and nuclear localization properties of a herpesvirus bZIP oncoprotein, MEQ," *Journal of virology*, vol. 71, no. 4, pp. 3188–3196, 1997.
- [493] M. Tsukahara, H. Suemori, S. Noguchi, Z. S. Ji, and H. Tsunoo, "Novel nucleolar protein, midnolin, is expressed in the mesencephalon during mouse development," *Gene*, vol. 254, pp. 45–55, aug 2000.
- [494] H. Suzuki, Y. Arakawa, M. Ito, S. Saito, N. Takeda, H. Yamada, and J. Horiguchi-Yamada, "MLF1-interacting protein is mainly localized in nucleolus through N-terminal bipartite nuclear localization signal," *Anticancer Research*, vol. 27, no. 3 B, pp. 1423–1430, 2007.
- [495] A. Birbach, S. T. Bailey, S. Ghosh, and J. A. Schmid, "Cytosolic, nuclear and nucleolar localization signals determine subcellular distribution and activity of the NF- $\kappa$ B inducing kinase NIK," *Journal of Cell Science*, vol. 117, pp. 3615–3624, jul 2004.

- [496] N. Ueki, M. Kondo, N. Seki, K. Yano, T. Oda, Y. Masuho, and M. A. Muramatsu, "NOLP: Identification of a novel human nucleolar protein and determination of sequence requirements for its nucleolar localization," *Biochemical and Biophysical Research Communications*, vol. 252, pp. 97–102, nov 1998.
- [497] T. Fujiwara, S. Suzuki, M. Kanno, H. Sugiyama, H. Takahashi, and J. Tanaka, "Mapping a nucleolar targeting sequence of an RNA binding nucleolar protein, Nop25," *Experimental Cell Research*, vol. 312, pp. 1703–1712, jun 2006.
- [498] M. Golding, C. Ruhrberg, J. Sandle, and W. J. Gullick, "Mapping nucleolar and spliceosome localization sequences of neuregulin1- $\beta$ 3," *Experimental Cell Research*, vol. 299, pp. 110–118, sep 2004.
- [499] M. Nagahama, Y. Hara, A. Seki, T. Yamazoe, Y. Kawate, T. Shinohara, K. Hatsuzawa, K. Tani, and M. Tagaya, "NVL2 is a nucleolar AAA-ATPase that interacts with ribosomal protein L5 through its nucleolar localization sequence," *Molecular Biology of the Cell*, vol. 15, pp. 5712–5723, dec 2004.
- [500] G. J. Pellar and P. J. DiMario, "Deletion and site-specific mutagenesis of nucleolin's carboxy GAR domain," *Chromosoma*, vol. 111, no. 7, pp. 461–469, 2003.
- [501] R. Y. L. Tsai and R. D. G. McKay, "A nucleolar mechanism controlling cell proliferation in stem cells and cancer cells," *Genes and Development*, vol. 16, pp. 2991–3003, dec 2002.
- [502] B. C. Valdez, L. Perlaky, D. Henning, Y. Saijo, P.-K. Chan, and H. Busch, "Identification of the nuclear and nucleolar localization signals of the protein p120. interaction with translocation protein b23.," *Journal of Biological Chemistry*, vol. 269, no. 38, pp. 23776–23783, 1994.
- [503] A. Kakuk, E. Friedländer, G. Vereb, D. Lisboa, P. Bagossi, G. Tóth, P. Gergely, and G. Vereb, "Nuclear and nucleolar localization signals and their targeting function in phosphatidylinositol 4-kinase PI4K230," *Experimental Cell Research*, vol. 314, pp. 2376–2388, aug 2008.
- [504] T. S. Kuroda, H. Maita, T. Tabata, T. Taira, H. Kitaura, H. Ariga, and S. M. Iguchi-Ariga, "A novel nucleolar protein, PAPA-1, induces growth arrest as a result of cell cycle arrest at the G1 phase," *Gene*, vol. 340, pp. 83–98, sep 2004.
- [505] M. A. Hahn and D. J. Marsh, "Nucleolar localization of parafibromin is mediated by three nucleolar localization signals," *FEBS Letters*, vol. 581, pp. 5070–5074, oct 2007.
- [506] W. Condemine, Y. Takahashi, M. Le Bras, and H. de Thé, "A nucleolar targeting signal in PML-I addresses PML to nucleolar caps in stressed or senescent cells," *Journal of Cell Science*, vol. 120, pp. 3219–3227, sep 2007.
- [507] R. R. Rowland, R. Kervin, C. Kuckleburg, A. Sperlich, and D. A. Benfield, "The localization of porcine reproductive and respiratory syndrome virus nucleocapsid protein to the nucleolus of infected cells and identification of a potential nucleolar localization signal sequence," *Virus Research*, vol. 64, pp. 1–12, oct 1999.

- [508] J. E. Henderson, N. Amizuka, H. Warshawsky, D. Biasotto, B. M. Lanske, D. Goltzman, and A. C. Karaplis, "Nucleolar localization of parathyroid hormone-related peptide enhances survival of chondrocytes under conditions that promote apoptotic cell death.," *Molecular and Cellular Biology*, vol. 15, pp. 4064–4075, aug 1995.
- [509] L. L. Woo, K. Futami, A. Shimamoto, Y. Furuichi, and K. M. Frank, "The Rothmund-Thomson gene product RECQL4 localizes to the nucleolus in response to oxidative stress," *Experimental Cell Research*, vol. 312, pp. 3443–3457, oct 2006.
- [510] L. A. Stark and M. G. Dunlop, "Nucleolar Sequestration of RelA (p65) Regulates NF- $\kappa$ B-Driven Transcription and Apoptosis," *Molecular and Cellular Biology*, vol. 25, pp. 5985–6004, jul 2005.
- [511] H. Siomi, H. Shida, S. H. Nam, T. Nosaka, M. Maki, and M. Hatanaka, "Sequence requirements for nucleolar localization of human t cell leukemia virus type i px protein, which regulates viral rna processing," *Cell*, vol. 55, no. 2, pp. 197–209, 1988.
- [512] T. Annilo, J. Jelina, I. Pata, and A. Metspalu, "Isolation and characterization of the mouse ribosomal protein S7 gene," *IUBMB Life*, vol. 46, pp. 287–295, oct 1998.
- [513] N. Jarrous, J. S. Wolenski, D. Wesolowski, C. Lee, and S. Altman, "Localization in the nucleolus and coiled bodies of protein subunits of the ribonucleoprotein ribonuclease P," *Journal of Cell Biology*, vol. 146, pp. 559–571, aug 1999.
- [514] K. Viiri, H. Korkeamäki, M. Kukkonen, L. Nieminen, K. Lindfors, P. Peterson, M. Mäki, H. Kainulainen, and O. Lohi, "Sap30l interacts with members of the sin3a corepressor complex and targets sin3a to the nucleolus," *Nucleic acids research*, vol. 34, no. 11, pp. 3288–3298, 2006.
- [515] É. Caron, C. Côté, M. Parisien, F. Major, and C. Perreault, "Identification of two distinct intracellular localization signals in stt3-b," *Archives of biochemistry and biophysics*, vol. 445, no. 1, pp. 108–114, 2006.
- [516] Z. Song and M. Wu, "Identification of a novel nucleolar localization signal and a degradation signal in Survivin-deltaEx3: A potential link between nucleolus and protein degradation," *Oncogene*, vol. 24, pp. 2723–2734, apr 2005.
- [517] S. T. Winokur and R. Shiang, "The Treacher Collins syndrome (TCOF1) gene product, treacle, is targeted to the nucleolus by signals in its C-terminus," *Human Molecular Genetics*, vol. 7, no. 12, 1998.
- [518] H. Pluk, J. Soffner, R. Lührmann, and W. J. van Venrooij, "cDNA Cloning and Characterization of the Human U3 Small Nucleolar Ribonucleoprotein Complex-Associated 55-Kilodalton Protein," *Molecular and Cellular Biology*, vol. 18, pp. 488–498, jan 1998.
- [519] C. von Kobbe and V. A. Bohr, "A nucleolar targeting sequence in the Werner syndrome protein resides within residues 949-1092," *Journal of Cell Science*, vol. 115, pp. 3901–3907, oct 2002.

- [520] I. S. de Melo, M. D. Jimenez-Nuñez, C. Iglesias, A. Campos-Caro, D. Moreno-Sanchez, F. A. Ruiz, and J. Bolívar, “Noa36 protein contains a highly conserved nucleolar localization signal capable of directing functional proteins to the nucleolus, in mammalian cells,” *PLoS One*, vol. 8, no. 3, p. e59065, 2013.
- [521] K. E. Amrein, B. Takacs, M. Stieger, J. Molnos, N. A. Flint, and P. Burn, “Purification and characterization of recombinant human p50csk protein-tyrosine kinase from an escherichia coli expression system overproducing the bacterial chaperones groes and groel,” *Proceedings of the National Academy of Sciences*, vol. 92, no. 4, pp. 1048–1052, 1995.