Uncovering the principles of membrane effector translocation in *Legionella pneumophila*

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat)

vorgelegt von

Silke Malmsheimer

aus Laichingen

Tübingen

2021

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	17.12.2021
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter:	Prof. Samuel Wagner, PhD
2. Berichterstatter:	Prof. Dr. Doron Rapaport

Table of Contents

Abbreviations IV

Zusammenfassung VIII

Summary X

1 Introduction ..... - 1 -

1.1 Secretion systems ..... - 1 -

1.2 Attacking eukaryotic host cells ..... - 2 -

1.2.1 The Type III Secretion system (T3SS) ..... - 2 -

1.2.2 The Type IV secretion system (T4SS) ..... - 3 -

1.3 The Dot/Icm system in *L. pneumophila* ..... - 5 -

1.3.1 Legionella spp. ..... - 5 -

1.3.2 *Legionella*-macrophage interaction and the role of the Dot/Icm system ..... - 7 -

1.3.3 Regulation of the Dot/Icm system within *L. pneumophila* ..... - 7 -

1.3.4 Assembly of the Dot/Icm system and effector translocation ..... - 8 -

1.4 TMD-effectors – a new effector subgroup ..... - 14 -

1.4.1 Why are TMD-effectors important for the pathogenicity of *L. pneumophila* ..... - 15 -

1.4.2 TMD-effector targeting within the pathogen ..... - 16 -

2 Aim of the work ..... - 22 -

3 Materials and Methods ..... - 24 -

3.1 Enzymes, chemicals, media and buffers ..... - 24 -

3.2 Antibodies ..... - 25 -

3.3 Bacterial and eukaryotic strains and their growth conditions ..... - 26 -

3.4 Molecular cloning ..... - 28 -

3.4.1 Gibson cloning ..... - 39 -

3.4.2 QuikChange site-directed mutagenesis ..... - 40 -

3.4.3 Transformation of electro-competent *L. pneumophila* ..... - 40 -

3.4.4 Construction of *L. pneumophila* mutants ..... - 40 -

3.4.5 Allelic exchange ..... - 41 -

3.5 *In vivo* photo-crosslinking ..... - 41 -

3.6 Crude membrane preparation ..... - 42 -

3.7 Membrane fractionation ..... - 43 -

3.8 Urea extraction ..... - 43 -

3.9 SDS-PAGE ..... - 43 -

3.10 BN PAGE ..... - 43 -
3.11 Coomassie staining of proteins ..... - 44 -
3.12 Western blot analysis and immunodetection ..... - 44 -
3.13 Mass spectrometry analysis ..... - 44 -
3.14 Split Nanoluc-based translocation assay ..... - 45 -
3.15 Split Nanoluc-based topology assay ..... - 46 -
3.16 Bioinformatics ..... - 47 -
4 Results ..... - 48 -
4.1 Targeting of TMD-effectors to the Dot/Icm system in *L. pneumophila* ..... - 48 -
4.2 Selection of TMD-effectors for in-depth analysis ..... - 49 -
4.3 Do TMD-effectors behave differently in their translocation in comparison to soluble effectors? (Effector characteristics) ..... - 51 -
4.3.1 Development of a split NanoLuc-based translocation assay ..... - 51 -
4.3.2 TMD-effectors cannot be secreted without their C-terminal secretion signal ..... - 55 -
4.3.3 Translocation of the TMD-effectors LegC2, LegC3, SidF and Ceg4 depends on the T4 chaperones IcmS and IcmW ..... - 56 -
4.4 How is the SRP targeting potential of TMD-effectors (IM protein characteristics) ..... - 57 -
4.4.1 Quantification of TMD-effectors in *L. pneumophila* ..... - 58 -
4.4.2 TMD-effectors of the Dot/Icm system can be found in the bacterial inner membrane ..... - 60 -
4.4.3 Can chaperone binding prevent membrane targeting? ..... - 68 -
4.4.4 Investigation of possible IcmSW binding sites in SidF ..... - 70 -
4.5 Recognition of TMD-effectors by the Dot/Icm system ..... - 73 -
4.5.1 Are TMD-effectors recognized in the cytoplasm or periplasm? ..... - 73 -
4.5.2 Insertion of additional protein domains into the periplasmic loop of SidF negatively affects with its translocation ..... - 78 -
4.6 Unraveling the mechanism of TMD-effector recognition in the cytoplasm after inner membrane insertion ..... - 81 -
4.6.1 TMD-effectors LegC2 and SidF can be identified in complex with the coupling complex of the Dot/Icm system ..... - 81 -
4.6.2 TMD-effectors seem to be recognized by IcmSW and not DotM of the coupling complex ..... - 83 -
4.7 Is a C-terminal signal located in the cytoplasm and the presence of all Dot/Icm components sufficient for translocation of an inner membrane protein? ..... - 85 -
4.8 Which additional substrate-intrinsic properties are important for successful translocation? ..... - 87 -

4.8.1 Which substrate-intrinsic properties mediate successful translocation of the TMD-effector SidF ..... - 90 -

5 Discussion ..... - 94 -

5.1 Targeting of TMD-effectors to the Dot/Icm system in *L. pneumophila* ..... - 94 -

5.2 Recognition of TMD-effectors by the Dot/Icm system ..... - 98 -

5.3 Outlook ..... - 104 -

6 Literature ..... - 105 -

7 Appendix ..... - 118 -

7.1 List of TMD-containing T3SS substrates with ΔG_{app} predictions (window 10-35; length correction ON) and their amino acid sequence ..... - 118 -

7.2 List of ΔG_{app} SRP predictions in TMD-containing T3SS substrates with sliding window 12-17 and length correction OFF ..... - 120 -

7.3 Figure S1: Expression levels of effector proteins in *L. pneumophila* ..... - 122 -

Abbreviations

aa	amino acid(s)
AAA	ATPase associated with diverse cellular activities
ACA	amino-caproic acid
ACES	N-(2-Acetamido)-2-aminoethanesulfonic acid
Arf1	ADP-ribosylation factor 1
ATP	adenosine triphosphate
B. pertussis	*Bordetella pertussis*
BN	blue native
BRET	bioluminescence resonance energy transfer
CagA/F	cytotoxin-associated gene A/F
CBD	chaperone binding domain
Ceg4	coregulated with the effector encoding gene 4
CM	crude membrane(s)
Cm	chloramphenicol
CP	cytosolic protein
Cryo-EM	cryo-electron microscopy
CsrA	carbon storage regulator A
Cya	adenylate cyclase
ddH_2O	double-distilled water
DMEM	Dulbecco's Modified Eagle Medium
Dot/Icm	Defective for organelle trafficking/ intracellular multiplication
DNA	deoxyribonucleic acid
dNTP	deoxy nucleoside triphosphate
DrrA	defects in Rab1 recruitment protein A
Dtr	DNA and replication accessory factor
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetate
ER	endoplasmic reticulum
et al.	and others
FCS	fetal calf serum
Ffh	fifty-four homologue
FRET	fluorescence resonance energy transfer
g	gram
GTP	guanosine-5'-triphosphate
HA	human influenza hemagglutinin
HCl	hydrochloric acid
hpi	hours post infection
IM	inner membrane
IMP	inner membrane proteins
IPTG	isopropyl-beta-D-thiogalactopyranoside
KCl	potassium chloride

KH_2PO_4	potassium dihydrogen phosphate
KOH	potassium hydroxide
l	liter
LegC	*Legionella* eukaryotic-like genes
LetA/S	*Legionella* transmission activator and sensor
Kan	kanamycin
kDa	kilodalton
LB	Luria Bertani
*L*CV	*Legionella* containing vacuole
L. pneumophila	*Legionella pneumophila*
LMNG	lauryl maltose neopentyl glycol
M	molar
MAPK	mitogen-activated protein kinase
mg	milligram
$MgCl_2$	magnesium chloride
ml	milliliter
mM	millimolar
MM	molecular mass
NaCl	sodium chloride
Na_2HPO_4	Disodium phosphate
NaOH	sodium hydroxide
NanoLuc	NanoLuc luciferase
nm	nanometer
OD_{600}	optical density at 600 nm
ODU	optical density per milliliter
OM	outer membrane
o/n	over night
PAGE	polyacrylamide gel electrophoresis
pBpa	para-Benzoyl-phenylalanine
PBS	phosphate-buffered saline
PI(3,4,5)P3	phosphatidyl-inositol-4-phosphate
Pmr	polymyxin resistance
PVDF	polyvinylidene difluoride
PCR	polymerase chain reaction
pEvol	plasmid suppressor tRNA and aminoacyl-tRNA synthetase
pT10 or pT12	plasmid TACO 10 or 12
QC	QuikChange site-directed mutagenesis
Rab1	Ras-related protein Rab-1A
RalF	regulation of ARF protein signal transduction
RNA	ribonucleic acid
RNC	ribosome nascent chain
rpm	revolutions per minute
RT	room temperature
S. Typhimurium	*Salmonella enterica* subspecies *enterica* serovar Typhimurium
SdeA	effector of the SidE family
Sec	secretion

Sid	substrate of Icm/Dot transporter
SDS	sodium dodecyl sulfate
Sn	supernatant
SNAP	soluble N-ethylmaleimide-sensitive-factor attachment protein
SNARE	SNAP receptors
SOB	super optimal broth
SOC	super optimal broth with catabolite repression
spp	species
SRP	signal recognition particle
Strep	streptomycin
T1SS	type I secretion system
T3SS	type III secretion system
T4SS	type IV secretion system
T6SS	type VI secretion system
T9SS	type IX secretion system
TAE	Tris-acetate buffer supplemented with EDTA
Tat	twin arginine translocation
Taq	*Thermus aquaticus*
TBS	Tris-buffered saline
TBS-T	Tris-buffered saline supplemented with Tween 20
TCA	trichloroacetic acid
TEA	Tris-acetate-EDTA
TF	trigger factor
TMD	transmembrane domain
TMS	transmembrane segment
Tra	transfer
Tris	tris(hydroxymethyl)aminomethane
UV	ultraviolet light
Vir	virulence
v/v	volume per volume
wc	whole cells
WipA	IcmW-interacting protein A
WT	wild type
w/v	weight per volume
w/w	weight per weight
µg	microgram
µl	micro liter
µM	micro molar
ΔG	Gibbs free energy
$\Delta G_{app,pred}$	predicted apparent Gibbs free energy
°C	degree Celsius

The common three letter code for amino acids is used in this thesis.

Zusammenfassung

Die Typ III und Typ IV Sekretionssysteme sind für viele Bakterien ein wesentlicher Faktor für ihre Pathogenität. Von der Sekretionsmaschinerie translozierte Effektorproteine werden verwendet, um Wirtszellfunktionen zu unterwandern und anschließend die Wirtszelle zu kolonisieren und eine Überlebensnische zu schaffen. Neben löslichen Effektoren können Effektoren mit Transmembransegmenten erfolgreich durch die Typ III und Typ IV Sekretionssysteme sezerniert werden. Diese werden anschließend in eine Wirtszellmembran inseriert, um dort ihre Funktion auszuüben. Effektorproteine müssen vor der Translokation zum jeweiligen Sekretionssystem rekrutiert werden. Im Allgemeinen beginnt dieser Vorgang im bakteriellen Zytoplasma, wo die Proteinbiosynthese stattfindet, und hängt von verschiedenen Signalen ab, die sich innerhalb des Proteins befinden. Effektoren des Typ III Sekretionssystems enthalten ein N-terminales Sekretionssignal, gefolgt von einer Chaperon-Bindungsdomäne. Im Gegensatz dazu liegt das Translokationssignal für Effektoren des Typ IV Sekretionssystems am C-terminalen Ende des Proteins. Darüber hinaus enthalten Transmembraneffektoren beider Sekretionssysteme hydrophobe Segmente, die für ihre korrekte Lokalisierung innerhalb der Wirtszelle essentiell sind. Das Vorhandensein dieser beiden inkompatiblen Signale innerhalb desselben Proteins stellt einen möglichen Erkennungskonflikt dar, da Transmembransegmente vom „*Signal Recognition Particle*“ (SRP) erkannt werden und zu einer Insertion in die innere Membran führen können. Für Transmembraneffektoren des Typ III Sekretionssystems konnte gezeigt werden, dass Insertion in die innere Membran durch eine ausgewogene Hydrophobizität des Transmembransegments (passive Vermeidung) sowie durch die Abschirmung der Transmembransegmente durch Chaperone (aktive Vermeidung) vermieden wird. Im Gegensatz dazu zeigen Vorhersagemodelle, dass einige Transmembraneffektoren des Dot/Icm-Systems in *L. pneumophila* ein ausreichend hydrophobes Signal besitzen, um durch SRP zur bakteriellen Innenmembran rekrutiert zu werden.

Das Ziel dieser Studie war, zu untersuchen, ob Transmembraneffektoren des Typ IV Sekretionssystems ähnlich wie Effektoren des Typ III Sekretionssystems SRP-Erkennung vermeiden können oder einen hypothetischen zweistufigen Sekretionsweg mit einem Intermediat in der bakteriellen Innenmembran verwenden.

Ich konnte zeigen, dass Transmembraneffektoren in *L. pneumophila* nicht nur in Vorhersagemodellen von SRP erkannt werden, sondern tatsächlich in der bakteriellen Innenmembran gefunden werden können. Die gleichen Beobachtungen konnten gemacht werden, wenn die Typ IV Chaperone IcmSW überexprimiert wurden. Dies deutet darauf hin, dass es keine aktive Vermeidung der SRP-Erkennung und keine direkte Rekrutierung zur Sekretionsmaschinerie durch Chaperon-Bindung gibt. Die Untersuchung der Membrantopologie von Transmembraneffektoren zeigte, dass diese mit einer N_{in}-

C_{in}-Topologie in der bakteriellen Innenmembran verankert sind und sich nur kleine Abschnitte aus wenigen Aminosäuren im Periplasma befinden. Außerdem konnte ich zeigen, dass, ähnlich wie bei löslichen Effektoren, das im Zytoplasma lokalisierte C-terminale Translokationssignal und mögliche interne Signale sowie die Anwesenheit der Chaperone IcmSW notwendig für die erfolgreiche Translokation von Transmembraneffektoren in die Wirtszelle sind. Basierend auf diesen Ergebnissen postuliere ich, dass Transmembraneffektoren in *L. pneumophila* einem zweistufigen Sekretionsweg mit einem Innenmembranintermediat als ersten Schritt folgen. Sobald die Effektoren in der bakteriellen Innenmembran verankert sind, werden diese von IcmSW, die an den „*coupling complex*“ des Dot/Icm-Systems gebunden sind, als Substrate des Dot/Icm-Systems erkannt. Dies führt dazu, dass sie zur zytoplasmatischen Seite hin extrahiert werden, bevor sie anschließend in die Wirtszelle transloziert werden.

Summary

The type III and type IV secretion systems are a major factor of pathogenicity for many bacteria. Effector proteins translocated by the secretion machinery are used to hijack and subvert host cell functions in order to colonize and live within the host cell. Next to soluble effectors, transmembrane domain effectors can successfully be secreted by the type III and type IV secretion system and carry out their function within a host cell membrane. Before being translocated, effector proteins must be targeted to the respective secretion system. Generally, protein targeting starts in the bacterial cytoplasm where protein biosynthesis takes place, and depends on various signals residing within the protein. Type III effectors harbor a N-terminal secretion signal followed by a chaperone binding domain. In contrast, the translocation signal for effectors of the T4SS resides at the C-terminal end of the protein. Additionally, TMD-effectors of both secretion systems contain hydrophobic segments which are essential for their proper localization within the host cell. The presence of these two incompatible signals within the same protein poses a possible targeting conflict as transmembrane segments can be recognized by the Signal Recognition Particle (SRP) and result in subsequent inner membrane insertion. For transmembrane domain effectors of the T3SS it was shown that inner membrane (mis-)targeting is avoided by a balanced hydrophobicity of the TMS (passive avoidance) as well as protection by chaperone binding (active avoidance). In contrast, some transmembrane domain effectors of the Dot/Icm system in *L. pneumophila* are predicted to possess a sufficiently hydrophobic signal for targeting to the bacterial inner membrane by SRP. The aim of this study was to investigate if transmembrane domain effectors of the type IV secretion system can, similarly to effectors of the type III secretion system, avoid SRP targeting or uses a hypothetical two-step secretion pathway through the Dot/Icm machinery with an inner membrane intermediate.

Using membrane fractionation by a sucrose gradient centrifugation protocol as well as urea extraction, I could show that transmembrane domain effectors in *L. pneumophila* can indeed be found properly integrated in the bacterial inner membrane. The same results were obtained when the type IV chaperones IcmSW were overexpressed suggesting that there is no active avoidance of SRP targeting and no direct delivery to the secretion machinery by chaperone binding. Investigation of the membrane topology of transmembrane domain effectors showed that transmembrane domain effectors are “anchored” in the bacterial inner membrane with a N_{in}-C_{in} topology and only small loops of few amino acids located in the periplasm. Furthermore, I could show that, similarly to soluble effectors, the C-terminal translocation signal located in the cytoplasm and possible internal signals as well as the presence of the chaperones IcmSW are crucial for the successful translocation of transmembrane domain effectors into host cells. Based on these results, I propose that transmembrane domain effectors in *L. pneumophila* follow a two-step secretion pathway with SRP

targeting as the first step. Once “anchored” in the bacterial inner membrane, transmembrane domain effectors are recognized as substrates of the Dot/Icm system by IcmSW, resulting in their extraction towards the cytoplasmic side before being translocated into host cells.

1 Introduction

“All living beings, not just animals, but plants and microorganisms, perceive. To survive, an organic being must perceive - it must seek, or at least recognize, food and avoid environmental danger” (Lynn Margulis). In the course of evolution, microorganisms, however, have not only optimized their ways of perception but also successfully developed and adapted various molecular mechanisms in order to benefit from or even control their environment and to ensure their survival (Green and Mecsas, 2016). In bacteria, one example of such a survival strategy are secretion systems.

1.1 Secretion systems

Secretion systems are large, specialized macromolecular machines which span one or both bacterial membranes in order to transport substrates such as small molecules, proteins, DNA or DNA-protein complexes from the bacterial cell to the extracellular space or into another bacterial or eukaryotic cell. Thereby, bacteria can compete for space, nutrients and ecological niches as well as ensure their survival under stress conditions (Pena *et al.*, 2019). At least nine different secretion systems have been identified in gram-negative bacteria which can be divided into two groups based on their ability to translocate substrates out of the cell in one or two steps (Fig. 1). The type II, V and VIII secretion systems use the Sec or the twin-arginine Translocation (Tat) machinery to execute the first step of their secretion mechanism. In contrast, the type I, III, IV, VI, VII and IX secretion systems (T1SS, T3SS, T4SS, T6SS, T9SS) span the bacterial inner and outer membranes as well as the periplasmic space and secrete substrate proteins in one step from the cytosol to the outside.

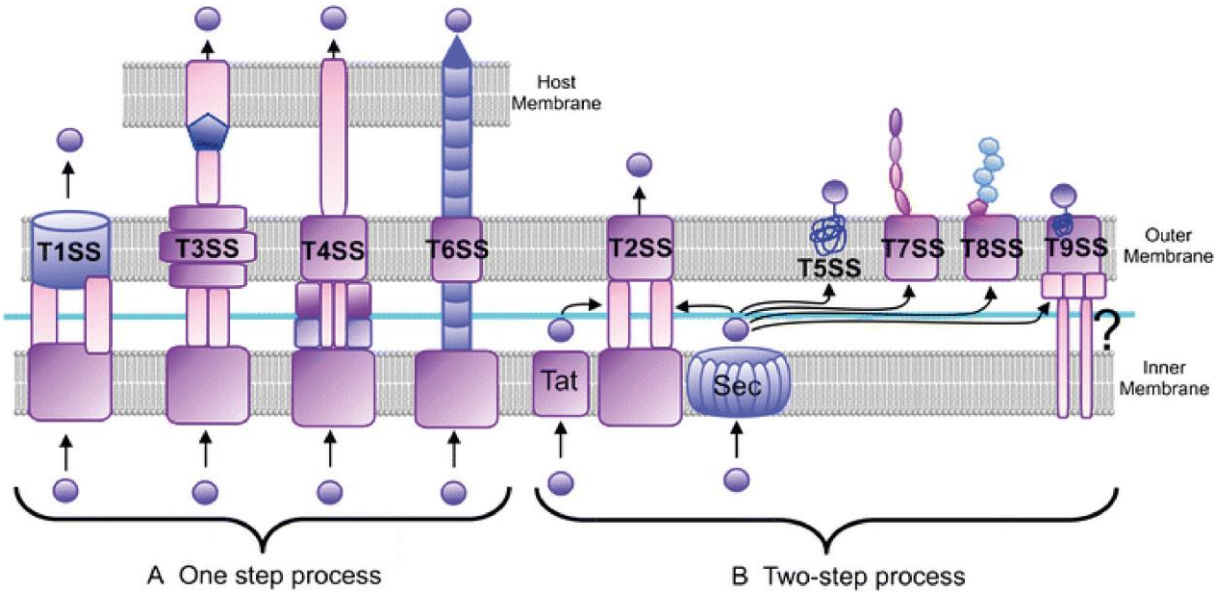


Fig. 1: Secretion systems of gram-negative bacteria
A) Substrate secretion through the type I, III, IV and VI secretion follows a one-step secretion mechanism by which the effector molecule is translocated directly into the extracellular milieu or the cytosol of a target cell. B) The type II, V, VII, VIII, IX secretion systems depend on the help of the Sec or Tat (twin arginine transportation) pathways for translocation of the effector molecules across the bacterial inner membrane. (Adopted from Bocian-Ostrzycka *et al.*, 2017)

1.2 Attacking eukaryotic host cells

T3SS, T4SS and T6SS are well known to not only export substrates from the bacterial cytoplasm in one step across the entire cell envelope but additionally translocate their substrates directly into other bacterial or eukaryotic host cells (Galan and Waksman, 2018) and thus play a major factor in bacterial virulence. The T6SS is most commonly used for protein secretion into neighboring bacteria, whereas the T3SS and some T4SS translocate effector proteins into host cells in order to remodel the host cell architecture and establish a replicative niche. With this, T3SS and T4SS play a central role in the pathogen-host cell interaction (Costa *et al.*, 2015; Galan and Waksman, 2018).

1.2.1 The Type III Secretion system (T3SS)

T3SS-secreted effector proteins are used by gram-negative bacteria such as *Salmonella spp., Yersinia spp.,* enteropathogenic *E. coli, Shigella spp.* and *Pseudomonas spp*. (Büttner, 2012) in order to invade and replicate within host cells, leading to diarrhea, pneumonia or systemic infections (Coburn, Sekirov and Finlay, 2007). The T3SS is evolutionarily and structurally related to bacterial flagella (Abby and

Rocha, 2012) and consists of up to 20 different proteins spanning both bacterial membranes. Effector proteins harbor a N-terminal 20-25 amino acid long secretion signal. Its role in the targeting mechanism is however still unclear (Krampen *et al.*, 2018). The secretion signal is often followed by a 20-50 residue-long chaperone binding domain (CBD). Cognate T3SS chaperones, which are often encoded adjacent to their T3SS substrate, bind the CBD and target the effector protein to the sorting platform of the injectosome (Lara-Tejero *et al.*, 2011). The effector is released from chaperone binding and unfolded with the help of T3SS's ATPase (Akeda and Galán, 2005; Yoshida *et al.*, 2014).

The T3SS uses substrate secretion in order to assemble the needle of the system until it reaches a defined length. A second class of T3SS substrates are translocator proteins forming a pore in the host cell membrane after host cell contact. Finally, effector proteins are translocated into the cytosol of the host cell manipulating different processes, such as actin polymerization, resulting in the uptake of the bacterium and bacterial colonization (Notti and Stebbins, 2016).

1.2.2 The Type IV secretion system (T4SS)

T4SSs are the most diverse of all known secretion systems (Wallden, Rivera-Calzada and Waksman, 2010) and play a key role in bacterial virulence of various animal and plant pathogens. They can be found in gram-negative and gram-positive bacteria as well as archaea and encompass mainly conjugation machines for interbacterial DNA delivery as well as effector translocator systems for injection of proteins into eukaryotic target cells in a cell-to-cell contact dependent manner.

Conjugation systems play an important role in the exchange and spreading of antibiotic resistance genes (Huddleston, 2014; Cabezón, de la Cruz and Arechaga, 2017; Koraimann, 2018) and thus are of great medical concern. A few T4SS export DNA or proteins to the extracellular milieu e.g. from *Neisseria gonorrhoeae* or import exogenous DNA e.g. in *Helicobacter pylori*. Furthermore, it was recently described that T4SSs which usually translocate effector proteins can also be used for the transfer of DNA, peptidoglycan or other macromolecules (Grohmann *et al.*, 2018; Bleves, Llosa and Haven, 2021).

The delivery of effector proteins into host cells by a T4SS plays a key role in the infectious cycles of many bacteria including *Xanthomonas, Bartonella, Legionella and Coxiella* (Cascales and Christie, 2003; Christie, 2016; Grohmann *et al.*, 2018). Effector proteins can have diverse activities, structures and sizes. Once inside the host cell, effector proteins hijack host cell machineries and limit the immune response, thus providing an environment which the bacteria can colonize. Interestingly, the number of translocated effector proteins is highly variable between different T4SS. The oncoprotein CagA is the only effector secreted by the T4SS in *Helicobacter* pylori (Hatakeyama, 2017). In contrast,

Legionella pneumophila translocates more than 300 effector proteins during the course of their infection (Hubber and Roy, 2010; Nagai and Kubori, 2011; Gomez-Valero *et al.*, 2019).

In gram-negative bacteria, two subgroups of T4SS can be distinguished: The type IVA secretion system includes *E. coli* conjugation apparatuses and the VirB/D system of *Agrobacterium tumefaciens, Xanthomonas spp, Bartonella spp.* (Grohmann *et al.*, 2018) whereas *Legionella spp.* and *Coxiella spp*. express the type IVB or Dot/Icm (Defective for organelle trafficking/ intracellular multiplication) system (Nagai and Kubori, 2011; Grohmann *et al.*, 2018) which will be discussed in greater detail in the next chapters of this thesis.
Similar to its functional diversity, T4SSs show also great variability in their structure. The VirB/D system consists of only 12 different proteins including three ATPases, the inner membrane complex (IMC) and the outer membrane complex (OMC) whereas there are at least 27 structural proteins known for the Dot/Icm system in *Legionella pneumophila*. One crucial component is the coupling protein or coupling complex.

1.2.2.1 Coupling components of the T4SS

Most conjugation systems as well as effector translocation systems depend in their function on a type IV (T4) coupling protein (T4CP) (Alvarez-Martinez and Christie, 2009; Zechner, Lang and Schildbach, 2012) which "couples" the DNA or protein substrate to the transport machinery. The T4CP mostly consists of a N-terminal transmembrane domain and a large C-terminal region with a nucleotide-binding domain, an all-α-domain and in some cases a C-terminal extension which can interact with other proteins (Gomis-Rüth *et al.*, 2001; Alvarez-Martinez and Christie, 2009). It serves as a recognition platform for either DNA or protein effector substrates which are recognized by means of their mostly unstructured C-terminal or internally located secretion signals (Nagai *et al.*, 2005; Vergunst *et al.*, 2005). In many cases docking of substrates requires additional adaptor proteins or chaperones such as the DNA and replication (Dtr) accessory factor in DNA conjugation or the VirE1 chaperone which is co-expressed to its substrate VirE2 (Zhao *et al.*, 2001). As the structure of the VirE1/E2 complex revealed, binding of the chaperone leads to the exposure of the disordered C-terminus of VirE2 (Dym *et al.*, 2008). Similar behavior was observed for the effector protein CagA and its chaperone CagF in *H. pylori* (Couturier *et al.*, 2006; Pattis *et al.*, 2007). In *Legionella pneumophila* the type IV coupling protein DotL was identified based on its sequence similarity to other AAA+ ATPases. Together with DotN and DotM as well as the chaperones or adaptor proteins IcmS, IcmW and LvgA and the just recently identified proteins DotY and DotZ, they form a subcomplex named as the Type IV coupling complex (T4CC) of the

Dot/Icm system. Further details of the T4CC in *Legionella pneumophila* will be discussed in subsequent chapters.

Interestingly, T4SS without a T4CP or T4CC seem to translocate their substrate proteins in a two-step secretion mechanism. The pertussis toxin in *B. pertussis* for instance is translocated across the bacterial inner membrane using the Sec system. After proper folding and oligomerization, the toxin is translocated across the outer membrane in a T4SS-dependent manner (Burns *et al.*, 2003; Locht, Coutte and Mielcarek, 2011). Although the exact mechanism of such a two-step mechanism still remains to be elucidated, it is likely to be similar to the T2SS (Ayers *et al.*, 2010) as all other A/B toxins from the A/B toxin superfamily are exported via T2SSs. One advantage of this route rather than being secreted like other T4SS substrates would be that after oligomerization the huge pertussis toxin does not need to go through the narrow inner membrane channel of the T4SS. The VirB T4SS of *Brucella* spp. also lacks a T4CP. Effector proteins use their N-terminal signal to be translocated by the Sec-machinery into the periplasmic space where they engage with the T4 channel using their C-terminal signal and are subsequently transported across the outer membrane of the bacterium (de Jong *et al.*, 2008; De Barsy *et al.*, 2011; Marchesini *et al.*, 2011).

1.3 The Dot/Icm system in *L. pneumophila*

The focus of this work will be on the T4BSS in *Legionella pneumophila* also called Dot/Icm system. In the following sections the function of the Dot/Icm system and its role in the pathogenicity of *Legionella* as well as the structural details will be explained in more detail. It is, however, worth noting, that in *L. pneumophila* two additional T4SS are expressed: The *lvh* system and the *tra* system which show homology to Tra proteins of the *E. coli* F plasmid (Brassinga *et al.*, 2003).

1.3.1 Legionella spp.

Legionella spp. are gram-negative, flagellated bacteria which can be found in freshwater ecosystems and in human-made aquatic environments such as showers, air conditioning systems, water fountains or cooling towers. They can grow in natural biofilms or replicate intracellularly in free-living amoeba, their natural host.

All *Legionella* strains sequenced so far employ a highly conserved Dot/Icm system (Burstein *et al.*, 2016; Gomez-Valero *et al.*, 2019) with which they are capable to subvert host cell functions after being

engulfed. To this end, they build a replication niche, the so-called *Legionella* containing vacuole (*L*CV), in which they are protected from hostile environments and can live and replicate (Hubber and Roy, 2010). The same mechanisms are used to infect human alveolar macrophages (Fig. 2).

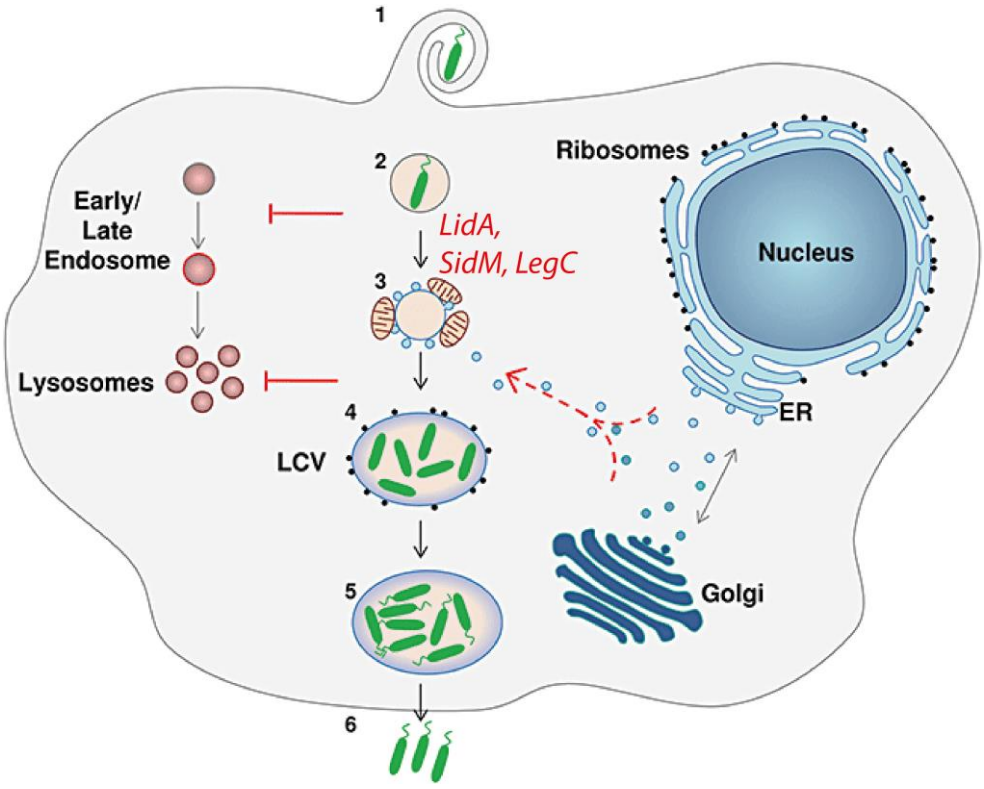


Fig. 2: Intracellular life cycle of *Legionella pneumophila* *(a*dapted from Franco, Shuman and Charpentier, 2009)**.** **1)** *Legionella pneumophila* is phagocyzed by an eukaryotic host cell (amoeba or alveolar macrophages) More than 300 effector proteins are translocated into the host cell cytoplasm in order to establish a *Legionella*-containing vacuole and subvert various host cell function. **2)** Fusion of *L*CV with endosomes is avoided. **3)** The *L*CV interacts with mitochondria. Moreover, ER-derived vesicles trafficking to the Golgi are recruited and fused with the *L*CV. **4)** A rough ER-like replicative vacuole surrounded with ribosomes is formed. **5)** Replication of bacteria takes place. **6)** The lack of nutrients leads to differentiation of the bacterium to its flagellated transmissive form. Bacteria escape the host cell.

Although 65 *Legionella* species have been identified to date, approximately 90 % of all clinical infections are caused by *L. pneumophila*. *Legionella pneumophila* strain Philadelphia-1 was discovered in 1976 when several American Legionnaires attending a meeting in Philadelphia suffered from an atypical pneumonia (Fraser *et al.*, 1977; McDade *et al.*, 1977), thus called Legionnaires' disease. Legionnaires' disease has an incubation time of 2-14 days and symptoms such as headache, asthenia, anorexia and myalgia (Edwards, Fry and Harrison, 2008; Cunha, Burillo and Bouza, 2016). So far, no transmission from human to human has been reported (Correia *et al.*, 2016). The bacteria are inhaled

with contaminated aerosols and start their infection cycle with their attachment to macrophages present in human lungs (McDade *et al.*, 1977; Payne and Horwitz, 1987). After replication within the *L*CV in the host cell, they switch to their "transmissive" form and are released from the macrophage (Fig. 2).

1.3.2 *Legionella*-macrophage interaction and the role of the Dot/Icm system

Already in 1994, it was shown that after their phagocytosis, *L. pneumophila* can only survive intracellularly with a functional Dot/Icm system (Berger, Merriam and Isberg, 1994; Franco, Shuman and Charpentier, 2009). In order to evade phagosome-lysosome fusion and subsequent degradation, various cellular processes of the host have to be modified. For instance, modification of phospholipids in the *L*CV membrane bilayer leads to a substantially different appearance of the *L*CV in comparison to the late endosome and lysosomes. In order to achieve this, effectors bind to phosphatidylinositol molecules or modify their phosphorylation state (Weber and Menko, 2006). SidM, for instance, binds to phosphatidylinositol-4-phosphate, whereas SidF is a phosphatidylinositol-2 phosphatase (Toulabi *et al.*, 2013). Additionally, ribosomes are recruited to conceal the *L*CV as rough endoplasmic reticulum (ER) (Tilney *et al.*, 2001).

Furthermore, a substantial subversion of host vesicle trafficking takes place in order to ensure *L*CV biogenesis. To this end, the regulation of several small host GTPases is crucial. The Dot/Icm effectors SidM and RalF, for example, acts as guanine exchange factors for the small GTPases Rab1 and Arf1, respectively, and recruit them to the *L*CV immediately after *L. pneumophila* has entered the host cell. By this they manipulate vesicle trafficking from the endoplasmic reticulum (ER) to the Golgi apparatus (Schoebel *et al.*, 2009; Neunuebel *et al.*, 2011).

To ensure sufficient energy supply mitochondria are recruited to the *L*CV (Chong *et al.*, 2009). Moreover, the actin pathway is modulated (Franco, Shohdy and Shuman, 2012) and effector proteins interfere with host protein synthesis (Belyi *et al.*, 2008).

1.3.3 Regulation of the Dot/Icm system within *L. pneumophila*

As expressing and maintaining a highly complex nanomachine like the Dot/Icm system comes with high energetic costs for the bacterium, *L. pneumophila* developed a bi-phasic life cycle with major transcriptomic changes in order to focus only on the imminent task. Nutrient-rich conditions, for example within host cells, induces the differentiation of *L. pneumophila* into its replicative form,

whereas in nutrient-limiting environments *L. pneumophila* enters the transmissive phase and with this alters the expression of up to 50 % of its genes including the upregulation of protein expression associated with host entry, virulence and survival including the Dot/Icm machinery and effector proteins.

The switch between these two phases depends on various regulatory systems. Starvation of amino acids, for example, leads to the activation of the two-component system LetA/S as well as the regulator of general stress response RpoS (Dalebroux, Edwards and Swanson, 2009; Dalebroux *et al.*, 2010) by the alarmone nucleotide guanosine pentaphosphate (pppGpp) which subsequently results in binding of the global repressor CsrA (Tiaden *et al.*, 2007, 2008; Rasis and Segal, 2009; Sahr *et al.*, 2009, 2017; Kessler *et al.*, 2013). CsrA regulates the expression of transmissive traits involving motility and virulence factors including more than 40 effector proteins. At the same time replicative traits are repressed (Rasis and Segal, 2009; Sahr *et al.*, 2009, 2012). Moreover, several other two-component systems like PmrA/B and CpxR/A are involved in the post-transcriptional regulation of virulence factors (Hammer, Tateda and Swanson, 2002; Zusman *et al.*, 2007; Altman and Segal, 2008).

Although modulation of expression of *dot/icm* genes as well as effector proteins has been observed, the Dot/Icm system seems to be present at different phases of the *Legionella* life cycle (Aurass *et al.*, 2016; Park *et al.*, 2020). This suggests a constant expression of *dot/icm* encoding genes. Moreover, different effector proteins are expressed in different growth phases of *L. pneumophila*. They are, however, not permanently secreted (Nagai *et al.*, 2005; Chen *et al.*, 2007), suggesting a post-translational control of the secretion complex. Although it is generally acknowledged that host cell contact plays an important role, the exact mechanism behind possible regulations of the Dot/Icm system has yet to be elucidated (Gonzalez-Rivera, Bhatty and Christie, 2016).

1.3.4 Assembly of the Dot/Icm system and effector translocation

Evolutionarily and structurally related to bacterial conjugation machines, the Dot/Icm system of *L. pneumophila* (Segal, Purcell and Shuman, 1998; Vogel *et al.*, 1998), comprises about 27 different structural components (Ghosal *et al.*, 2018) encoded in two separate regions on the chromosome. Although the function of some Dot/Icm proteins is still unknown, several studies showed that single deletions of most of the Dot/Icm components results in a decrease or loss of activity of the Dot/Icm system (Segal, Purcell and Shuman, 1998; Segal and Shuman, 1999; Zuckman, Hung and Roy, 1999; Coers *et al.*, 2000; Luo and Isberg, 2004; Kubori *et al.*, 2014). It is however important to mention that this can vary between different host cells (Segal and Shuman, 1999).

In contrast to the T4ASS, the architecture of the T4BSS was only poorly understood for many years. However, just recently exciting structural insights were obtained (Ghosal *et al.*, 2017, 2019; Kwak *et al.*, 2017; Jeong *et al.*, 2018; Meir *et al.*, 2018). By using fluorescence microscopy, it was shown that Dot/Icm systems in *L. pneumophila* are primarily located at the poles of the bacterium and that polar localization is important for the intracellular survival of the bacterium (Ghosal *et al.*, 2019). The two Dot/Icm proteins DotU and IcmF are suggested to be responsible for recruiting the other components of the system to the bacterial poles for subsequent assembly (Jeong *et al.*, 2018). Although so far there is not much known about the assembly process of the Dot/Icm system, Park *et al.* could give a first idea of the assembly mechanism in 2020 by identifying five distinct subassembly intermediates by high-throughput cryo-electron tomography. With this it was proposed that first, the core complex also called outer membrane core complex (OMCC) is assembled, suggesting an "outside-in" mechanism (Fig. 3).

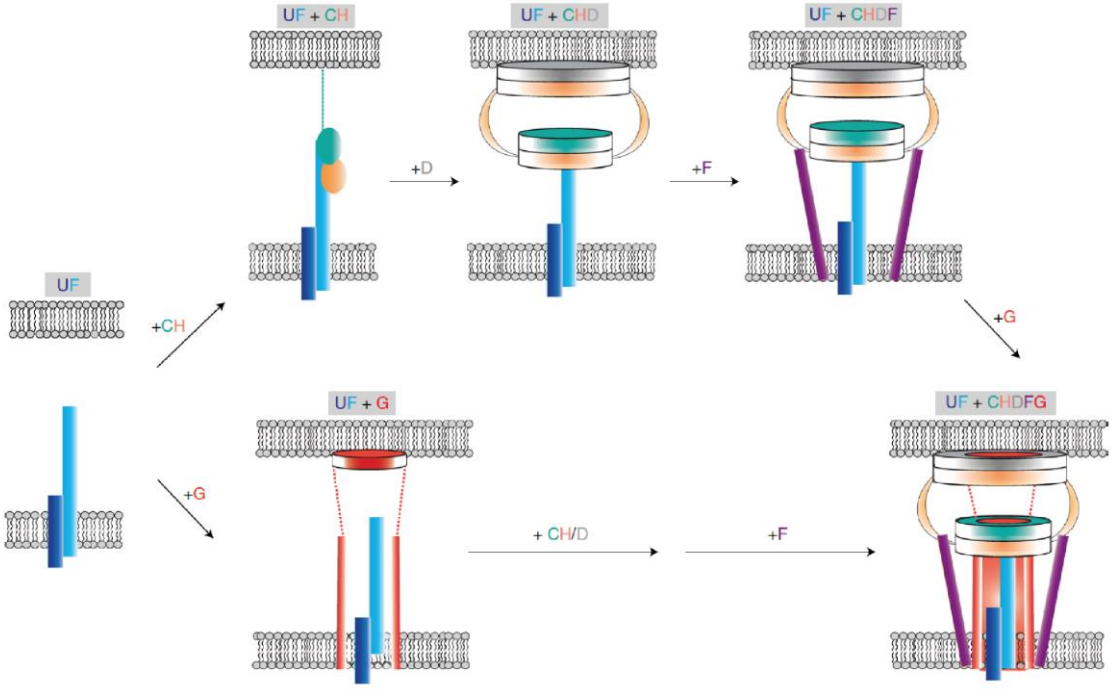


Fig. 3: Polar targeting and assembly of the Dot/Icm core complex in *L. pneumophila* (adopted from Ghosal 2019).

DotU and IcmF recruit other components of the Dot/Icm system to the poles of *L. pneumophila*. The Dot/Icm system is subsequently assembled by an outside-in mechanism (upper panel). In an alternative pathway (lower panel) DotG is recruited in a first step of assembly.

U DotU (dark blue), F IcmF (light blue), G DotG (red), H DotH (orange), C DotC (green), D DotD (grey), F DotF (violet)

1.3.4.1 The Core complex of the Dot/Icm system

The Dot/Icm OMCC consists of at least five proteins (DotG, DotH, DotF, DotD and DotC) forming a ring- or barrel-shaped structure with two distinct layers, one larger layer just below the outer membrane and a smaller one in the periplasm (Chetrit *et al.*, 2018). In the assembly process the lipoprotein DotC is anchored in the outer membrane. Next, the outer membrane protein DotH is recruited. Interestingly, DotU and IcmF seem to be necessary for this step as their absence significantly affects DotH localization (Jeong *et al.*, 2018). Upon proper targeting of DotC and DotH, another lipoprotein, DotD, is recruited followed by DotG and DotF. DotD seems to play an important role in mediating the outer membrane association of DotH resulting in a pore formation of DotH in the outer membrane. DotF and DotG are inner membrane proteins with a single transmembrane helix in their N-terminal region, but can be found in inner membrane as well as outer membrane fractions (Vincent *et al.*, 2006). DotG was shown to be a central channel subunit linking the OMCC to the cytoplasmic complex (Chetrit *et al.*, 2018). DotF has a large periplasmic domain making up the "wings" surrounding and probably interacting with DotG and DotH (Vincent *et al.*, 2006; Ghosal *et al.*, 2018). As polar localization of DotG can take place independently of DotCDFH to some extent, there might be an alternative way of starting assembly with the inner membrane integration of DotG (Fig. 3, lower panel).

Together these five proteins form the central channel of the Dot/Icm system that extends to the inner membrane complex (IMC). However, additional components were discovered interacting with this subcomplex of the Dot/Icm system. DotK for instance, another lipoprotein, interacts with DotD, DotH and the peptidoglycan and by this probably stabilizes the core complex in the cell envelope (Ghosal *et al.*, 2018), although its absence has no big influence on the intracellular growth of *Legionella* (Segal and Shuman, 1999).

1.3.4.2 The IMC and the ATPases of the Dot/Icm system

Although next to DotG and DotF there are a lot of Dot/Icm components located in the bacterial inner membrane, not much is known about the assembly of the IMC. Kuroda *et al.* showed that DotI and its partial paralog DotJ form a stable heterocomplex, which, however, seems not to form a stable association with the core complex (Kuroda *et al.*, 2015).

After proper assembly of the OMCC and IMC, the ATPases DotO and DotB are recruited to the Dot/Icm system. DotB is homologous to a large family of nucleotide-binding proteins and has well-known Walker motifs for nucleotide binding (Purcell and Shuman, 1998). In 2018, Chetrit *et al.* could show that DotB associates with the Dot/Icm system and forms a cytoplasmic ATPase subcomplex which

consists of two “hexamers of DotB and DotO stacked on one another and centrally positioned at the base of the inner membrane-spanning channel” (Chetrit *et al.*, 2018). Their model suggests that an already assembled OMCC as well as IMC are required for recruitment of the ATP-bound form of DotO to the cell poles. Upon activating signals, DotB is recruited by DotO, which in turn leads to a ~20° clockwise rotation of the entire cytoplasmic complex (Park *et al.*, 2020) decreasing the distance between outer and inner membrane. Moreover, a V-shaped cylinder domain above the DotO complex in the inner membrane appears to open up upon DotB binding. This suggests that prior to DotB binding, the Dot/Icm machine is in an inactive state, whereas DotB binding triggers an early step of activation of the system.

1.3.4.3 The type IV coupling complex of the Dot/Icm system

Another essential subcomplex of the Dot/Icm system is its coupling complex. Although the atomic structure of the fully assembled T4CC in *L. pneumophila* was solved recently (Meir *et al.*, 2020), its assembly and position relative to the already described subcomplexes has not yet been investigated.

In the T4CC of the Dot/Icm system DotL, a member of the VirD4 family, plays the central role. Like other T4CP DotL is a membrane-embedded AAA+ ATPase with a long C-terminal extension with which it interacts with the soluble protein DotN, the chaperones or adaptor proteins IcmS, IcmW and LvgA, the inner membrane protein DotM as well as with the newly identified proteins DotY and DotZ (Kwak 2018, Meir 2020). Based on the homology of DotL to the AAA+ ATPase VirD4 of the T4ASS Meir *et al.* proposed a model of a hexamer of pentameric units (Meir *et al.*, 2020). The model shows a cavity between DotM, DotZ and DotN, which is supposed to be used by a subgroup of effector proteins harboring a glutamate-rich C-terminal signal. Details of effector recognition and loading are described in chapter 1.3.4.7.

It is important to note that components of the T4CC depend on each other, especially in the stationary growth phase of *L. pneumophila*. The presence of DotM, IcmS or IcmW seems to be required for the stability of DotL. Moreover, in strains lacking DotL no DotM could be detected and IcmS and IcmW require each other for their stability (Vincent *et al.*, 2012). Interestingly, deletion of the chaperones results in a partial intracellular growth defect, but no inhibition of phagosome-lysosome fusion (Coers *et al.*, 2000).

1.3.4.4 Chaperones of the Dot/Icm system

IcmS, IcmW and LvgA were identified as chaperones of the Dot/Icm system given that they are small acidic proteins and by this resemble the physical properties of type III (T3) chaperones (Segal and Shuman, 1997; Zuckman, Hung and Roy, 1999; Coers *et al.*, 2000; Xu *et al.*, 2017). Several studies showed that IcmS-IcmW or IcmS-LvgA can bind to substrates of the Dot/Icm system as heterodimers and thereby facilitate their translocation (Bardill, Miller and Vogel, 2005; Ninio *et al.*, 2005; Vincent and Vogel, 2006; Cambronne and Roy, 2007). Although the binding of IcmSW and LvgA to the C-terminal extension of DotL as well as to different binding sites within several T4 effector proteins such as WipA, SdeA, SidH and SidG was shown (Bardill, Miller and Vogel, 2005; Ninio *et al.*, 2005; Cambronne and Roy, 2007; Lifshitz *et al.*, 2013) the exact function of the Dot/Icm chaperones is still unknown. As binding of IcmSW to SidG does not change its solubility, stability or localization within the bacterial cell, it is assumed that one of their functions is to keep effector proteins in a translocation-competent state (Bennett and Hughes, 2000; Parsot, Hamiaux and Page, 2003). Furthermore, Cambronne and Roy could show that IcmSW binds to a region distinct from the C-terminal translocation signal of SidG which probably leads to a conformational change and the recognition of the signal (Cambronne and Roy, 2007). It is, however, unclear at which step in the targeting and recognition process the chaperones interact with effector proteins. So far, there is no indication that IcmSW bind effector proteins during translation and thus are involved in effector targeting. In contrast, a significant amount of IcmS and IcmW was found associated with membrane fractions (Vincent *et al.*, 2012). Moreover, for the soluble effector SidG it was shown that deletion of the identified IcmSW binding site within the effector did not abrogate effector translocation which implies that IcmSW are likely not involved in the targeting process.

1.3.4.5 Effectors of the Dot/Icm system

Although the Dot/Icm machinery itself is highly conserved throughout different *Legionella* species, they demonstrate a great variety in the size and composition of their effector repertoires. A comparative genome sequence analysis of 80 *Legionella* strains revealed that there is no common set of effectors which exists in strains robustly replicating in human macrophages (Gomez-Valero *et al.*, 2019). Overall, 18,000 effector proteins encoded by *Legionella* spp. were found of which only 8 were conserved in all analyzed *Legionella* genomes.

L. pneumophila expresses a large repertoire of more than 300 effectors of which many are not yet characterized in detail (Bi *et al.*, 2013). Although they have a wide variety of functions such as

prevention of phagosome-lysosome fusion, creation of the *L*CV and evasion of the host immune system (Ensminger and Isberg, 2009, 2010; Luo, 2012), effectors of the Dot/Icm system in *L. pneumophila* seem to be highly redundant as a deletion of 71 effectors leads to no growth defect in macrophages (Luo and Isberg, 2004; Ninio and Roy, 2007; O'Connor *et al.*, 2011).

T4 substrates were identified either by bioinformatics, for example based on the presence of a C-terminal secretion signal (Cazalet *et al.*, 2004), by predicted transcriptional regulatory properties (Zusman *et al.*, 2007) or by screening for translocation of protein fusions with assayable enzymatic functions. Furthermore, it was shown that T4 effectors often have sequence similarities to proteins primarily found in eukaryotes (De Felipe *et al.*, 2005). These can be eukaryotic domains/motifs such as ankyrin repeats, which are present in several *Legionella* effectors or eukaryotic-like proteins which were likely acquired directly from protozoa (Cazalet *et al.*, 2004; Gomez-Valero *et al.*, 2019).

1.3.4.6 The translocation signal of Dot/Icm effectors

It is generally acknowledged that similar to proteins involved in conjugation processes by related T4 systems, the translocation signal of effectors of the Dot/Icm system resides at their C-terminal end and is about 20 amino acids long (Vergunst *et al.*, 2000; Luo and Isberg, 2004).

In *L. pneumophila* a large number of effectors were identified harboring a C-terminal translocation signal with a preference for short polar and negatively charged amino acids (Nagai *et al.*, 2005; Kubori, Hyakutake and Nagai, 2008; Burstein *et al.*, 2009). In 2001, Huang *et al.* proposed a first translocation motif for a number of Dot/Icm substrates which is 3-6 residue long and located 10-17 amino acids from the C-terminus of the protein. It is composed of several consecutive glutamic acid residues and is therefore called E-block motif.

Although some studies show that the last 20 amino acids are most important and sufficient for successful effector translocation by the Dot/Icm system (Nagai *et al.*, 2005), some additional signals have been described. Huang *et al.* could still show low levels of translocation after deleting the 100 C-terminal amino acids of the effector SidC, suggesting an internal signal. In SidJ two independently acting signals were identified. A conventional C-terminal signal is needed for translocation of this effector at an early point in infection, whereas an internal motif modulates secretion at later time points (Jeong, Sexton and Vogel, 2015).

Internal signals were also found in effectors of other T4 systems such as the VirB/D4-like system in *Bartonella* spp. (Schulein *et al.*, 2005). In *H. pylori*, the effector CagA requires an additional N-terminal motif as well as an internal motif (Couturier *et al.*, 2006; Hohlfeld *et al.*, 2006; Pattis *et al.*, 2007). Intracellular delivery domains can also be found in relaxases involved in the conjugative process

(Schulein *et al.*, 2005).
As, however, internal signals are mostly found in addition to a C-terminal signal, translocation of effector proteins is expected to be post-translational.

1.3.4.7 Effector recognition and loading

In general, effectors in *L. pneumophila* are thought to be recognized by the coupling complex of the Dot/Icm system. The structural study of Meir *et al.* provides mechanistic models for the recruitment of two different subgroups of effector proteins. As already proposed in 2018, E-block motif-containing substrates bind to the positively charged surface of the coupling complex component DotM in an IcmSW-independent manner (Meir *et al.*, 2018). Alternatively, a significant number of effectors are IcmSW-dependent and harbor IcmSW binding sites. Bound to DotL, IcmSW as well as LvgA seem to act as adaptor proteins scanning the environment of the Dot/Icm system for effector proteins. If they only capture effector proteins or additionally promote exposure of their C-terminal or possibly internal signals, as it was shown for SidG (Cambronne and Roy, 2007), has not yet been investigated. Nevertheless, the coupling complex of the Dot/Icm system seems to be the central receptor platform providing different interaction sites for different kinds of effector subgroups.
After being recognized by the T4CC effector proteins are believed to be handed over to the T4 ATPases, which in turn transfer them to the secretion channel (Ghosal *et al.*, 2018).

1.4 TMD-effectors – a new effector subgroup

Until now, effectors of the Dot/Icm system have been classified in various subgroups based on their chaperone-dependence, the presence of an E-block motif, sequence similarity to other proteins or their function within the host cell.

In 2018, Krampen *et al.* divided the effector proteins of the T3SS in *Salmonella* Typhimurium, *Chlamydia* and *E. coli* into two new groups: soluble effectors and transmembrane domain-containing effectors (TMD-effectors) (Krampen *et al.*, 2018). TMD-effectors have their final destination in one of the membranes of the eukaryotic host cell and execute various functions during the course of infection.

1.4.1 Why are TMD-effectors important for the pathogenicity of *L. pneumophila*

A key event in the course of infection of *L. pneumophila* is the inhibition of phagosome-lysosome fusion (Horwitz *et al.*, 1983). As briefly mentioned above, effector proteins translocated by the Dot/Icm system are used in order to hijack host cell machineries such as vesicle trafficking, membrane fusion and signaling pathways and with this promote the concealment as well as maturation of the *L*CV.

This chapter briefly summarizes what is already known about the function of four TMD-effectors, which were selected for this study as well as two soluble effectors which serve as comparison.

Eukaryotic membranes are characterized, among other properties, by the presence of phosphatidylinositol molecules which are not only important for signal transduction and actin remodeling but especially for defining organelle identity and membrane dynamics (Odorizzi, Babst and Emr, 2000; De Matteis and Godi, 2004; Di Paolo and De Camilli, 2006). To this end, effector proteins which can either bind or modify these molecules have an important function during the course of infection.

The TMD-effector SidF, anchored to the *L*CV membrane via its two TMS, is involved in maintaining the phosphatidylinositol composition of the *L*CV by hydrolyzing the D3 phosphate of PI(3,4)P_2 and PI(3,4,5)P_3 which are generated in the phagosome in an early stage of infection (Vieira *et al.*, 2001; Kamen, Levinsohn and Swanson, 2007), and by this inhibits endosome-lysosome fusion (Vergne, Chua and Deretic, 2003; Flannagan, Cosío and Grinstein, 2009).

Ceg4, another TMD-effector, attenuates the activation of the mitogen-activated protein kinase (MAPK) pathway which is involved in stress response and apoptosis of eukaryotic cells. It manipulates the phosphorylation cascades with its phosphotyrosine-specific phosphatase activity.

SidM, also called DrrA, and RalF are well studied soluble effector proteins. Upon translocation into the host cell cytoplasm, they are recruited to the *L*CV. Subsequently, acting as guanine nucleotide exchange factors, they activate small GTPases that are involved in many signal transduction pathways and in the regulation of the actin cytoskeleton (Jaffe 2005). SidM binds to phosphatidytl-inositol-4-phosphate molecules on the *L*CV, and activates the small GTPase Rab1 which in turn regulates vesicle transport from the ER to the Golgi apparatus (Machner and Isberg, 2006; Ingmundson *et al.*, 2007; Faucher, Mueller and Shuman, 2011). RalF recruits the small GTPase ADP-ribosylation factor 1 (Arf1) which is involved in retrograde vesicle transport from the Golgi apparatus to the ER (Nagai *et al.*, 2002).

The TMD-effectors LegC2 and LegC3 are eukaryotic-like proteins which interfere with organelle trafficking (De Felipe, Glover, Charpentier, Anderson, Reyes, Pericone and H. A. Shuman, 2008; Bennett *et al.*, 2013a). Harboring coiled-coil super secondary structural elements they resemble eukaryotic Q-SNARE proteins and are able to form stable hybrid complexes with endogenous R-SNARE proteins and

by this promote vesicle fusion with and maturation of the LCV (Fig. 4) (Shi *et al.*, 2016). At the same time endosomes are prevented from maturing. It was shown that together with LegC7, LegC2 and LegC3 can form fully functional SNARE acceptor complexes that can substitute endogenous Q-SNAREs and thus mediate membrane fusion. As the once formed complex cannot be disassembled anymore, this proves to be an irreversible hijacking of the SNARE machinery by the pathogen. In this context, it was shown that the presence of their TMS is important for proper effector localization within different host cell membranes (Bennett *et al.*, 2013b). While LegC2 localizes to the *L*CV in mammalian cells, LegC3 localizes to the plasma membrane and induces fragmentation of the intracellular degradative vacuole when ectopically expressed in *S. cerevisiae* (De Felipe *et al.*, 2008).

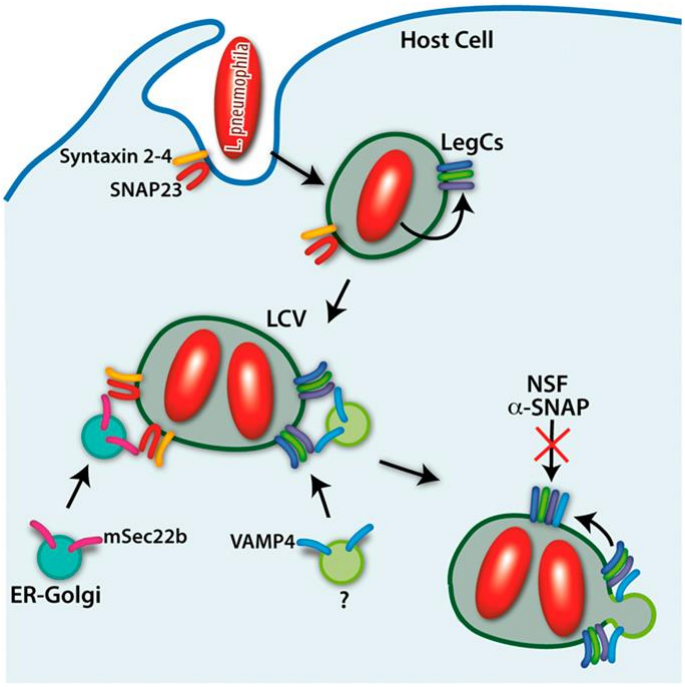


Fig. 4: Manipulation of membrane trafficking in the host cell.
LegC effectors translocated by the Dot/Icm system in *L. pneumophila* uses SNARE mimicry in order to manipulate membrane trafficking in the host cell. LegC proteins are integrated in the *L*CV membrane in order to recruit VAMP4 vesicles and with this membranes for vacuolar growth. (Adopted from Shi *et al.*, 2016)

1.4.2 TMD-effector targeting within the pathogen

As shown above, TMD-effectors carry out different important functions within the host cell. Therefore, their successful translocation out of the pathogen is vital for the course of infection of *L. pneumophila*. This does not only depend on successful effector recognition and subsequent transport through the secretion channel but starts with proper effector targeting to the secretion machinery. In contrast to

soluble effectors, TMD-effectors possess not only their secretion signal but additionally one or more hydrophobic transmembrane segments that can possibly be recognized by SRP and result in inner membrane integration by the Sec system.

1.4.2.1 Bacterial inner membrane proteins: targeting and insertion

Beside receptor functions for cell signaling or enzymatic activities membrane-integral as well as periplasmic proteins are components of protein transport systems or secretion systems. In order to remain stable in membranes, the proteins have two basic structures: They are either composed of α-helix bundles spanning the lipid bilayer in all kinds of cellular membranes, or β-barrels, mostly found in outer membrane proteins. In all cases, the membrane spanning region consists of hydrophobic amino acids flanked by hydrophilic residues (Elofsson and Von Heijne, 2007). The bacterial inner membrane can easily be spanned by peptides with a length of approximately 20 amino acids, however, hydrophobic stretches up to 40 residues are possible, including slopes and kinks (Wallin and Von Heijne, 1998; Papaloukas *et al.*, 2008).
In order to reach their final destination, membrane proteins have to undergo two major events: First of all, they have to be recognized as membrane proteins and targeted correctly to the appropriate membrane. In a second step, they must be inserted into the membrane and undergo proper folding.

1.4.2.1.1 SRP-targeting pathway

Inner membrane proteins harbor their hydrophobic signal in one of their transmembrane segments (TMS) which, after inner membrane insertion, span the lipid bilayer. This hydrophobic signal consists of three regions: A positively charged N-tail, a hydrophobic core and a polar C-terminus (Papanikou, Karamanou and Economou, 2007). As soon as the signal emerges from the ribosome, the Signal Recognition Particle (SRP), containing a small 4.5S RNA bound to the fifty-four homologue (Ffh) protein (Luirink and Sinning, 2004), binds to a hydrophobic stretch of 12-17 amino acids of the nascent polypeptide chain (White and Heijne, 2004; Schibich *et al.*, 2016). The SRP-ribosome-nascent-chain (RNC) complex is then bound by the SRP receptor FtsY resulting in the delivery of the complex to the SecYEG translocase (Fig. 5) (Luirink *et al.*, 1994; Papanikou, Karamanou and Economou, 2007). Next to the twin arginine translocation pathway, the Sec system is crucial for the general housekeeping of the bacterial cell including integration of inner membrane proteins as well as protein translocation into the bacterial periplasm (Lycklama a Nijeholt and Driessen, 2012).
The SRP-FtsY complex assembly is a crucial checkpoint in the co-translational targeting process in

which the trigger factor (TF) reduces the rate of translation for stable SRP-FtsY complex assembly (Holtkamp *et al.*, 2012). Hydrolysis of the GTP molecules bound by SRP and FtsY promotes the dissociation of the complex and the transfer of the RNC to the SecYEG translocon (Kol, Nouwen and Driessen, 2008). Integration into the bacterial inner membrane involves the membrane embedded proteins SecY, SecE and SecG, which form a protein-conducting channel, as well as the membrane protein insertase YidC (Fig. 5). Translation drives the secretion of the nascent chain through the SecYEG channel while SecY opens laterally allowing the TMS to interact with the lipids of the inner membrane (Paetzel *et al.*, 2002). The energy for this process is provided by the ATPase SecA, which is peripherally attached to the SecY channel. The topology of inner membrane proteins is mostly determined by the preferential occurrence of positively charged amino acids, such as lysine and arginine, at the cytoplasmic side of their TMS (von Heljne, 1989).

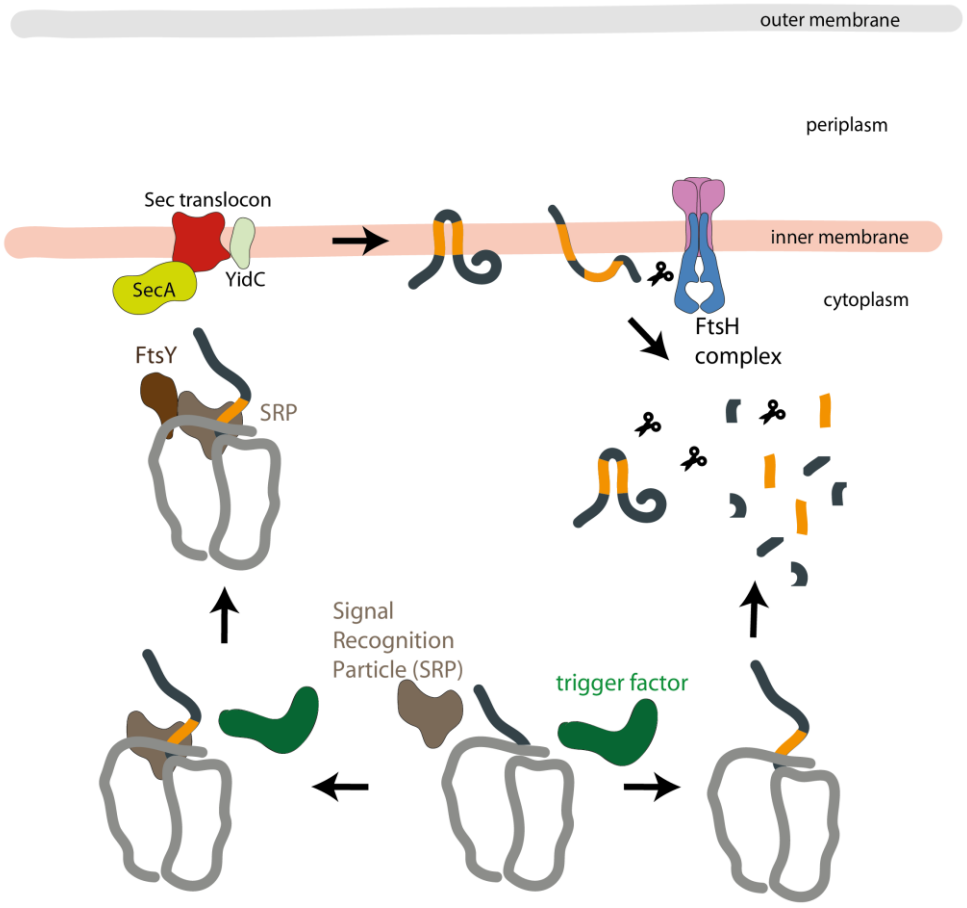


Fig. 5: SRP targeting and protein quality control of inner membrane proteins.
Inner membrane proteins are co-translationally targeted by SRP to the Sec translocon. SRP as well as the trigger factor are bound at the exit tunnel of the ribosome waiting for a nascent chain to emerge. After SRP binds to the hydrophobic segments (orange) of a polypeptide chain, the SRP-ribosome-nascent-chain complex is then bound by the SRP receptor FtsY resulting in the delivery of the complex to the SecYEG translocase and the integration of the protein into the bacterial inner membrane.
Unprotected hydrophobic patches result in protein aggregation and subsequent degradation.

Specific proteases such as FtsH and GlpG, which are located within the bacterial inner membrane, degrade misfolded protein.

1.4.2.1.2 Protein quality control

Although inner membrane targeting and insertion pathways are well regulated, it is crucial that mis-localization of membrane proteins is avoided. Hydrophobic stretches within proteins are prone for aggregation, and thus must be protected as soon as they emerge from the ribosome. Chaperones play an important role in this protein quality control. They can act as holdases, keeping a protein in a specific state, foldases or unfoldases, helping in the process of protein folding, or unfolding wrongly folded proteins (Mayer and Bukau, 2005; Elad *et al.*, 2007; Liberek, Lewandowska and Ziętkiewicz, 2008; Saibil, 2008). TF is the only known ribosome-associated chaperone in bacteria and localized directly at the ribosome exit site (Kramer *et al.*, 2002; Craig, Eisenman and Hundley, 2003; Baram *et al.*, 2005), helping in the initial folding steps and thus protecting hydrophobic stretches of membrane proteins from misfolding and aggregation when the nascent polypeptide chain emerges from the ribosome. Other chaperones such as DnaK and the GroEL/ES protein complex act further downstream after translation of the protein. Moreover, damaged membrane proteins are recognized by specific proteases such as FtsH and GlpG which are located within the bacterial inner membrane and degrade misfolded protein (Kihara, Akiyama and Koreaki, 1999; Ito and Akiyama, 2005).
As 25-30 % of all proteins are destined for translocation, factors such as SecB and SRP are present at the ribosomal exit tunnel in order to bind to periplasmic, outer membrane or inner membrane proteins, respectively (Driessen and Nouwen, 2008). Moreover, there are factors influencing the stability and processing of the nascent chain making the ribosome exit tunnel a rather crowded area.

1.4.2.1.3 TF vs. SRP

It was shown that SRP and TF bind to the same L23 ribosomal protein (Kramer *et al.*, 2002; Baram *et al.*, 2005; Ariosa *et al.*, 2015). There are, however, contradictory statements about how SRP and TF influence each other. On the one hand, it is stated that TF and SRP bind to the ribosome at the same time and by this have to compete for ribosome and substrate, and weaken each other's binding about ten-fold (Sandikci *et al.*, 2013; Bornemann, Holtkamp and Wintermeyer, 2014) whereas, on the other hand, Schibich *et al.* found in 2016 that SRP and TF have only limited substrate overlap, with TF primarily binding to cytoplasmic, periplasmic and outer membrane proteins and SRP binding periplasmic and inner membrane proteins. Furthermore, they could show that binding of inner membrane proteins by TF or SRP is temporally separated with SRP binding mostly before TF does. In

contrast, Bornemann *et al.* shows that TF prevents SRP binding to the majority of ribosomes (Bornemann, Holtkamp and Wintermeyer, 2014). Moreover, the hydrophobicity of inner membrane proteins might influence the binding of the TF and SRP. With increasing hydrophobicity, the affinity to the TF seems to decrease whereas it increases with respect to SRP (Beck *et al.*, 2000; Lee and Bernstein, 2001).

1.4.2.1.4 Characteristic protein features of inner membrane proteins

In order to prevent misfolding and aggregation in the cytoplasm, proper recognition and targeting of inner membrane proteins is crucial. A linear correlation between SRP recognition and the hydrophobicity of nascent TMS were shown in crosslinking experiments with different PhoA mutants (Valent *et al.*, 1995). Moreover, the amino acid composition of the signal sequence seems to be important for substrate specificity as the introduction of leucine residues promotes SRP binding (Adams *et al.*, 2002).

Over the years, prediction programs such as TOPCONS (Tsirigos *et al.*, 2015) or the ΔG predictor (Hessa, Nadja M Meindl-Beinker, *et al.*, 2007) were developed in order to predict membrane proteins and their topology based on their hydrophobicity and the localization of positive amino acids flanking their TMS. In addition, thermodynamics studies were conducted revealing that single amino acids prefer distinct positions across the lipid bilayer. Moreover, the length of the TMS affects their insertion as very long TMS can tilt and interact along their full length with the lipids (Hessa, Nadja M Meindl-Beinker, *et al.*, 2007). This interaction of polypeptide chains with the lipids is called membrane partitioning and can be described thermodynamically as "apparent Gibbs free energy" of insertion (ΔG_{app}). Using *in vitro* expression systems with model membrane proteins, Hessa *et al.* were able to quantitatively measure the position-dependent ΔG_{app} for each natural amino acid in a model TMS containing various numbers of alanine and leucine residues. As hydrophobic amino acids can readily interact with lipid molecules, they have the lowest ΔG_{app} value. The probability of the partitioning of a transmembrane segment into the lipid bilayer is higher, the lower the sum of the ΔG_{app} of its constituting amino acids is.

1.4.2.2 TMD-effector targeting to the T3SS and T4SS

As TMD-effectors can in principle be recognized by SRP due to their hydrophobic TMS, these substrates face a possible targeting conflict. Protein mistargeting does not only lead to a loss of effectors but might also be toxic for the bacterium and therefore must be avoided. As mentioned in chapter 1.2.1, soluble substrates of the T3SS are guided to the injectisome by an N-terminal 20-25 residue-long

secretion signal (Arnold, Jehl and Rattei, 2010), often followed by a 20-50 residue-long chaperone-binding domain. In contrast, T4SS effectors harbor a C-terminal signal sequence. It is, however, not yet understood if this signal or rather chaperones and thus chaperone binding sites are responsible for proper targeting of effector proteins to the Dot/Icm system. If only the C-terminal T4 signal is responsible for effector targeting, effector binding can only take place post-translationally whereas SRP already binds to hydrophobic segments of polypeptide chains as soon as they emerge from the ribosome.

As cognate T3SS chaperones are often encoded adjacent to their T3SS substrate they can bind the T3 TMD-effector co-translationally and thus shield hydrophobic segments from being recognized by SRP (Krampen *et al.*, 2018). Furthermore, Krampen *at al.* could show that a balanced hydrophobicity of the TMS of T3SS substrates supports targeting discrimination. While being sufficiently hydrophobic for principal membrane integration, SRP-dependent membrane targeting was in general not facilitated by these segments. These results suggest that a TMS-specific co-translational targeting mechanism by T3 chaperones prevents co-translational mistargeting by SRP for subsequent post-translational secretion of membrane proteins through the T3SS injectosome.

Although a reduced hydrophobicity also seems to be true for most of the TMD-effectors in *L. pneumophila*, the Dot/Icm system is also able to secrete TMD-substrates of higher hydrophobicity (Krampen *et al.*, 2018). Moreover, one can only hypothesize about an active avoidance mechanism similar to the T3SS, as very little is known about chaperone binding domains in effector proteins and their targeting to the Dot/Icm system.

As mentioned above, other T4SS are able to secrete effector proteins that are targeted to and inserted into the bacterial inner membrane or translocated into the periplasmic space prior to recruitment to the T4SS machine and secretion out of the bacterium into the host cell (Dolezal *et al.*, 2012). It is at present entirely unclear, how prevalent this two-step secretion pathway of TBSS substrates is, how the T4BSS might recognize intramembrane substrates, how and to which side these substrates are extracted from the membrane and where they enter the secretion machinery for subsequent translocation into the eukaryotic host.

2 Aim of the work

Effector proteins translocated by the secretion machinery are used to hijack and subvert host cell function in order to live and colonize within the host cell. Not only soluble effectors but also TMD-effectors can successfully be secreted by the Dot/Icm system. Although essential for their proper localization within the host cell, the presence of hydrophobic transmembrane segments within the protein poses a possible targeting conflict as TMS can be recognized by SRP and result in subsequent inner membrane insertion. Krampen *et al.* previously showed that some secreted effector proteins of the type IV secretion system contain TMS of lower hydrophobicity and by this presumably avoid SRP targeting, however, also higher hydrophobic TMD-effectors exist.

In the study I aimed to elucidate if TMD-effectors with higher hydrophobicity follow a two-step secretion pathway with an inner membrane intermediate (Fig. 6) or can avoid inner membrane targeting by chaperone binding.

To follow this hypothesis, the SRP targeting potential as well as the inner membrane integration of TMD-effectors will be determined using membrane fractionation followed by mass spectrometry. To investigate a potential active avoidance mechanism by chaperone binding, I analyzed if TMD-effectors depend on the chaperones IcmSW for their successful translocation into host cells and if chaperone binding sites can be identified.

Another focus of this work is the question how TMD-effectors are recognized by the Dot/Icm system once integrated in the bacterial inner membrane. For this the membrane topology of the TMD-effectors LegC3 and SidF were determined and potential protein receptors in the periplasm as well as in the cytoplasm were investigated.

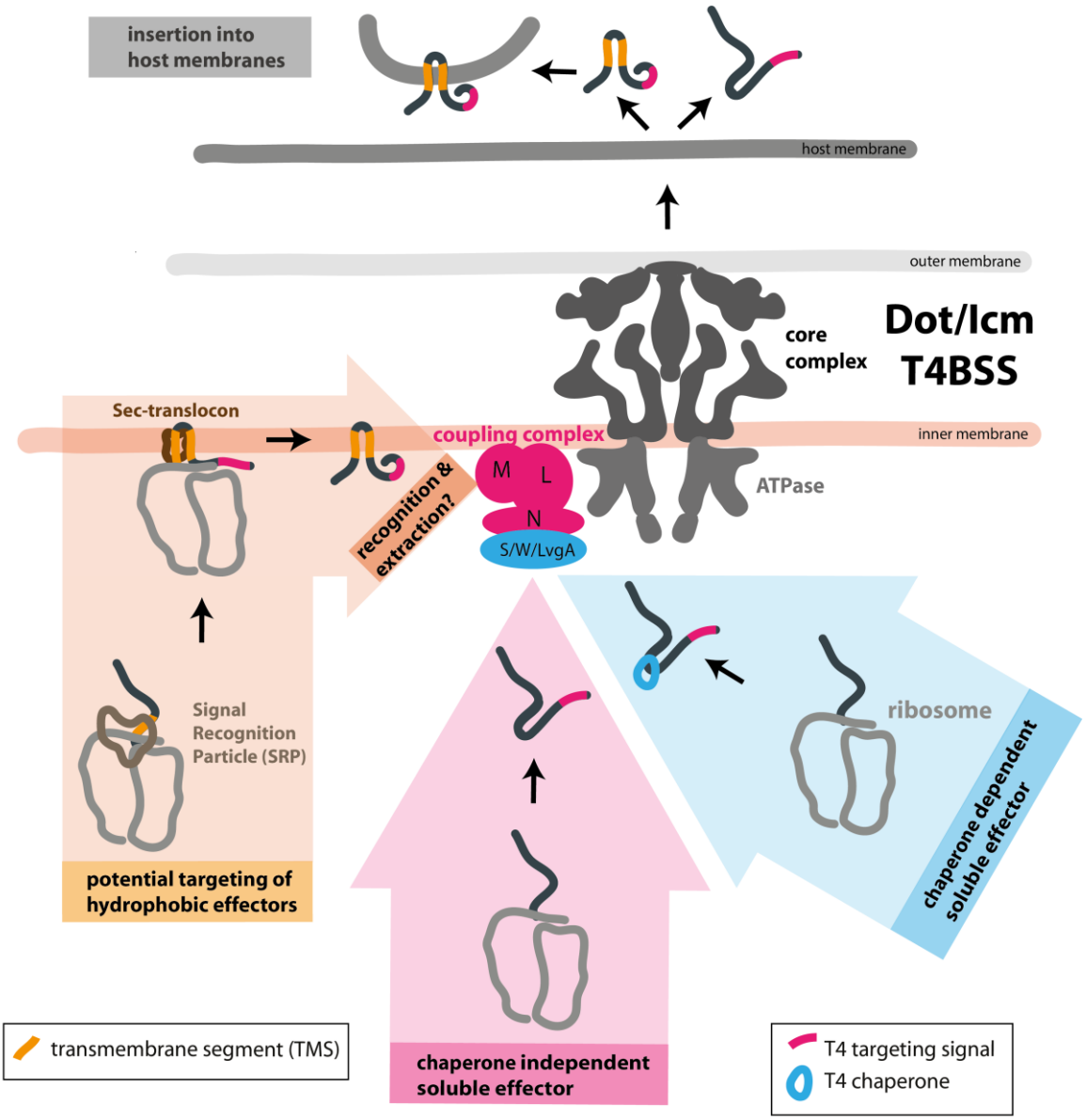


Fig. 6: Effector targeting pathways to the Dot/Icm system.
Soluble effector proteins are targeted to the coupling complex (rose) in a chaperone-independent (rose arrow) or chaperone-dependent (blue arrow) targeting pathway. Effector proteins harboring a transmembrane segment (orange) may follow a hypothetical two-step secretion pathway with an inner membrane intermediate (orange arrow).

3 Materials and Methods

3.1 Enzymes, chemicals, media and buffers

If not mentioned otherwise, all enzymes and chemicals used in this study were purchased from Sigma-Aldrich (München), Novagen (Darmstadt), New England Biolabs (Frankfurt am Main), Merck (Darmstadt), Carl Roth (Karlsruhe), Becton Dickinson (Le Pont De Claix, France), Biozym (Oldendorf), Thermo Fisher Scientific (Rockford,USA) or AppliChem (Darmstadt). All media, buffers, and solutions used in this study are listed in table 1. The media used in this study were autoclaved or filter sterilized.

Table 1: Media, buffers and solutions used in this study

Name	Contents
LB medium	5 g NaCl, 10 g tryptone and 5 g yeast extract, dissolved and filled up to 1 l with H_20
LB agar	6 g LB Lennox and 4.5 g agar, dissolved and filled up to 300 ml with H_20.
SOB medium	40 g bacto-tryptone, 10 g yeast extract, 1 g NaCl, 0.373 g KCl, dissolved and filled up to 2 l with H_20 and autoclaved. Sterile filtered (0.22 µm) 1 M $MgCl_2$ and 1 M $MgSO_4$ solutions were added to a final concentration of 10 mM.
SOC medium	20 ml sterile filtered (0.22 µm) 1 M glucose were added to 1 l SOB medium
BCYE agar	2 g activated charcoal, 10 g yeast extract, 13 g agar, dissolved and filled up to 1 l with H_20. For sucrose agar medium components were dissolved and filled up to only 900 ml with H_20. 100 ml sterile filtered 50 % (w/v) sucrose was added after autoclaving.
BCYE supplements	5 g ACES (N-(2-Acetamido)-2-aminoethanesulfonic acid) buffer, 0.125 g ferric pyrophosphate, 0.2 g L-cysteine hydrochloride, 0.5 g α-Ketoglutarate, dissolved in 50 ml H_20
ACES-buffered yeast extract broth	10 g yeast extract and 10 g ACES buffer dissolved and filled up to 1 l with H_20; pH was adjusted to 6.9 with KOH
Stock medium	63 g glycerol, 10 g peptone, dissolved and filled up to 500 ml with H_20
DMEM	DMEM medium supplemented with 10 % (v/v) FCS, 1 % (v/v) L-glutamine 200 mM (all from Gibco)
10 x PBS	80 g NaCl, 2 g KCl, 14.4 g Na_2HPO_4 2 H_20, 2.4 g KH_2PO_4, dissolved and filled up to 1 l with H_20, pH was adjusted to 7.4 with NaOH
Buffer K	50 mM TEA, 250 mM sucrose, 1 mM EDTA, dissolved and filled up to 500 ml with H_20. The pH was adjusted to 7.5 with acetic acid.
50x TAE buffer	242 g Tris base, 57.1 ml glacial acetic acid, 37.2 g EDTA, dissolved and filled up to 1 l to H_20
4x SB buffer	10 ml 0.5 M Tris-HCl, pH 6.8, 5 ml 80 % (v/v) glycerol, 1.6 g SDS, 10 mg bromophenol blue, filled up to 16 ml with H_20. Before

	use the buffer was supplemented with a final concentration of 5 % (v/v) β-mercaptoethanol
6x DNA loading buffer	1.9 ml 80 % (v/v) glycerol, 20 µl Tris-HCl pH 8.8, small crumb of bromophenol blue, dissolved and filled up to 3 ml with H_20
10x SDS running buffer	30 g Tris base, 144 g glycine, 10 g SDS, dissolved and filled up to 1 l with H_20
Coomassie staining solution	0.4 % (w/v) Coomassie, 50 % (v/v) ethanol, 5 % (v/v) acetic acid
10x Transfer buffer	30 g Tris base, 144 g glycine, 2.5 g SDS, dissolved and filled up to 1 l with H_20
Transfer buffer	10x transfer buffer was diluted 1:10 in H_20 and supplemented with a final concentration of 10 % (v/v) methanol
10x TBS	94 g NaCl, 30 g Tris base, dissolved and filled up to 1 l with H_20. pH was adjusted to 8.0 with HCl
TBS-T	10x TBS was diluted 1:10 with H_20 and supplemented with 0.05 % (v/v) Tween20
5x ISO mix	300µl 1M Tris-HCl pH 7.5, 30 µl 1 M $MgCl_2$ 60 µl 10 mM dNTP Mix, 30 µl 1 M DTT, 150 mg PEG 8000, 30 µl 100 mM NAD, 600 µl H_20
Gibson master mix	100 µl 5x ISO mix, 0.2 µl T5 exonuclease, 6.25 µl Phusion DNA polymerase, 50 µl Taq DNA Ligase, 218.6 µl H_20
10x Anode buffer	52.3 g Bis-Tris dissolved and filled up to 500 ml with H_20
10x Cathode buffer I	44.79 g tricine (Biomol, Hamburg), 15.69 g Bis-Tris, 1 g Coomassie Serva Blue G, dissolved and filled up to 500 ml with H_20
10x Cathode buffer II	44.79 g tricine (Biomol, Hamburg), 15.69 g Bis-Tris, dissolved and filled up to 500 ml withH_20
Blue native loading buffer	25 mg Coomassie Serva Blue G dissolved in 450 µl 250 mM amino-caproic acid (ACA), supplemented with 50 % (v/v) glycerol

3.2 Antibodies

All antibodies used in the present study (table 2) were diluted in TBS-T before use. The primary antibodies α-FLAG and α-HA were purchased from Sigma-Aldrich (München). The secondary antibodies are conjugated to either DyLight 680 or DyLight 800 4x PEG and were purchased from Thermo Fisher Scientific (Rockford, USA). The conjugation with a fluorophore enables the detection of the secondary antibodies by a LI-COR Odyssey system (ODY-3191).

Table 2: List of antibodies used in this study

	Antibody	Clonality	Dilution	Origin
Primary antibodies	α-HA	monoclonal	1:1000	mouse
	α-HA	monoclonal	1:1000	rabbit
	α-FLAG	M2, monoclonal	1:5000	mouse
Secondary antibodies	α-Mouse680	Polyclonal	1:10000	goat
	α-Mouse800	Polyclonal	1:10000	goat
	α-Rabbit680	Polyclonal	1:10000	goat
	α-Rabbit800	Polyclonal	1:10000	goat

3.3 Bacterial and eukaryotic strains and their growth conditions

Legionella pneumophila strain Philadelphia-1 (LP01) was used as wild type strain. *Escherichia coli* strains NEB5α and TOP10 were used for molecular cloning procedures. A full list of bacterial strains used in this study is provided in table 3.

E. coli strains were either grown on LB agar medium at 37 ° C or in liquid LB medium at 37 ° C, 180 rpm. All media were supplemented as required with streptomycin (50 µg/ml), ampicillin (100 µg/ml), kanamycin (25 µg/ml), or chloramphenicol (50 µg/ml) for plasmid or strain selection.

L. pneumophila strains were grown on BCYE-Agar plates supplemented with "BCYE Growth Supplements" or in ACES-buffered yeast extract broth at pH 6.9 supplemented with 0.4 mg/ml L-cysteine and 0.25 mg/ml ferric pyrophosphate. All media were supplemented with necessary antibiotics in the following concentration: 50 µg/ml streptomycin, 10 µg/ml kanamycin (15 µg/ml were used for the plasmid pSR47S) or 6.8 µg/ml chloramphenicol. Protein overexpression from the low-copy number plasmid pxDC61 was induced by supplementing the media with 0.5 mM Isopropyl-β-D-thiogalactopyranosid (IPTG).

Bacterial strains were stored at -80 °C in stock media. Raw 264.7 macrophages expressing LgBiT were grown in DMEM containing 10 % (v/v) fetal calf serum (FCS) and 1 % (v/v) L-glutamine at 37 °C and 5 % CO_2.

Table 3: *Escherichia coli* and *Legionella pneumophila* strains used in this study

Name	Species	Genotype	Source
NEB5α	*Escherichia coli*	fhuA2Δ(argF-lacZ)U169 phoA ginV44 Φ80Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17	lab collection
Top10	*Escherichia coli*	F mcrA Δ(mrr-hsdRMS-mcrBC) ϕ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(Str^R) endA1 λ-	Kindly provided from the Hantke lab
CR14	*Escherichia coli* DH5α λpir	Sup E44, ΔlacU169 (ΦlacZΔM15), recA1, endA1, hsdR17, thi-1, gyrA96, relA1, λpir phage lysogen (Zuckman, Hung and Roy, 1999)	Kindly provided from the lab of Craig Roy
CR19	*Escherichia coli*	MT607 E. coli containing plasmid pRK600; ColE1 replicon with RK2 transfer genes	Kindly provided from Hiroki Nagai
LP01	*Legionella pneumophila*	rpsL	Kindly provided from Hiroki Nagai
MIBL1002	*Legionella pneumophila*	LP01 Δ*dotO*	Kindly provided from Hiroki Nagai
MIBL1005	*Legionella pneumophila*	LP01 Δ*dotK*	Kindly provided from Hiroki Nagai
MIBL1009	*Legionella pneumophila*	LP01 Δ*icmS*	Kindly provided from Amit Meir
MIBL1011	*Legionella pneumophila*	LP01 Δ*dotG*	This study
MIBL1012	*Legionella pneumophila*	LP01 Δ*icmW*	Kindly provided from Amit Meir
MIBL1013	*Legionella pneumophila*	LP01 Δ*dotF*	Kindly provided from Hiroki Nagai
MIBL1014	*Legionella pneumophila*	LP01 (ΔT4SS) chromosomal deletions of three loci: icmX–dotA, dotB–dotD, and icmT–dotU	Kindly provided from Amit Meir
MIBL1015	*Legionella pneumophila*	LP01 *dotM* R196E/R197E	Kindly provided from Amit Meir
MIBL1019	*Legionella pneumophila*	LP01 $DotM^{FLAG}$	This study
MIBL1021	*Legionella pneumophila*	LP01 ^{FLAG}DotN	This study
MIBL1022	*Legionella pneumophila*	LP01 ^{FLAG}DotY	This study
MIBL1023	*Legionella pneumophila*	LP01 ^{FLAG}DotZ	This study

3.4 Molecular cloning

DNA modifications in plasmids were introduced using Gibson-Cloning (Gibson *et al.*, 2010) or the QuikChange site-directed mutagenesis protocol (Agilent, 2018).

A full list of plasmids used and constructed in this study is provided in table 4. Oligonucleotides needed for plasmid generation were synthetized by Eurofins Genomics and are listed in table 5.

Table 4: List of plasmids used in this study (Cm chloramphenicol, Kan kanamycin)

Name	**Description**	**Resistance**	**Source**	**Construction** (Made by Gibson assembly of PCR products of the following two primer/template pairs)
pEvol-*p*Bpa	Amber suppressor plasmid (constitutive)	Cm	Lab stock	
pxDC61	Broad range expression plasmid	Cm	pXDC61 was a gift from Howard Shuman (Addgene plasmid #21841; http://n2t.net/addgene:21841; RRID:Addgene_21841)	
pMIB6814	pxDC61-^{2HA}SidF	Kan	Lab stock	
pMIB6815	pxDC61-^{2HA}LegC3	Kan	Lab stock	
pMIB7802	pT10-SseF$^{HiBit-3FLAG}$	Kan	Lab stock	
pMIB5505	pT12-Lep-inv	Kan	Lab stock	

pMIB5327	PT10-Sicp-ubiquitin-GFP-3XFLAG	Kan	Lab stock	
pMIB5692	pT10-Sicp-3xFlag-SopA-SopE-Sptp-ubiquitin^{E3G,E13G}-GFP	Kan	Lab stock	
pMIB6828	pxDC61-$^{HiBiT-2HA}$LegC2	Cm	This study	Plasmid: gib_uni_pxDC61_f + gib_uni_pxDC61_r from pxDC61; Insert: gib_pxDC61_Hibit_f + gib_pxDC61_LegC2_r from pMIB6821
pMIB6829	pxDC61-$^{HiBiT-2HA}$LegC3	Cm	This study	Plasmid: gib_uni_pxDC61_f + gib_uni_pxDC61_r from pxDC61; Insert: gib_pxDC61_Hibit_f + gib_pT12_LegC3_r from pMIB6816
pMIB6830	pxDC61-$^{HiBiT-2HA}$SidF	Cm	This study	Plasmid: gib_uni_pxDC61_f + gib_uni_pxDC61_r from pxDC61; Insert: gib_pxDC61_Hibit_f + gib_pxDC61_SidF_r from pMIB6818
pMIB6848	pxDC61-$^{HiBiT-2HA}$Ceg4	Cm	This study	Plasmid: gib_uni_pxDC61_f + gib_uni_pxDC61_r from pxDC61; Insert: gib_pxDC61_Hibit_f + gib_pT12_Ceg4_r from pMIB6819
pMIB6841	pxDC61-$^{HiBiT-2HA}$SidM	Cm	This study	Plasmid: gib_uni_pxDC61_f + gib_uni_pxDC61_r from pxDC61; Insert: gib_pxDC61_Hibit_f + gib_pxDC61_SidM_r from pMIB6820
pMIB7196	pxDC61-$^{HiBiT-2HA}$RalF	Cm	This study	Plasmid: gib_uni_pxDC61_f + gib_uni_pxDC61_r from pxDC61;

				insert: gib_pxDC61_Hibit_f + gib_pxDC61_RalF_r frompMIB7193
pMIB7898	pxDC61-$^{HiBiT-2HA}$Lep-inv	Cm	This study	Insert: gib_HA_Lep-inv_f + gib_pxDC61_Lep-inv_r from pMIB5505; Plasmid: gib_uni_HA_r + gib_uni_pxDC61_f from pMIB6830
pMIB6855	pxDC61-$^{HiBiT-2HA}$LegC2Δ30	Cm	This study	Plasmid: gib_uni_pxDC6_f + gib_pxDC61_LegC2-30aa_r from pMIB6829
pMIB7198	pxDC61-$^{HiBiT-2HA}$LegC3Δ30	Cm	This study	Plasmid: gib_uni_pxDC6_f + gib_pxDC61_LegC3dc30_r from pMIB6829
pMIB7197	pxDC61-$^{HiBiT-2HA}$SidFΔ30	Cm	This study	Plasmid: gib_uni_pxDC6_f + gib_pxDC61_SidFdc30_r from pMIB6830
pMIB7199	pxDC61-$^{HiBiT-2HA}$Ceg4Δ30	Cm	This study	Plasmid: gib_uni_pxDC6_f + gib_pxDC61_Ceg4dc30_r from pMIB6848
pMIB6860	pxDC61-$^{HiBiT-2HA}$SidMΔ30	Cm	This study	Plasmid: gib_uni_pxDC6_f + gib_pxDC61_SidM-30aa_r from pMIB6829
pMIB7651	pxDC61-$^{HiBiT-2HA}$RalFΔ30	Cm	This study	Plasmid: gib_uni_pxDC6_f + gib_pxDC61_RalFdc30_r from pMIB7196
pMIB7895	pxDC61-$^{HiBiT-2HA}$SidFΔTMS1ΔTMS2	Cm	This study	Plasmid: gib_uni_SidFC_f + gib_SidFC_SidF760_r from pMIB76830
pMIB7909	pxDC61-$^{HiBiT-2HA}$SidFΔTMS1ΔTMS2Δc30	Cm	This study	Plasmid: gib_uni_pxDC6_f + gib_pxDC61_SidFdc30_r from pMIB7895
pMIB7757	pxDC61-$^{HiBiT-2HA}$SidFΔTMS1ΔT	Cm	This study	Plasmid: gib_uni_pxDC61_f + gib_pxDC61_SidF760_r from pMIB6830

	MS2Δ822-862Δc30			
pMIB7769	pxDC61-$^{HiBiT-2HA}$SidF2Δ862-862Δc30	Cm	This study	Plasmid: gib_uni_SidFc30_f +gib_SidFc30_SidF862_r from pMIB6830
pMIB7770	pxDC61-$^{HiBiT-2HA}$SidF2Δ842-862Δc30	Cm	This study	Plasmid: gib_uni_SidFc30_f +gib_SidFc30_SidF842_r from pMIB6830
pMIB7768	pxDC61-$^{HiBiT-2HA}$SidF2Δ822-862Δc30	Cm	This study	Plasmid: gib_uni_SidFc30_f +gib_SidFc30_SidF822_r from pMIB6830
pMIB7676	pxDC61-$^{HiBiT-2HA}$SidF2ΔTMS2	Cm	This study	Plasmid: gib_SidFTMS1_SidF_f + gib_uni_SidFTMS1_r from pMIB6830
pMIB7755	pxDC61-$^{HiBiT-2HA}$SidF2ΔTMS2Δdc30	Cm	This study	Plasmid: gib_uni_pxDC6_f + gib_pxDC61_SidFdc30_r from pMIB7676
pMIB7892	pxDC61-$^{HiBiT-2HA}$SidF2ΔTMS2Δ862-882	Cm	This study	Plasmid: gib_uni_pxDC6_f + gib_pxDC61_SidFdc30_r from pMIB7769
pMIB7951	pxDC61-$^{HiBiT-2HA}$SidF2ΔTMS2Δ862-882Δdc30	Cm	This study	Plasmid: gib_uni_pxDC6_f + gib_pxDC61_SidFdc30_r from pMIB7892
pMIB7952	pxDC61-$^{HiBiT-2HA}$Lep-inv-SidMc30	Cm	This study	Plasmid: gib_uni_SidMc30_f + gib_uni_HA_r from pMIB6841; Insert: gib_HA_Lep-inv_f + gib_SidMc30_Lep-inv_r from pMIB5505
pMIB7893	pxDC61-$^{HiBiT-2HA}$SidF-TMS1-HiBiT-TMS2	Cm	This study	Plasmid: gib_Hibit-GSSG_SidFTMS2_f + gib_GSSG-Hibit_SidFTMS1_r from pMIB6830
pMIB7896	pxDC61-$^{HiBiT-2HA}$SidF-TMS1-ubiquitin-TMS2	Cm	This study	insert: gib_SidFTMS1-GSSG_Ubi_f + gib_GSSG-SidFTMS2_Ubi_r from pMIB5327; plasmid: gib_uni_GSSG-

				SidFTMS2_f + gib_uni_GSSG-SidFTMS1_r from pMIB7893
pMIB7897	pxDC61-$^{HiBiT-2HA}$SidF-TMS1-ubiquitinE3GE13G-TMS2	Cm	This study	insert: gib_SidFTMS1_UbiE3GE13G_f + gib_GSSG-SidFTMS2_Ubiquitin_r from pMIB5692; plasmid: gib_uni_GSSG-SidFTMS2_f + gib_uni_GSSG-SidFTMS1_r from pMIB7894
pMIB7679	pxDC61-LepB^{HiBiT}	Cm	This study	Plasmid: gib_uni_pxDC61_f + gib_uni_pxDC61_r from pxDC61; Insetr: gib_pxDC61_LepB_f + gib_pxDC61_Hibit_r from pMIB7573
pMIB7680	pxDC61-YidC^{HiBiT}	Cm	This study	Plasmid: gib_uni_pxDC61_f + gib_uni_pxDC61_r from pxDC61; Insetr: gib_pxDC61_YidC_f + gib_pxDC61_Hibit_r from pMIB7572
pMIB7708	pxDC61-SidF^{HiBiT}	Cm	This study	Plasmid: gib_Hibit_SidFc30_f + gib_Hibit_SidF_r from pMIB6814
pMIB7709	pxDC61-LegC3^{HiBiT}	Cm	This study	Plasmid: gib_Hibit_LegC3c30_f + gib_Hibit_LegC3_r from pMIB6815
pMIB7714	pxDC61-SidFΔTMS2^{HiBiT}	Cm	This study	Plasmid: gib_SidFTMS1_SidF_f + gib_uni_SidFTMS1_r from pMIB7708
pMIB7709	pxDC61-LegC3ΔTMS2^{HiBiT}	Cm	This study	Plasmid: gib_LegC3TMS1_LegC3_f + gib_uni_LegC3TMS1_r from pMIB7709
pMIB7751	pxDC61-Lep-inv^{HiBiT}	Cm	This study	Plasmid: gib_uni_pxDC61_f + gib_uni_pxDC61_r from pxDC61; Insert: gib_pxDC61_Lep-inv_f +

				gib_Hibit_Lep-Inv_r from pMIB5505
pMIB7712	pxDC61(kan)	Kan	This study	insert: gib_pxDC61Res_kan_f + gib_pxDC61Res_kan_r from pT10; Plasmid: gib_uni_pxDC61Res_f + gib_uni_pxDC61Res_r from pxDC61
pMIB7703	pxDC61-IcmSW^{3FLAG}	Cm	This study	Insert: gib_pxDC61_IcmS_f + gib_pxDC61_FLAG_r from synthesized oligonucleotide "IcmSW-3FLAG"; plasmid: gib_uni_pxDC61_f + gib_uni_pxDC61_r from pxDC61
pMIB7720	pMIB7703	Kan	This study	insert: gib_pxDC61Res_kan_f + gib_pxDC61Res_kan_r from pT10; Plasmid: gib_uni_pxDC61Res_f + gib_uni_pxDC61Res_r from pMIB7703
pMIB6830(kan)	pxDC61-$^{HiBiT-2HA}$SidF	Kan	This study	insert: gib_pxDC61Res_kan_f + gib_pxDC61Res_kan_r from pT10; Plasmid: gib_uni_pxDC61Res_f + gib_uni_pxDC61Res_r from pMIB6830
psr47S	Suicide plasmid	Kan	Kindly provided by Craig Roy	
psr47SΔ*dotG*	Suicide plasmid for the deletion of DotG	Kan	Kindly provided by Craig Roy	
pMIB7957	psr47S-DotM^{FLAG} +−1000 bp of DotM	Kan	This study	insert1: gib_DotM_3xFLAG_f + gib_DotM-down_3xFLAG_r from pMIB7802; insert2: gib_psr47S_DotM_f + gib_psr47S_DotM-down_r from genomic DNA; plasmid:

				gib_uni_pSR47S_f + gib_uni_pSR47S_r from psr47S
pMIB7830	psr47S-^{FLAG}DotN +−1000 bp of DotN	Kan	This study	insert1: gib_DotN-up_3xFLAG_f + gib_DotN_3xFLAG_r from FLAG-peptide; insert2: gib_psr47S_DotN-up_f + gib_psr47S_DotN-down_r from genomic DNA; plasmid: gib_uni_pSR47S_f + gib_uni_pSR47S_r from psr47S
pMIB7903	psr47S-^{FLAG}DotY +−1000 bp of DotY	Kan	This study	insert1: gib_DotY-up_3xFLAG_f + gib_DotY_3xFLAG_r from FLAG-peptide; insert2: gib_psr47S_DotY-up_f + gib_psr47S_DotY-down_r from genomic DNA; plasmid: gib_uni_pSR47S_f + gib_uni_pSR47S_r from psr47S
pMIB7827	psr47S-^{FLAG}DotZ +−1000 bp of DotZ	Kan	This study	insert1: gib_DotZ-up_3xFLAG_f + gib_DotZ_3xFLAG_r from FLAG-peptide; insert2: gib_psr47S_DotZ-up_f + gib_psr47S_DotZ-down_r from genomic DNA; plasmid: gib_uni_pSR47S_f + gib_uni_pSR47S_r from psr47S
				Made by QuikChange
pMIB7863	pxDC61kan-IcmS$_{X115*,F147X}$-IcmW^{FLAG}	Kan	This study	QC_SidFV147X_f + QC_SidFV147X_r from pMIB6830(kan)
pMIB7864	pxDC61kan-IcmS$_{X115*,L427X}$-IcmW^{FLAG}	Kan	This study	QC_SidFL427X_f + QC_SidFL427X_r from pMIB6830(kan)
pMIB7865	pxDC61kan-IcmS$_{X115*,G620X}$-IcmW^{FLAG}	Kan	This study	QC_SidFG620X_f + QC_SidFG620X_r from pMIB6830(kan)

pMIB7866	pxDC61kan-$IcmS_{X115*,Q721X}$-$IcmW^{FLAG}$	Kan	This study	QC_SidFQ721X_f + QC_SidFQ721X_r from pMIB6830(kan)
pMIB7867	pxDC61kan-$IcmS_{X115*,K742}$-$IcmW^{FLAG}$	Kan	This study	QC_SidFK742X_f + QC_SidFK742X_r from pMIB6830(kan)
pMIB7868	pxDC61kan-$IcmS_{X115*,L762}$-$IcmW^{FLAG}$	Kan	This study	QC_SidFL762X_f + QC_SidFL762X_r from pMIB6830(kan)
pMIB7876	pxDC61-$^{HiBiT-2HA}SidF_{L893X}$	Kan	This study	QC_SidFL893X_f + QC_SidFL893X_r from pMIB6830(kan)
pMIB7877	pxDC61-$^{HiBiT-2HA}SidF_{L894X}$	Kan	This study	QC_SidFT894X_f + QC_SidFT894X_r from pMIB6830(kan)
pMIB7878	pxDC61-$^{HiBiT-2HA}SidF_{L895X}$	Kan	This study	QC_SidFA895X_f + QC_SidFA895X_r from pMIB6830(kan)
pMIB7879	pxDC61-$^{HiBiT-2HA}SidF_{L896X}$	Kan	This study	QC_SidFA896X_f + QC_SidFA896X_r from pMIB6830(kan)

Table 5: List of primer used in this study

Primer name	**Sequence (5' to 3')**
gib_uni_pxDC61_f	GCTTGGCTGTTTTGGCGGATG
gib_uni_pxDC61_r	ATGTATATCTCCTTCTGAATTCTGTTTC
gib_uni_HA_r	AGAGGATCCTGCATAATCAGG
gib_pxDC61_Hibit_f	GGAAACAGAATTCAGAAGGAGATATACATATGGTGAGCGGCTGGCGG
gib_pxDC61_LegC2_r	CTTCTCTCATCCGCCAAAACAGCCAAGCTTAACCTGTGAGAGTTTGAGTTG
gib_pT12_LegC3_r	CATCCGCCAAAACAGCCAAGCTTTACGCTATCTCATTAACTGTTTC
gib_pxDC61_SidF_r	CTTCTCTCATCCGCCAAAACAGCCAAGCTTAGAAGTTTACTGGCGTGG
gib_pT12_Ceg4_r	CATCCGCCAAAACAGCCAAGCTCTATGAATTATGGAGTGGATTAG

gib_pxDC61_SidM_r	CTTCTCTCATCCGCCAAAACAGCCAAGCTTATTTTATCTTAATGGTTTGTCTTTC
gib_pxDC61_RalF_r	CTCTCATCCGCCAAAACAGCCAAGCTTAAAATTTTAATTGTCTACC
gib_HA_Lep-inv_f	CATACGATGTTCCAGATTACGCTGGAAGCGGAGCGAATATGAAGAAGAAGTTTGCC
gib_pxDC61_Lep-inv_r	GAAAATCTTCTCTCATCCGCCAAAACAGCCAAGCTTAATGGATGCCGCCAATGCG
gib_pxDC61_LegC2-30aa_r	CATCCGCCAAAACAGCCAAGCTTAACCAAGTACTTTTGCATTACTATATG
gib_pxDC61_LegC3dc30_r	CATCCGCCAAAACAGCCAAGCTTATAATGATGGATCACTGGATTGGG
gib_pxDC61_SidFdc30_r	CTTCTCTCATCCGCCAAAACAGCCAAGCTTATTCAAACTCTTGTTTTGGAGTGGATTG
gib_pxDC61_Ceg4dc30_r	CTTCTCTCATCCGCCAAAACAGCCAAGCCTAGAAATGACCAACCGGATGAATATTG
gib_pxDC61_SidM-30aa_r	CATCCGCCAAAACAGCCAAGCTTATAACCCTAAAAGCTGGGTTGTTAAC
gib_pxDC61_RalFdc30_r	CTTCTCTCATCCGCCAAAACAGCCAAGCTTAAATGAGTTTTTTCTGGTTATCATATGC
gib_uni_SidFC_f	GGTACCGCTATTGATGCTATTAATAACAAG
gib_SidFC_SidF760_r	CTTCTTCTATCTTGTTATTAATAGCATCAATAGCGGTACCTGCTTTAATTTTTGTACCCATATTTTTGG
gib_pxDC61_SidF760_r	GAAAATCTTCTCTCATCCGCCAAAACAGCCAAGCTTATGCTTTAATTTTTGTACCCATATTTTTGG
gib_uni_SidFc30_f	AAACCCAAAGATATGACCTCGAAAG
gib_SidFc30_SidF862_r	GCTGTTAAGTCAACTTTCGAGGTCATATCTTTGGGTTTTTGTTGATGAAGTTTGGCATTCAC
gib_SidFc30_SidF842_r	GCTGTTAAGTCAACTTTCGAGGTCATATCTTTGGGTTTTGAATTTTCTGATGCCGCTTTC
gib_SidFc30_SidF822_r	GCTGTTAAGTCAACTTTCGAGGTCATATCTTTGGGTTTTATCTTGTTATTAATAGCATCAATAGCGG
gib_SidFTMS1_SidF_f	CTTACTTTGCTATTCCCTCCGGCAGGCGGTACCGCTATTGATGCTATTAATAAC
gib_uni_SidFTMS1_r	GCCTGCCGGAGGGAATAGC
gib_uni_SidMc30_f	AAGACAAGTTCAGTGTCTTCATTTG
gib_HA_Lep-inv_f	CATACGATGTTCCAGATTACGCTGGAAGCGGAGCGAATATGAAGAAGAAGTTTGCC

gib_SidMc30_Lep-inv_r	CCTCAACCATTTTCTCAAATGAAGACACTGAACTTGTCTTATGGATGCCGCCAATGCG
gib_Hibit-GSSG_SidFTMS2_f	GAGCGGCTGGCGGCTGTTCAAGAAGATTAGCGGAAGCAGCGGAGCAGGCTTGGCTGTGGCC
gib_GSSG-Hibit_SidFTMS1_r	CTTCTTGAACAGCCGCCAGCCGCTCACTCCGCTGCTTCCCGGAGGGAATAGCAAAGTAAGG
gib_SidFTMS1-GSSG_Ubi_f	GCTATTCCCTCCGGGAAGCAGCGGACAGATCTTTGTGAAGACCCTCAC
gib_GSSG-SidFTMS2_Ubi_r	CAGCCAAGCCTGCTCCGCTGCTTCCTCCACCGCGGAGGCGCAAC
gib_uni_GSSG-SidFTMS2_f	GGAAGCAGCGGAGCAGGC
gib_uni_GSSG-SidFTMS1_r	TCCGCTGCTTCCCGGAGG
gib_SidFTMS1_UbiE3GE13G_f	GCTATTCCCTCCGGGAAGCAGCGGACAGGGCTTTGTGAAGACCC
gib_Hibit_SidFc30_f	GCGGCTGGCGGCTGTTCAAGAAGATTAGCGGAAGCGGAAAACCCAAAGATATGACCTCGAAAG
gib_Hibit_SidF_r	CTAATCTTCTTGAACAGCCGCCAGCCGCTCACTCCGCTTCCTTCAAACTCTTGTTTTGGAGTGGATTG
gib_Hibit_LegC3c30_f	GCGGCTGGCGGCTGTTCAAGAAGATTAGCGGAAGCGGATCCACGCATTCAGAAGAGAGCG
gib_Hibit_LegC3_r	CTAATCTTCTTGAACAGCCGCCAGCCGCTCACTCCGCTTCCTAATGATGGATCACTGGATTGGG
gib_LegC3TMS1_LegC3_f	CTTTATATTTTGGTCTCAGCGCCTCCACATTCCAATGAATCAGCTATTAAATC
gib_uni_LegC3TMS1_r	TGGAGGCGCTGAGACCAAAATATAAAG
gib_pxDC61_Lep-inv_f	CAATTTCACACAGGAAACAGAATTCAGAAGGAGATATACATATGGCGAATATGAAGAAGAAGTTTG
gib_Hibit_Lep-Inv_r	CCGCTCACtccactcgaaccatggatgccgccaatgcga
gib_pxDC61_IcmS_f	GGAAACAGAATTCAGAAGGAGATATACATATGGAGCGAGATATTAGCAAGTG
gib_pxDC61_FLAG_r	CTTCTGCGTTCTGATTTAATCTGTATCAGTTACTTGTCATCGTCATCCTTG
gib_uni_pxDC61Res_f	TTTTTTTAAGGCAGTTATTGTCTTCAAATTCCCG
gib_uni_pxDC61Res_r	TTTAGCTTCCTTAGCTCCTGAAAATCTC
gib_pxDC61Res_kan_f	GAGATTTTCAGGAGCTAAGGAAGCTAAAATGCTGATTTTATATGGGTATAAATGGG

gib_pxDC61Res_kan_r	CGGGAATTTGAAGACAATAACTGCCTTAAAAAAATTAGAAAAACTCATCGAGCATC
gib_uni_pSR47S_f	GGTGGAGCTCCAGCTTTTG
gib_uni_pSR47S_r	CTGCAGGAATTCGATATCAAGC
gib_DotM_3xFLAG_f	GATTAACACCACGTCAAATGGAGGAATTAGAGCCAGACTACAAAGACCATGACGGTG
gib_DotM-down_3xFLAG_r	GTTGGATCTAATTCATGACGAGAATCAATACCCCGCATTACTTGTCATCGTCATCCTTG
gib_psr47S_DotM_f	CGGTATCGATAAGCTTGATATCGAATTCCTGCAGAGATGGCACAACAACAGCAG
gib_psr47S_DotM-down_r	CTAAAGGGAACAAAAGCTGGAGCTCCACCGCATACTGGAAGAGCCTTGG
gib_DotN-up_3xFLAG_f	GTCAGTAGTTAAGGATAGAACACATAATTTTTTGAAAATTGAATGGACTACAAAGACCATGACGGTG
gib_DotN_3xFLAG_r	GTTGTTGATTATCCGCCATAGTTTGGTTCACATTCAGCTTGTCATCGTCATCCTTGTAATC
gib_psr47S_DotN-up_f	CGGTATCGATAAGCTTGATATCGAATTCCTGCAGATTTCCCGTCATTGGTTTCC
gib_psr47S_DotN-down_r	CTAAAGGGAACAAAAGCTGGAGCTCCACCCAATTTGCCATCCTTTGTCC
gib_DotY-up_3xFLAG_f	CTATATTTCCTGGACCAAATATACATTGTTGAGAGTAAGTTATGGACTACAAAGACCATGACGGTG
gib_DotY_3xFLAG_r	GGCATCTCTTGTGGGCAGTGTGTATTTTGGCATCTTGTCATCGTCATCCTTGTAATC
gib_psr47S_DotY-up_f	CGGTATCGATAAGCTTGATATCGAATTCCTGCAGTGGCTGCTTTACGAAATGTG
gib_psr47S_DotY-down_r	CTAAAGGGAACAAAAGCTGGAGCTCCACCTGCCTGCTATAACCCAAAGG
gib_DotZ-up_3xFLAG	GCCCTTATATCACATACTTTAAAGAATTTAAATTGGATAAAAAATGGACTACAAAGACCATGACGGTG
gib_DotZ_3xFLAG_r	GCCATTGACTCAATTCATCATCTTTTTTGATCTCGTCCATCTTGTCATCGTCATCCTTGTAATC
gib_psr47S_DotZ-up_f	CGGTATCGATAAGCTTGATATCGAATTCCTGCAGTTGCAATTCAGACCAAGCAG
gib_psr47S_DotZ_r	CTAAAGGGAACAAAAGCTGGAGCTCCACCTCGGTTAATCGCATCTTGTG
QC_SidFV147X_f	CTTTCTGGAACAGTTGGCCAAATAGGAAGGCAGTTTACAGCAATTCAC
QC_SidFV147X_r	GTGAATTGCTGTAAACTGCCTTCCTATTTGGCCAACTGTTCCAGAAAG

QC_SidFL427X_f	GTTTAGACAGGTTGGAGCGTAGAATCCGCTCTCTGACAATAGTG
QC_SidFL427X_r	CACTATTGTCAGAGAGCGGATTCTACGCTCCAACCTGTCTAAAC
QC_SidFG620X_f	CTATTATGAGAACACCTGGCATCATTAGAAAAATAATTTCAAGCTGCAAACCTTAATGGC
QC_SidFG620X_r	GCCATTAAGGTTTGCAGCTTGAAATTATTTTTCTAATGATGCCAGGTGTTCTCATAATAG
QC_SidFQ721X_f	CGGCGGAAAGTTTGAAATTTAGGATAAAGCCAATTTCGCTGATAATG
QC_SidFQ721X_r	CATTATCAGCGAAATTGGCTTTATCCTAAATTTCAAACTTTCCGCCG
QC_SidFK742X_f	GCAAGGTCGCTAATTTTGCCAAAATATAGGAAATGGGGCCTGAAAACG
QC_SidFK742X_r	CGTTTTCAGGCCCCATTTCCTATATTTTGGCAAAATTAGCGACCTTGC
QC_SidFL762X_f	CCAAAAATATGGGTACAAAAATTAAAGCAGGTTAGGGGCTTGGGTTACTAGC
QC_SidFL762X_r	GCTAGTAACCCAAGCCCCTAACCTGCTTTAATTTTTGTACCCATATTTTTGG
QC_SidFL893X_f	GATATGACCTCGAAAGTTGACTAGACAGCAGCGATGGAG
QC_SidFL893X_r	CCTCCATCGCTGCTGTCTAGTCAACTTTCGAGGTCATATC
QC_SidFT894X_f	CAAAGATATGACCTCGAAAGTTGACTTATAGGCAGCGATGGAGGATG
QC_SidFT894X_r	GATTATCATCCTCCATCGCTGCCTATAAGTCAACTTTCGAGGTCATATCTTTG
QC_SidFA895X_f	GACCTCGAAAGTTGACTTAACATAGGCGATGGAGGATGATAATCG
QC_SidFA895X_r	CGATTATCATCCTCCATCGCCTATGTTAAGTCAACTTTCGAGGTCATATC
QC_SidFA896X_f	CTCGAAAGTTGACTTAACAGCATAGATGGAGGATGATAATCGAAGCG
QC_SidFA896X_r	CGCTTCGATTATCATCCTCCATCTATGCTGTTAAGTCAACTTTCGAG

3.4.1 Gibson cloning

Gibson assembly (Gibson *et al.*, 2010) of vector and insert DNA fragments was used to construct plasmids. These fragments were amplified by PCR using Gibson primers (Table 5) and Q5 Hot Start High-Fidelity DNA polymerase KOD polymerase or Phusion polymerase. Then, the PCR product was digested with DpnI for at least 2 h at 37 °C and analyzed on a 1 % (w/v) agarose gel in TAE buffer. Amplified vector and insert DNA fragments were assembled in 12.5 µl Gibson master mix comprising DNA ligase, DNA polymerase, and T5 exonuclease for 15-60 minutes at 50 °C. Afterwards, chemically competent NEB5α or Top10 cells were transformed with 5-15 µl of the reaction product containing the

assembled plasmids. Colony PCRs were performed by using *Taq* polymerase. Purification of PCR products and plasmids was performed according to the PCR-purification kit manual (QIAGEN). Plasmids were extracted from an o/n culture using QIAprep Spin Miniprep Kit (Qiagen) and positive mutants were identified by sequencing analysis (Eurofins Genomics).

3.4.2 QuikChange site-directed mutagenesis

In order to enable *in vivo* photo-crosslinking, amber stop codons were introduced in target genes by QuikChange site-directed mutagenesis (QC) (Agilent, 2018).The suicide plasmid psr47s was amplified using KOD Hot Start DNA Polymerase and QC primers (Table 5) according to manufacturer's protocol. Then, the PCR product was digested with DpnI for at least 2 h at 37 °C and chemically competent NEB5α or Top10 cells were transformed with the reaction product using a standard heat shock protocol. Plasmids were extracted from an o/n culture using the QIAprep Spin Miniprep Kit (Qiagen) and positive mutants were identified by sequencing analysis (Eurofins Genomics).

3.4.3 Transformation of electro-competent *L. pneumophila*

Plasmids used in this study were introduced into *Legionella* strains by electroporation. Electro-competence was obtained by washing the cells three times with ice cold Millipore water and once with 10 % (v/v) glycerol. Electroporation was performed with an Eppendorf electroporator at 2.2 kV, 200 Ω and 25 µF. Transformed cells were grown in antibiotic-free ACES-buffered yeast extract broth at pH 6.9 supplemented with 0.4 mg/ml L-cysteine and 0.25 mg/ml ferric pyrophosphate for 4-8 hours. Afterwards, cells were plated on BCYE agar plates with appropriate antibiotics.

3.4.4 Construction of *L. pneumophila* mutants

Chromosomal DNA mutations in *L. pneumophila* were introduced using allelic exchange. The suicide plasmid psr47s was kindly provided by the lab of Craig Roy. The suicide plasmid encodes the mutated gene of interest with approximately ± 1000 bp flanking regions, which are equivalent to the chromosomal destination locus. Cloning and propagation of the psr47s plasmid was performed in CR14 cells.

Oligonucleotides needed for strain construction as well as sequencing primers were produced by Eurofins genomics and are listed in Table 5.

3.4.5 Allelic exchange

Allelic exchange was used to modify genes in *L. pneumophila* strains on the chromosomal level. For this purpose, a suicide plasmid (psr47s) was used. The suicide plasmid encodes the mutated gene of interest with approximately ± 1000 bp flanking regions, which are equivalent to the chromosomal destination locus. Cloning and propagation of the psr47s plasmid was performed in CR14 cells. *L. pneumophila* wild type strains were grown for two days on BCYE agar plates. The donor strain CR14 and the *E. coli* helper strain CR19 were grown on o/n on LB agar plates. The suicide plasmid was mated from *E. coli* into *L. pneumophila* by mixing all three strains on a BCYE agar plate without any antibiotics for 4-8 hours at 37 °C. Mating mixtures were plated on BCYE plates containing kanamycin (15 µg/ml) and streptomycin for 2-4 days to select for *L. pneumophila* that had the suicide plasmid integrated into their chromosome by homologous recombination. Kanamycin-resistant colonies were then plated on BCYE agar plates containing 5 % (w/v) sucrose for 4-5 days to select for bacteria that had lost the plasmid. Single sucrose-resistant colonies were screened by colony PCR and verified by sequencing analysis (Eurofins Genomics).

3.5 *In vivo* photo-crosslinking

In vivo photo-crosslinking is a well-known method to show protein-protein interactions. However, so far nobody has used this method in *L. pneumophila.* It is based on a vector system, which encodes genes for an amber suppressor tRNA and an aminoacyl-tRNA synthetase. The artificial photoreactive amino acid *para*-benzoyl-phenylalanine (*p*Bpa), provided in the culture medium, is loaded onto the suppressor tRNA by the synthetase. Upon introduction of amber stop codons at distinct positions within a gene, *p*Bpa gets incorporated into the protein at those specific positions. After UV_{365} irradiation, covalent bonds (crosslinks) are formed to proteins nearby (Farell *et al.*, 2005; Liu *et al.*, 2007) (principle shown in Fig. 7).

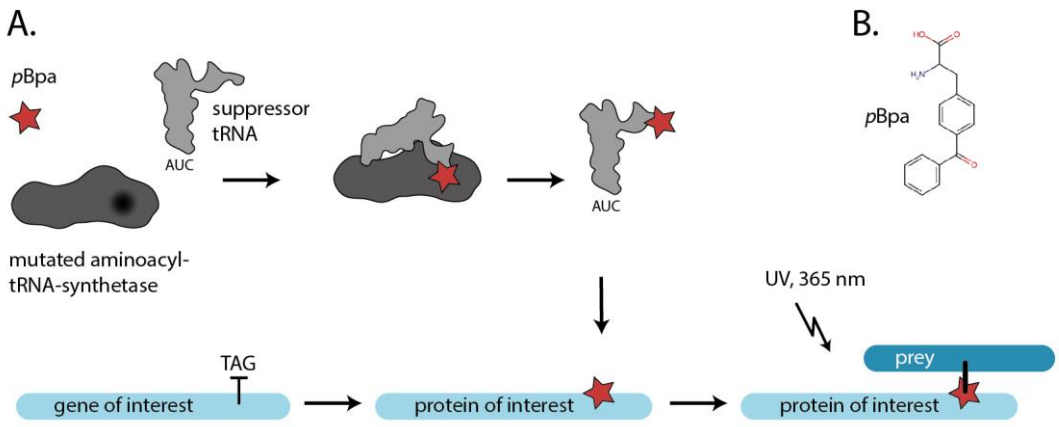


Fig. 7: Principle of the *in vivo* photo-crosslinking method (adopted by Singh and Wagner, 2019)
A) The artificial photoreactive amino acid *para*-benzoyl-phenylalanine (*p*Bpa) **(B)** is loaded on the suppressor tRNA by the tRNA-synthetase. UV irradiation results in covalent binding of *p*Bpa to proteins nearby.

L. pneumophila LPO1 was transformed with pEvol-*p*Bpa and the plasmid pxDC61(Kan) coding for the protein of interest. Bacteria were grown for 2-3 days on BCYE-Agar plates supplemented with "BCYE Growth Supplements" 6,8 µg/ml chloramphenicol and 10 µg/ml kanamycin. Afterwards they were restreaked on BCYE-Agar plates supplemented with "BCYE Growth Supplements", 1 mM para- benzoyl-phenylalanine (*p*Bpa, dissolved in 1 M NaOH, Bachem, Bubendorf, CH), 0.5 mM IPTG, 6,8 µg/ml chloramphenicol and 10 µg/ml kanamycin and incubated o/n at 37 °C. 8 ODU were harvested and resuspended in 2 ml ice cold PBS. For cross-linking of proteins, 1 ml of the sample was irradiated with ultraviolet light (UV) at λ = 365 nM on a CELLSTAR® 6 well cell culture plate (Greiner Bio-One) for 30 minutes (+UV). The other 1 ml cells were left untreated (-UV). Cells were pelleted by centrifugation at maximum speed in a tabletop centrifuge for 2 minutes at 4 °C and subjected to sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

3.6 Crude membrane preparation

8 ODU were harvested by centrifugation at 6000 x *g* for 15 minutes at 4 °C. The cell pellet was resuspended in 750 µl buffer K supplemented with 10 µg/ml DNase I, 2 mg/ml lysozyme, 1 mM EDTA and 1 mM $MgCl_2$, 75 µl of a protease inhibitor cocktail, and incubated for 30 min at 4 °C. Samples were lysed in a SpeedMill PLUS (Analytic Jena) (continuous mode, 2 min, 2 repetitions) and cell debris and aggregates were removed by centrifugation for 20 min at 20,000 x *g*. Crude membranes in the supernatant were pelleted at 52,000 rpm and 4 °C for 45 min in the Beckman TLA-55 rotor. Samples

were stored at -20 °C or directly subjected to SDS- or blue native polyacrylamide gel electrophoresis (BN PAGE).

3.7 Membrane fractionation

Approximately 1000 ODU grown in 300 ml culture were harvested in a JLA 8.1 Beckman rotor for 10 minutes at 6000 x *g*. The cell pellet was washed once wit ice-cold PBS and then resuspended in 15 ml buffer K supplemented with 10 µg/ml DNase, 2 mg/ml lysozyme, 1 mM EDTA and 1 mM $MgCl_2$, 150 µl of a protease inhibitor cocktail, 100 µg/ml gentamycin, and incubated for 30 min at 4 °C. Samples were lysed using a French press (Thermo Spectronic) at $1.1x10^6$ hPa. Debris were removed by centrifugation at 8,000 x *g* for 20 min at 4 °C. Crude membranes were centrifuged at 42,000 rpm for 45 min (70.1 Ti Beckman rotor) and homogenized in buffer M. Inner and outer membranes were separated in a Beckman SW41 Ti rotor for 15 h at 41,000 rpm on a sucrose gradient (50 % - 30 % w/w) created on a gradient station (Biocomp, Fredericton, Canada). 12 fractions were collected and protein amounts were measured using a BCA assay (Thermo Pierce). Proteins of the inner and outer membrane were analyzed by SDS-PAGE followed by Western blotting or mass spectrometry.

3.8 Urea extraction

Membrane samples were solubilized in 8 M urea (dissolved in 1x buffer M) for one hour at room temperature and centrifuged for 1.5 h at 23 °C in the Beckman TLA-55 rotor at 50,000 rpm.

3.9 SDS-PAGE

Proteins were denaturated by incubation in SB buffer for 10 minutes at 50 °C and subsequently separated according to their molecular weight by using ServaGel TG Prime 8-16 % (v/v) precast gels. As a reference Precision Plus Protein All Blue standards (Biorad) was used.

3.10 BN PAGE

Crude membrane of 3 ODU cells were resuspended in 80 µl PBS and homogenization was achieved by extensive pipetting. A final concentration of 1 % (w/v) lauryl maltose neopentyl glycol (LMNG, Anatrace) was added and the membrane proteins were solubilized by shaking at 500 rpm and 4 °C for

one hour. Unsolubilized material was removed by centrifugation at 50,000 rpm at 4 °C for 30 minutes (Beckman TLA-55 rotor). The supernatant was mixed with blue native loading buffer and native protein complexes were separated according to their molecular weight by using Native PAGE™ 3-12 % (v/v) Bis-Tris protein gels (Invitrogen). Electrophoresis was performed in anode buffer and cathode buffer I at 130 V, 200 mA, at 4 °C for 50 minutes. Then, the cathode buffer I was exchanged with cathode buffer II and electrophoresis was continued at 300 V, 200 mA, 4 °C for 90-110 minutes. Blue native gels were directly destained with 5 % (v/v) acetic acid solution or equilibrated in SDS running buffer for 20 minutes and subjected to Western blot analysis.

3.11 Coomassie staining of proteins

After protein separation via SDS-PAGE, gels were washed 3 times with ddH_2O. Proteins were stained o/n with Coomassie solution

3.12 Western blot analysis and immunodetection

After separation by SDS or BN PAGE, proteins were blotted on a polyvinylidene difluoride (PVDF) membrane (Bio-Rad) using a standard wet blot procedure. The transfer was performed at 35 V and 4 °C for 2.5 - 3.5 hours. Residual Coomassie of blotted BN gels was removed by several washing steps with methanol. To block unspecific binding sites, the membrane was incubated for 1 hours at RT in Blue Block PF (Serva) in TBS-T. Next, the membrane was incubated in a primary and secondary antibody solution for at least one hour at RT or o/n at 4 °C with several washing steps in between. Finally, the membrane was scanned with a Li-COR Odyssey system (ODY3191) to visualize proteins. Image analysis was performed with Image Studio 2.1 (Li-COR).

3.13 Mass spectrometry analysis

Identification of crosslinking positions was done by a LC-MS/MS approach in cooperation with the Proteome Center Tübingen (Dr. Mirita Franz-Wachtel).

3.14 Split Nanoluc-based translocation assay

In literature there can be found two assays in order to analyze effector translocation from *L. pneumophila* into host cells. On the one hand, effector proteins are fused to the enzymes β-lactamase and strains expressing this fusion protein are used for host cell infection. Host cells are loaded with CCF4/AM containing a coumarin and a fluorescein fluorophore and emit green fluorescence at 529 nm due to fluorescence resonance energy transfer (FRET) when exited at 409 nm. Successfully translocated β-lactamase cleaves the β-lactam ring of CCF4/AM changing the fluorescence emission from green to blue (447 nm). If the ratio of blue to green fluorescence is greater than 1 a β-lactamase effector fusion protein is considered as successfully translocated (De Felipe *et al.*, 2008). On the other hand, the enzyme adenylate cyclase (Cya) is used in order to measure translocation into host cells. Effector proteins are fused to the calmodulin-dependent adenylate-cyclase domain of the adenylate cyclase (CyaA) which gets activated by calmodulin once successfully translocated into the host cell. The subsequent conversion of ATP to cAMP can be measured with an enzyme-linked immunosorbent assay kit.

While the secretion of β-lactamase-fusions can be performed in a 96 well format and quantified by fluorescence, it lacks in sensitivity and cannot be used for measuring translocation kinetics. The most frequently used Cya-assay is much more time consuming due to the fact that samples have to be processed after infection. Westerhausen *et al.* showed the advantages of using a luciferase-based translocation assay (Westerhausen *et al.*, 2020). NanoLuc is a small and very bright luciferase. To minimize the size and possible interference of the attached reporter even more, a split NanoLuc approach was developed. Split NanoLuc is composed of the large fragment LgBiT which is 18 kDa and comprises most of the proteins β-barrel as well as a small fragment called HiBiT, which is only 1.3 kDa in size and has a very high affinity to LgBiT. Advantages of this assay are that first of all, there is no product accumulation of the luciferase reaction. The measured luminescence signal is directly proportional to the amount of accumulated translocated protein. Moreover, the luciferase-based assay allows a simple, quick, quantitative and highly specific as well as sensitive assessment of protein translocation over time.

DrkBiT was added in the infection buffer in order to quench every luminescence signal which may occur due to the presence of LgBiT outside of macrophages after cell lysis (Fig. 8). Similar to HiBiT, DrkBiT is a peptide that binds to free LgBiT with an equally high affinity but prevents catalysis once bound (Yamamoto *et al.*, 2019).

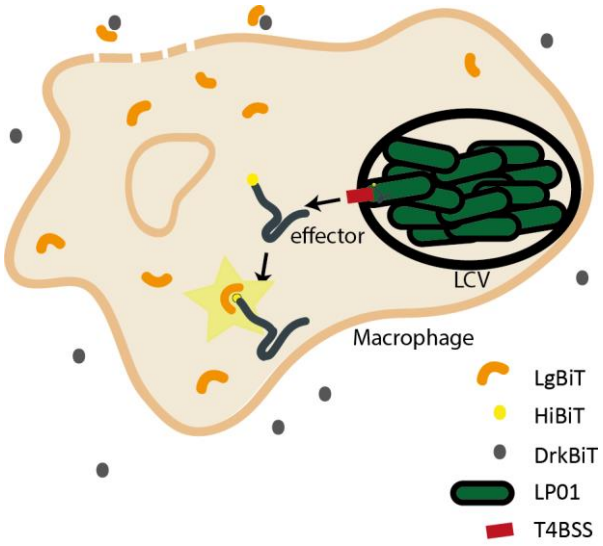


Fig. 8: Principle of a split NanoLuc translocation assay.
The principle of effector translocation into Raw 264.7 macrophage detected by luminescence measurement. LgBiT (orange) expressed in the host cell interacts with translocated HiBiT(yellow)-effector fusion proteins to form functional luciferase. Extracellular LgBiT due to cell lysis is quenched by the addition of DrkBiT (black).

The translocation into host cells of N-terminally HiBiT tagged effector proteins can by followed by using a split NanoLuc translocation assay. One day prior to infection, Raw 264.7 macrophages (kindly provided by Erwin Bohn) expressing LgBiT were seeded in a white clear-bottom 96 well cell culture plate (Greiner) at 8 x 10^4 cells/well.

L. pneumophila strains carrying different HiBiT-tagged effector proteins were grown for 2-3 days on BCYE-Agar plates supplemented with "BCYE Growth Supplements" and appropriate antibiotics. Gene expression was induced by incubation of bacteria o/n at 37 °C on BCYE agar plates including 0.5 mM IPTG. 100 µl of bacteria resuspended in HBSS + DrkBiT (1:1000) at 1.6x10^8 cfu/ml were used to infect Raw 264.7 macrophages (MOI = 200). After centrifugation at 300 *x g* for 8 minutes 25 µl of NanoGlo® live cell buffer (Promega) supplemented with 1:20 of the extended live cell substrate Endurazine™ (Promega) was added to each well. Luminescence signal was measured at 37 °C and 5 % CO_2 every 15 minutes for 12 hours by plate reader (Tecan Spark)

3.15 Split Nanoluc-based topology assay

L. pneumophila strains carrying different HiBiT-tagged proteins were grown for 2-3 days on BCYE-Agar plates supplemented with "BCYE Growth Supplements" and appropriate antibiotics. Gene expression was induced by incubation of bacteria o/n at 37 °C on BCYE agar plates including 0.5 mM IPTG. Bacteria

were resuspended in PBS and the OD was adjusted to 1. 0.5 ODU were harvested at 6000 x *g* for 5 minutes at 4 °C. The cell pellet was resuspended in either 500 µl PBS containing 10 µg/ml DNase I and 1 mM $MgCl_2$ or 500 µl PBS supplemented with 10 µg/ml DNase, 2 mg/ml lysozyme, 1 mM EDTA and 1 mM $MgCl_2$, 75 µl of a protease inhibitor cocktail, and 0.5 % (v/v) Triton-X 100. All samples were incubated for 30 min at 4 °C. The luminescence was measured in a 384 well plate using NanoGlo® live cell buffer (Promega) supplemented with 1:50 of either a membrane impermeable substrate (Promega), for the unlysed samples, or the lytic substrate, for the lysed samples.

3.16 Bioinformatics

TOPCONS was used to search for TMDs in T3SS substrates (http://topcons.cbr.su.se/) (Tsirigos *et al.*, 2015). Predictions of the hydrophobicity ΔG levels of the TMDs were analyzed with the ΔG predictor program (httpµ//dgpred.cbr.su.se/) (Hessa, Nadja M Meindl-Beinker, *et al.*, 2007). Protter was used for visualization of the TMDs in a membrane (http://wlab.ethz.ch/protter/start/).

4 Results

So far, more than 300 effectors of the Dot/Icm system have been identified in *Legionella pneumophila* strain Philadelphia-1. They play an important role in the modulation and manipulation of host cell functions during the course of infection. For some of these effectors it has been reported that they contain transmembrane segments (TMS) and thus are localized within host cell membranes (Vieira *et al.*, 2001; Kamen, Levinsohn and Swanson, 2007; De Felipe *et al.*, 2008; Dolezal *et al.*, 2012; Isaac *et al.*, 2015). Although for some TMD-effectors their functions within the host cell are well understood, it is still unclear how these effectors are successfully targeted to the Dot/Icm system. Similar to soluble effectors, TMD-effectors are characterized by their C-terminal T4 signal. Additionally, they harbor one or more hydrophobic stretches in order to be integrated into a host cell membrane after translocation. Before leaving the bacterial cell, TMD-effectors could be recognized by SRP which may result in inner membrane targeting prior to translocation into the host cell.

In this study, I investigated the targeting, recognition and translocation mechanisms of TMD-effectors in *Legionella pneumophila* strain Philadelphia-1. I determine if TMD-effectors are, similarly to soluble effectors, directly targeted and recognized by the Dot/Icm system or if they follow a hypothetical two-step mechanism with an inner membrane intermediate. To this end, I focused on their similarities as well as differences to well-known soluble effectors. Moreover, I investigated on how TMD-effectors are recognized by the Dot/Icm system and which substrate-intrinsic properties are important for their successful translocation into host cells.

4.1 Targeting of TMD-effectors to the Dot/Icm system in *L. pneumophila*

So far, some TMD-effector such as LegC2, LegC3, Ceg4 and SidF were investigated regarding their function within the host cell but not much is known of their translocation process. Therefore, I determined if the basic characteristics of effector proteins of the Dot/Icm system such as dependency on a C-terminal T4 signal as well as the possible interference of the Dot/Icm chaperones IcmSW also apply to TMD-effectors. Moreover, I elucidated the SRP targeting potential to the bacterial inner membrane of these TMD-effectors in order to assess a possible two-step secretion mechanism.

4.2 Selection of TMD-effectors for in-depth analysis

In order to investigate the targeting and translocation mechanism of TMD-effectors, I selected the four TMD-effectors LegC2, LegC3, SidF and Ceg4 to be studied in depth. First of all, I investigated the presence of a C-terminal T4 signal as well as hydrophobic TMS.

Although no general sequence-based motif is known for effector proteins of the Dot/Icm system, Huang *et al.* found a glutamic acid-rich sequence they called "E-block" motif which is present in a subgroup of effectors. Such an E-block motif is present in the selected TMD-effectors LegC2, LegC3, SidF and Ceg4 as well as the soluble effectors SidM and RalF, which were selected for comparison in their translocation behavior (Fig. 9A). All selected effector proteins were analyzed in respect of their translocation behavior with and without their C-terminal signal. The results are described in the next chapter.

Furthermore, LegC2, LegC3, SidF and Ceg4 were analyzed for TMS using the ΔG predictor program (http://dgpred.cbr.su.se/) (Hessa et al, 2007). Fig. 9B shows the predicted TMS and topology in the host cell membrane of the four selected Dot/Icm TMD-effectors. They are all predicted to have two TMS with only a very small loop of few amino acids in between. The quantification of all TMD-effectors in *L. pneumophila* and the ΔG_{app} values of the TMS of LegC2, LegC3, Ceg4 and SidF, respectively, are described and discussed in chapter 4.4.2.

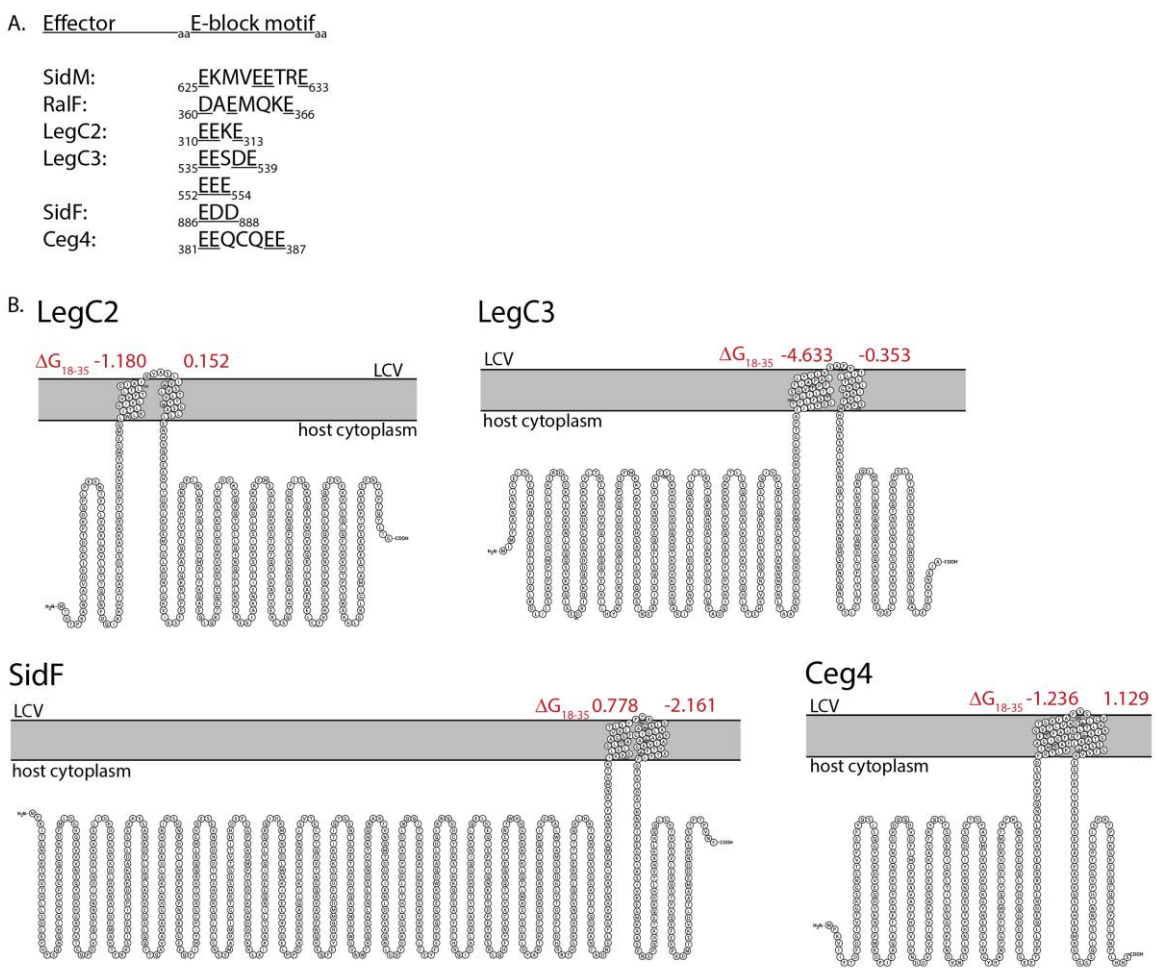


Fig. 9: C-terminal signal (E-block motif) and TMSs in selected substrates of the Dot/Icm system in *Legionella pneumophila.*

A) The C-terminal 30 amino acids of the selected TMD-effectors LegC2, LegC3, SidF and Ceg4 as well as the soluble effectors SidM and RalF were analyzed for the presence of an E-block motif according to Huang *et al.* (Huang *et al.*, 2011).

B) The sequence of the TMD-effectors LegC3, SidF and Ceg4 were analyzed using the full protein scan of the ΔG predictor (http://dgpred.cbr.su.se/) (Hessa, Nadja M Meindl-Beinker, *et al.*, 2007) (ΔG predictor settings: window of 18-35 for LegC3, SidF and Ceg4 and 18-25 for LegC2, length correction: ON). Visualization of the protein topology was done using Protter (http://wlab.ethz.ch/protter/start/).

4.3 Do TMD-effectors behave differently in their translocation in comparison to soluble effectors? (Effector characteristics)

In general, effector proteins of the Dot/Icm system in *L. pneumophila* are characterized by their C-terminal T4 signal, on which they depend for successful translocation into host cells (Vergunst *et al.*, 2000; Luo and Isberg, 2004). Furthermore, most of the effectors which have been studied in-depth so far depend on the presence of the Dot/Icm chaperones IcmS and IcmW (Bardill, Miller and Vogel, 2005; Cambronne and Roy, 2007; Jeong, Sutherland and Vogel, 2015; Xu *et al.*, 2017). To test, if this also applies to TMD-effectors, I used a split NanoLuc-based translocation assay to investigate the translocation efficiency of TMD-effectors *in vivo* with and without their C-terminal 30 amino acids as well as in single deletion mutants of *icmS* and *icmW*.

4.3.1 Development of a split NanoLuc-based translocation assay

In order to analyze the translocation behavior of different effectors of interest, I established a split NanoLuc-based translocation assay with which I could quantify effector translocation by luminescence measurement. Thus, the HiBiT peptide was fused to the N-terminus of the selected Dot/Icm substrates which were expressed from an IPTG-inducible low-copy number plasmid for 20 hours on BCYE plates supplemented with IPTG. Raw 264.7 macrophages constitutively expressing LgBiT were seeded in a 96 well plate and infected with a MOI of 200.

Fig. 10A depicts the luminescence measurement of the soluble effector SidM in a *L. pneumophila* wild type strain as well as in a ΔT4SS knock out mutant over 12 hours. A time frame of 12 hours was chosen to analyze effector translocation in early (< 6 hours post infection (hpi)) as well as later stages of infection without including the transition from replicative to transmissive phase (> 18 hpi). Hereafter, only the effector translocation after 4 and if relevant after 12 hours post infection will be shown.

The Dot/Icm dependent translocation of SidM into Raw 264.7 macrophages was followed over time with a high signal to noise ratio and reached its maximum approximately 8 hpi. HiBiT-LgBiT interaction was shown to be very specific to translocated SidM as no luminescence signal could be detected in a ΔT4SS strain.

To further test the newly established split NanoLuc-based translocation assay in *L. pneumophila*, I included the TMD-effectors LegC2 and LegC3 and other knock-out mutants of the Dot/Icm system such as a deletion mutant of the central channel protein DotG as well as single knock-outs of the ATPases DotO and DotB (Fig. 10B). As expected, no translocation took place for the tested effectors,

independently if they are soluble or have TMS, when essential proteins of the Dot/Icm system were not present.

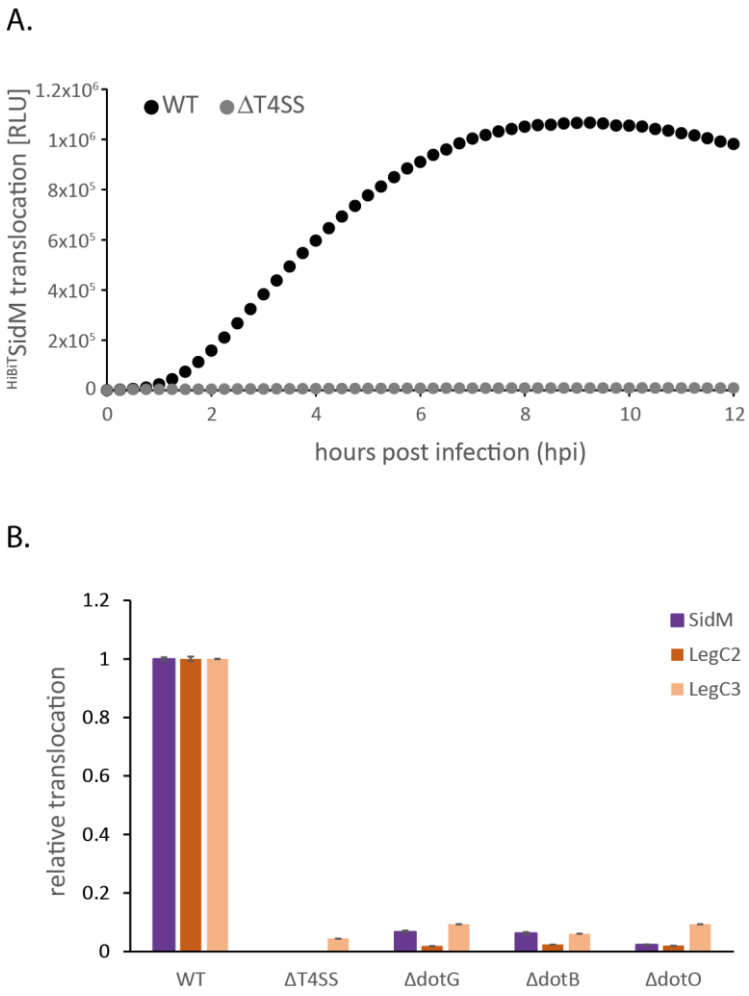


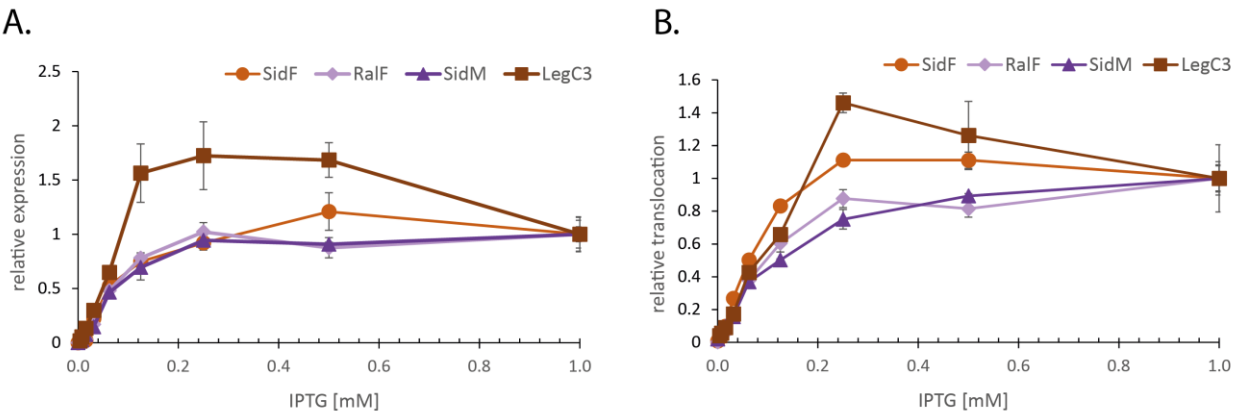
Fig. 10: Development of a split NanoLuc translocation assay.
A) *Legionella pneumophila* wild type and a ΔT4SS mutant strain expressing the soluble Dot/Icm effector protein ^{HiBiT}SidM were used to infect Raw 264.7 macrophages which constitutively express LgBiT. Luminescence was measured over time using a plate reader. A representative result of at least three independent experiments is shown.
B) *L. pneumophila* wild type, a ΔT4SS mutant strain and strains lacking *dotG*, *dotB* or *dotO*, respectively, expressing the Dot/Icm effector proteins ^{HiBiT}SidM, ^{HiBiT}LegC2 or ^{HiBiT}LegC3 were used to infect Raw 264.7 macrophages which constitutively express LgBiT. Luminescence was measured over time using a plate reader. Only the Luminescence signal after 4 hours post infection is shown. The data points represent the mean (± standard deviation) of at least three independent experiments.

Overall, the split NanoLuc-based translocation assay allowed the quantitative measurement of protein translocation into host cells in a very quick and simple manner and showed reproducible results with a high signal to noise ratio.

Although the assessment of effector translocation based on luminescence measurement seems to be straightforward, different influencing factors must be considered.

First of all, effector protein expression might directly influence effector translocation unless effector translocation is the rate limiting step. To test this, effector expression was titrated using different IPTG concentrations and protein translocation was again analyzed using the split NanoLuc-based translocation assay. Additionally, protein expression within the bacteria was measured after bacterial lysis and complementation with purified LgBiT. As expected, protein expression increases with increasing IPTG concentrations until being saturated around 0.25-0.5 mM IPTG (Fig. 11A). With higher IPTG concentrations of 1 mM a reduction of the luminescence signal could be observed for the membrane proteins LegC3 and SidF. Because of the hydrophobic nature of these proteins, overexpression might result in the formation of inclusion bodies or increased proteolysis. Translocation levels, however, seem to be directly proportional to the expression levels of the protein (Fig. 11C).

Based on these results, protein expression was induced with 0.5 mM IPTG and the translocation signal was normalized to the respective expression levels when proteins with different expression levels were compared.



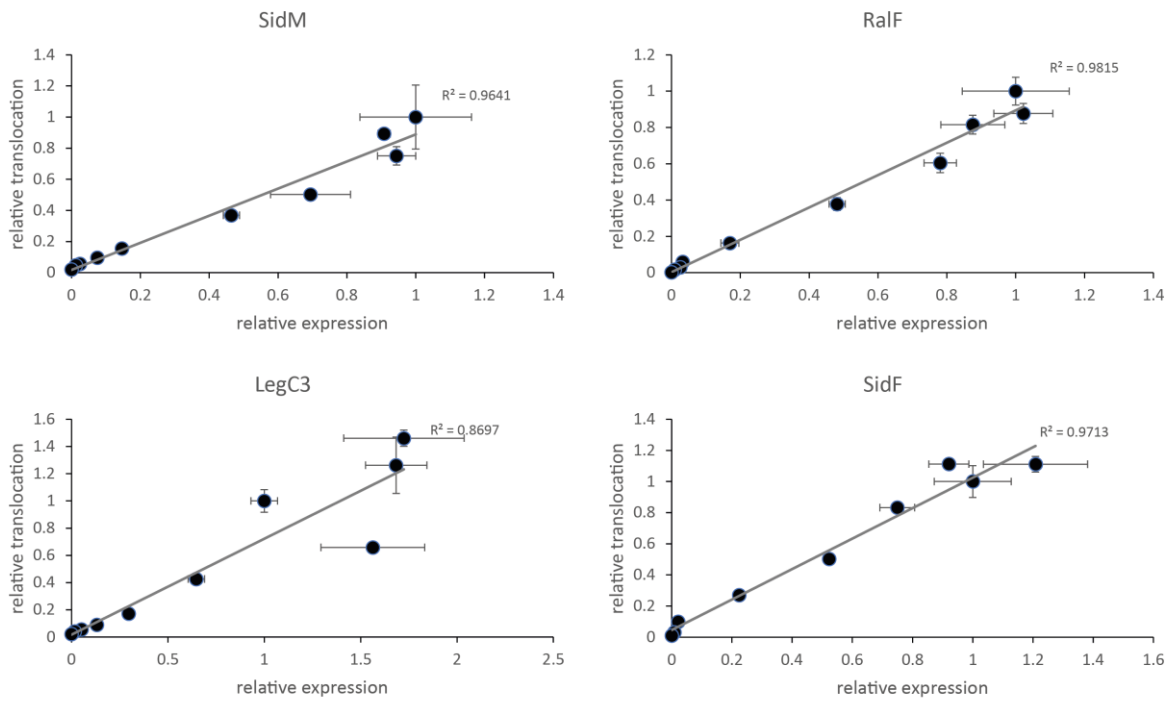


Fig. 11: Translocation levels of effector proteins are directly proportional to the expression levels of the protein.

Legionella pneumophila wild type strain expressing the soluble Dot/Icm effector proteins ^{HiBiT}SidM and ^{HiBiT}RalF or the TMD-effector ^{HiBiT}LegC3 or ^{HiBiT}SidF, respectively were used to infect Raw 264.7 macrophages constitutively expressing LgBiT. A representative result of at least three independent experiments is shown. All values in respect to one protein were normalized to the luminescence measured after induction with 1 mM IPTG.

A) Luminescence of effector expression was measured after whole cell lysis.

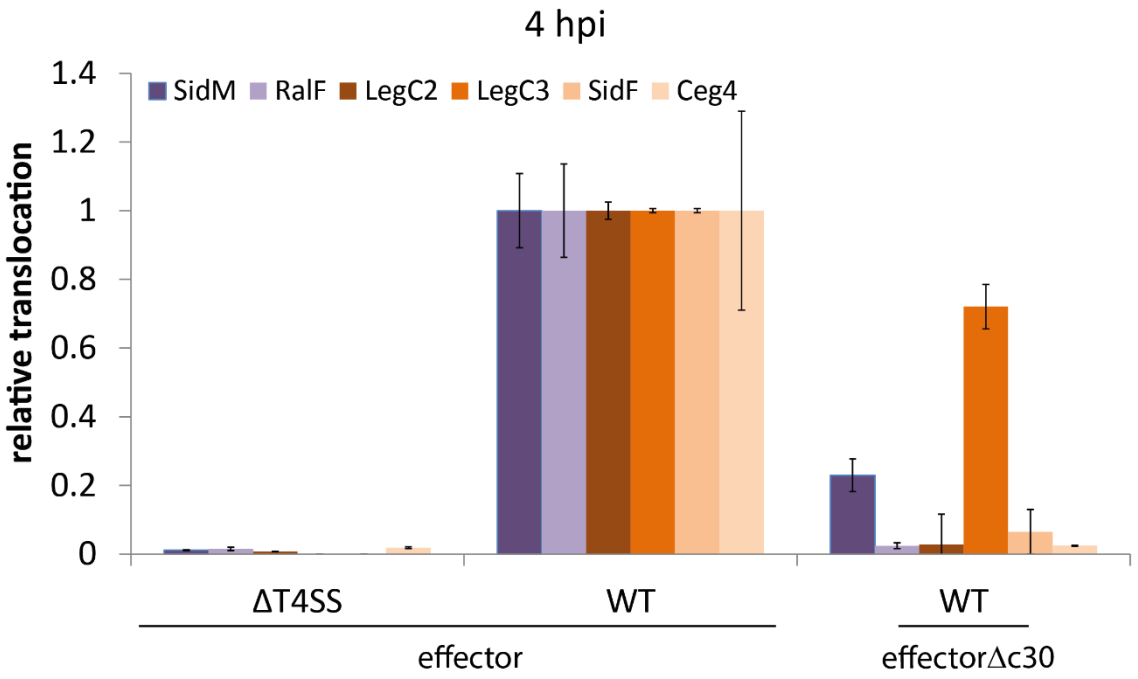
B) Luminescence of translocated effectors was measured 4 hours post infection using a plate reader.

C) Expression and translocation levels of effector proteins were plotted.

Another factor which can indirectly influence translocation signals is the intracellular growth rate of *L. pneumophila*. As *L. pneumophila* is only able to replicate inside host cells if it has a functional Dot/Icm system, mutations within the machinery may influence the intracellular survival of *L. pneumophila* and by this effector expression. As already shown in various studies, single deletions of different structural components of the Dot/Icm machinery very often result in a growth defect of *L. pneumophila* within host cells due to stability issues or insufficient effector translocation (Coers *et al.*, 2000; Sutherland *et al.*, 2013; Chauhan and Shames, 2021). In this context, it is not always clear whether decreased translocation signals are due to the loss of specific interactions of an effector protein with a component of the Dot/Icm system or due to general problems with effector translocation and subsequent growth defects. Thus, survival assays of relevant mutants were performed and will be discussed when relevant.

4.3.2 TMD-effectors cannot be secreted without their C-terminal secretion signal

Although LegC2, LegC3, SidF and Ceg4 are listed as effectors of the Dot/Icm system in *L. pneumophila* in the Secret4 database (Bi *et al.*, 2013)(http://db-mml.sjtu.edu.cn/SecReT4/) and an E-block motif has been identified within their C-terminal 30 amino acids (Fig. 9A), no one ever tested if deletion of their C-terminus results in a translocation deficiency. Thus, I analyzed the translocation efficiency of these four selected TMD-effectors as well as of the soluble effectors SidM and RalF with and without their C-terminal 30 amino acids (Δc30).



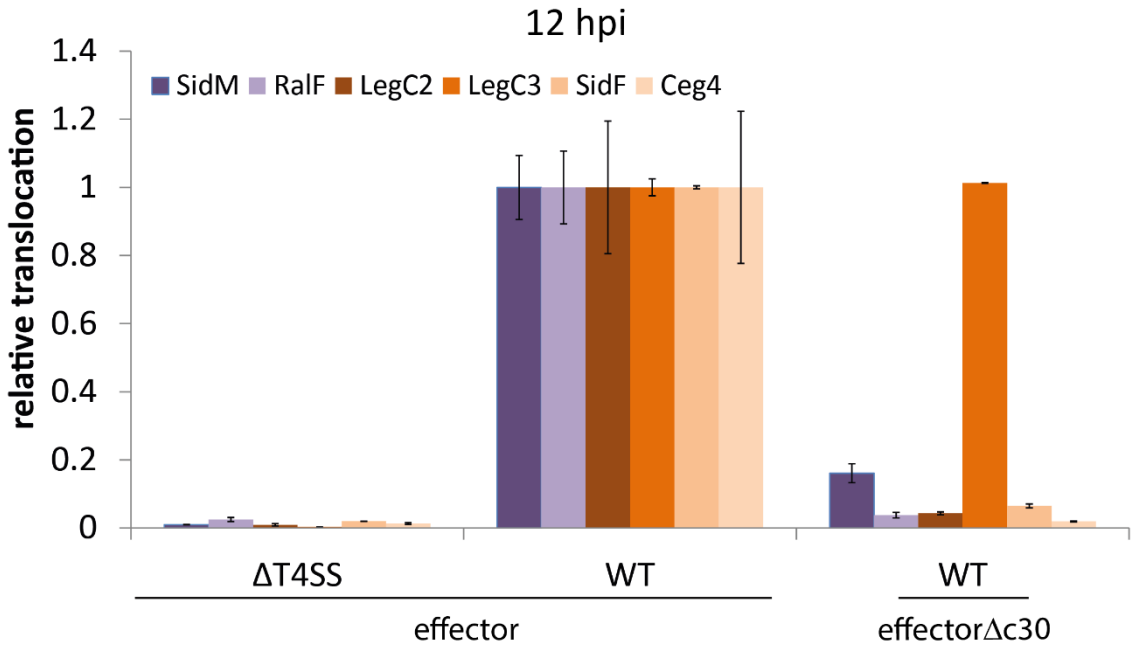


Fig. 12: TMD-effectors depend on the presence of their C-terminal translocation signal.
Legionella pneumophila wild type and a ΔT4SS mutant strain expressing the Dot/Icm effector proteins ^{HiBiT}SidM, ^{HiBiT}RalF, ^{HiBiT}LegC2, ^{HiBiT}LegC3, ^{HiBiT}SidF or ^{HiBiT}Ceg4 as well as the respective effectors lacking their C-terminal translocation signal were used to infect Raw 264.7 macrophages which constitutively express LgBiT. Luminescence was measured over time using a plate reader. The luminescence signal 4 hpi (upper panel) and 12 hpi (lower panel) is displayed. The data points represent the mean (± standard deviation) of at least three independent experiments. hpi hours post infection

As already shown by Nagai *et al.* in 2002 (Nagai *et al.*, 2002), I was able to confirm that RalF cannot be translocated into host cells without its C-terminus present (Fig. 12). The same effect was observed for LegC2, SidF and Ceg4. SidMΔc30 was translocated at levels of about 20 %. Surprisingly, LegC3Δc30 showed only a small decrease in translocation efficiency 4 hours post infection and behaved like the full-length protein 12 hours post infection (Fig. 12, lower panel), indicating that LegC3 must have another internal signal. All proteins showed similar expression in the wild type as well as in the mutant strains (appendix 7.3), indicating that reduced translocation levels are not due to lower protein amounts. In general, however, these results indicate that the C-terminal signal of TMD-effectors is, similarly to soluble effectors, essential for successful effector translocation into host cells.

4.3.3 Translocation of the TMD-effectors LegC2, LegC3, SidF and Ceg4 depends on the T4 chaperones IcmS and IcmW

As the chaperones IcmSW play an important role in the translocation process of many effectors (Amyot, DeJesus and Isberg, 2013; Jeong, Sutherland and Vogel, 2015; Xu *et al.*, 2017), I tested how single deletion mutants of *icmS* or *icmW* influence the translocation of the TMD-effectors LegC2, LegC3, SidF and Ceg4, respectively. Similar to the soluble effectors SidM and RalF, all four TMD-effectors showed a decrease in translocation of up to 80 % when IcmS or IcmW was not present (Fig. 13). This suggests that also the recruitment of TMD-effectors to the Dot/Icm system depends on these Dot/Icm chaperones.

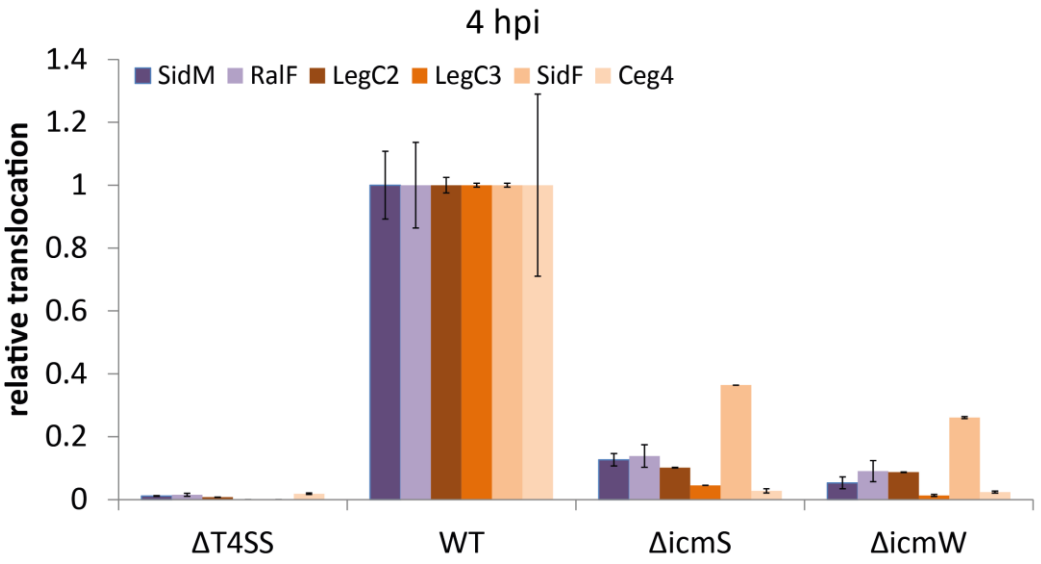


Fig. 13: TMD-effectors depend on the presence of the Dot/Icm chaperones IcmS and IcmW.
Legionella pneumophila wild type and a ΔT4SS mutant strain expressing the Dot/Icm effector proteins ^{HiBiT}SidM, ^{HiBiT}RalF, ^{HiBiT}LegC2, ^{HiBiT}LegC3, ^{HiBiT}SidF or ^{HiBiT}Ceg4 were used to infect Raw 264.7 macrophages which constitutively express LgBiT. Luminescence was measured over time using a plate reader. The luminescence signal 4 hpi is displayed. The data points represent the mean (± standard deviation) of at least three independent experiments. hpi hours post infection

In summary, the selected TMD-effectors LegC2, Ceg4 and SidF need their C-terminal T4 signal as well as the T4 chaperones IcmSW to be successfully translocated into the host cell and with this showed a similar behavior to the soluble effectors SidM and RalF.

However, it is still unclear whether IcmSW already act in the targeting process, and by this might be able to prevent a possible SRP binding by shielding the TMS of TMD-effectors, or simply recognize the effector proteins at the coupling complex.

4.4 How is the SRP targeting potential of TMD-effectors (IM protein characteristics)

After showing that TMD-effectors behave similarly to soluble effectors regarding basic characteristics of translocation by the Dot/Icm system such as C-terminal signal and IcmSW dependence, I analyzed the hydrophobic character of TMD-effectors. Unlike soluble effectors, TMD-effectors harbor one or more hydrophobic TMS, in addition to their C-terminal T4 signal, which might be recognized by SRP and subsequently result in inner membrane insertion.

4.4.1 Quantification of TMD-effectors in *L. pneumophila*

In order to determine if the presence of a TMS within an effector protein of the Dot/Icm system can result in inner membrane insertion and thus in either a potential two-step secretion mechanism or a dead end mistargeting, the membrane targeting potential of TMD-effectors by SRP was analyzed. As shown in Krampen *et al.*, initial analyses were entirely conducted using bioinformatics as well as well-established *E. coli*- based model systems (Krampen *et al.*, 2018). Briefly, for the quantification of TMD-effectors, I included all 286 substrates of the Dot/Icm system in *Legionella pneumophila* strain Philadelphia-1 listed in the T4 database Secret4 (Bi *et al.*, 2013).

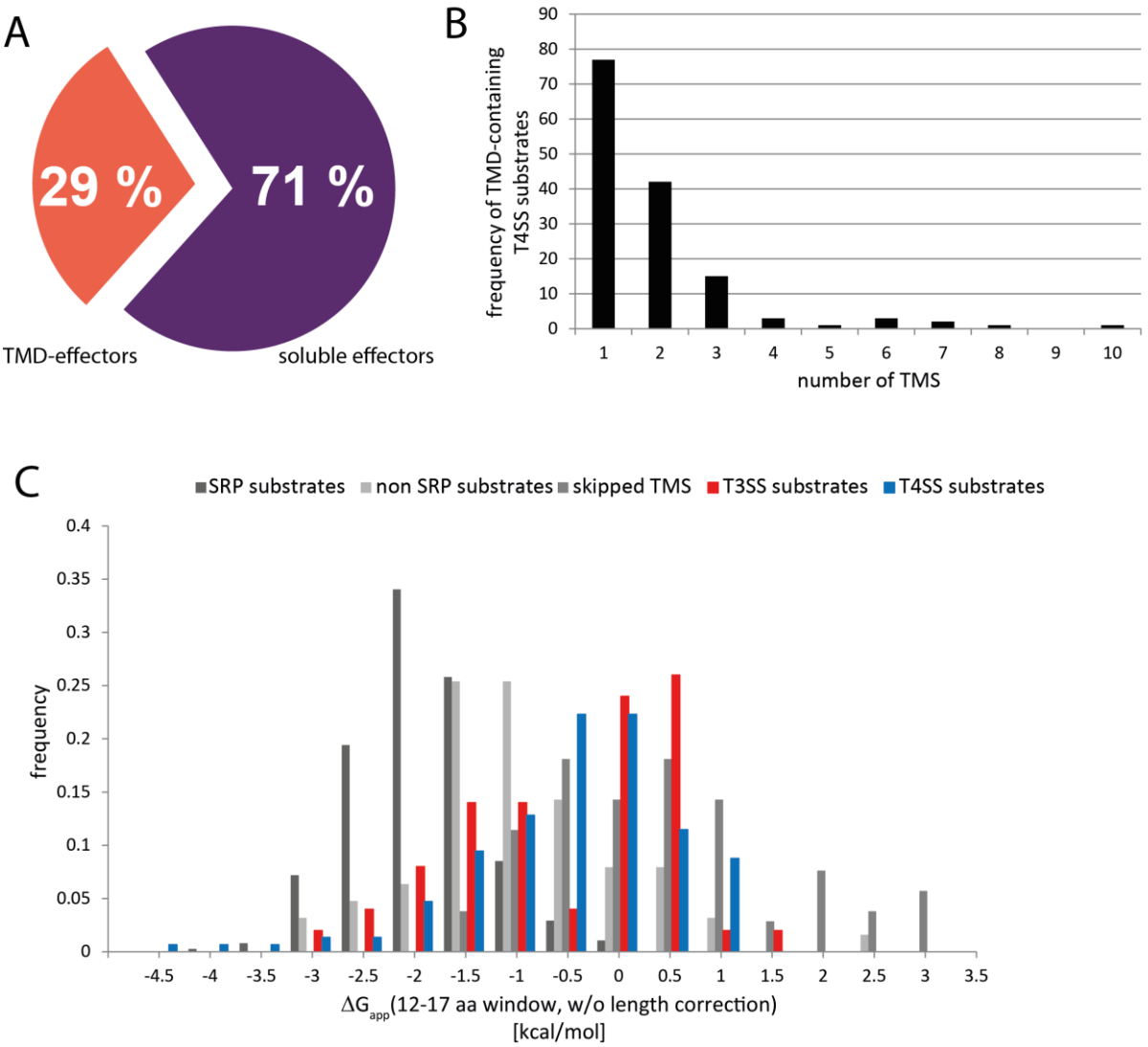


Fig. 14: Dot/Icm-secreted transmembrane proteins (TMD-effectors).
All effectors of the Dot/Icm system in *Legionella pneumophila* strain Philadelphia-1 which are listed in the Secret4 database (Bi *et al.*, 2013) were analyzed using the full protein scan of the ΔG predictor (Hessa et al, 2007).

A) The percentage of effectors of the Dot/Icm system with no (soluble effectors) or one and more (TMD-effectors) predicted transmembrane segments are depicted (ΔG predictor settings: window: 18-35, length correction: ON).
B) The histogram shows the distribution of the number of TMS of predicted TMD-effectors (ΔG predictor settings: window: 18-35, length correction: ON).
C) Calculation of ΔG_{app} for the SRP-targeting window of 12-17 amino acids for hydrophobic segments of type III-secreted transmembrane proteins (red) and TMD-effectors of the Dot/Icm system (blue) compared to *E. coli* transmembrane proteins (shades of grey). The classification of *E. coli* membrane proteins is according to Schibich *et al.* (Schibich *et al.*, 2016): SRP substrates (dark gray), non-SRP substrates (middle grey) and substrates, in which the first TMS was skipped by SRP (light grey). (ΔG predictor settings: window: 12-17, length correction: OFF). TMS transmembrane segment

A subgroup of about 30 % is predicted as membrane proteins with mostly one but in some cases up to ten TMS (Fig. 14A,B) (Krampen *et al.*, 2018). They show a highly diverse hydrophobicity with a minimal ΔG_{app} (window 18-35 amino acids) value of -4.5 kcal/mol. The selected TMD-effectors have ΔG_{app} values as follows: Except for SidF, the first TMS is always more hydrophobic with ΔG_{app} values of -1.1, -1.2 and -4.6 kcal/mol for LegC2, Ceg4 and LegC3, respectively, when using a window of 18-35 amino acids. The second TMS of SidF has a ΔG_{app} value of -2.2 kcal/mol (Fig. 9). As a $\Delta G_{app,}$ 18-35 with a value of 0 means an inner membrane insertion of 50 % of the molecules (Hessa, Nadja M. Meindl-Beinker, *et al.*, 2007), all four TMD-effectors are likely to be found in the bacterial inner membrane.

As already described in Krampen *et al.*, ΔG values can also be used as a proxy for SRP targeting, however a sliding window of 12-17 amino acids was chosen because SRP binds polypeptides of limited length of 12-17 residues (Schibich *et al.*, 2016). As reference, typical SRP-substrates, non-SRP-substrates and substrates, in which the first TMS was skipped, all described in Schibich *et al.*, were included in the analysis. The ΔG_{app} (window 12-17 amino acids) values of the most negative TMS are shown in Fig. 14C. The most hydrophobic TMS of Ceg4, LegC2, SidF and LegC3 in the SRP-targeting window of 12-17 amino acids have ΔG_{app} values of -0.08, -0.6, -1.2 and -1.4 kcal/mol, respectively. A list of all predicted TMD-effectors in *L. pneumophila* with their number of predicted TMS and the ΔG value of the most hydrophobic stretch in a sliding window of 18-35 as well as 12-17 amino acids is provided in appendix 7.1 and 7.2.

Similar to effectors of the T3SS, most TMD-effectors of the Dot/Icm system in L. pneumophila show a more intermediate hydrophobicity with a peak at -0.5 kcal/mol in the SRP-targeting window. However, in contrast to T3 substrates, there is also a significant number of substrates of the Dot/Icm system with a higher hydrophobicity down to values of -4 kcal/mol. The selected TMD-effectors show different hydrophobicities for their TMS. LegC2 and Ceg4 are predicted to be semi-hydrophobic, and with this might avoid SRP targeting. In contrast, SidF and especially LegC3 show high hydrophobicity values. With this, recognition and targeting to the bacterial inner membrane by SRP is very likely. Overall, the

different hydrophobic TMS of the four selected TMD-effectors makes them an interesting set of effectors to elucidate different possible targeting pathways to the Dot/Icm system and how their hydrophobicities play a role.

4.4.2 TMD-effectors of the Dot/Icm system can be found in the bacterial inner membrane

In order to confirm the predicted integration into the bacterial inner membrane shown above, the actual subcellular location of TMD-effectors in *L. pneumophila* was investigated.

In a first step, I prepared crude membranes of *L. pneumophila* lysates. Proteins in the pellet and supernatant, which should contain cytoplasmic proteins, were analyzed by mass spectrometry. The separation of the membranes from the cytosol worked very nicely, as inner membrane proteins as well as soluble proteins could be found specifically in their expected subcellular localization (data not shown). Unfortunately, only 17 effectors of the Dot/Icm system could be identified by mass spectrometry, none of them a TMD-effector.

To increase expression levels, I overexpressed LegC2, SidF and RalF as well as the inner membrane protein Lep-inv, each attached to a N-terminal HA-epitope tag to facilitate detection, from a low-copy number plasmid and repeated crude membrane preparations. As expected, Lep-inv could be solely found in the membrane pellet, whereas the soluble effector RalF was only present in the supernatant (Fig. 15). The TMD-effector SidF can be solely found in the crude membrane samples whereas LegC2 was detected in the membrane pellet as well as in the supernatant. It seems, however, to be more stable when it is integrated in the inner membrane than in the cytoplasm.

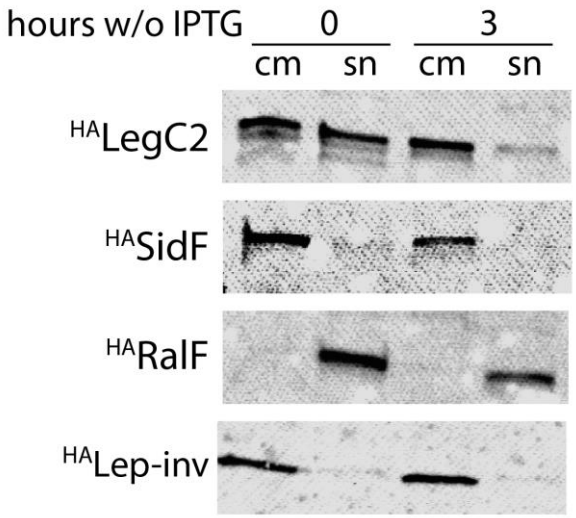


Fig. 15: Subcellular localization of the TMD-effectors LegC2 and SidF in comparison to the soluble effector SidM and the inner membrane protein Lep-inv.
Protein stability of the effectors LegC2, SidF, RalF or the inner membrane protein Lep-inv, respectively in cell membranes and the cytoplasm was analyzed by crude membrane preparation after 20 hours protein expression with 0.5 mM IPTG or additional 3 hours growth without IPTG. Crude membrane samples as well as the respective supernatants were separated by SDS-PAGE and visualized by immunodetection. A representative result of at least three independent experiments is shown.
Cm crude membranes, Sn supernatant

To analyze the subcellular localization of TMD-effectors in more detail, I adapted protocols for membrane fractionation in *L. pneumophila*. Similar to crude membrane preparations, I expressed the selected TMD-effectors LegC2, LegC3 as well as the soluble control SidM from low-copy number plasmids. The inner and outer membranes were separated using sucrose density gradient equilibrium centrifugation. The contents of 12 fractions were analyzed by SDS-PAGE followed by Western blotting. To exclude overexpression effects, I also analyzed the endogenous expression and localization of TMD-effectors using mass spectrometry (experimental procedure is depicted in Fig. 16).

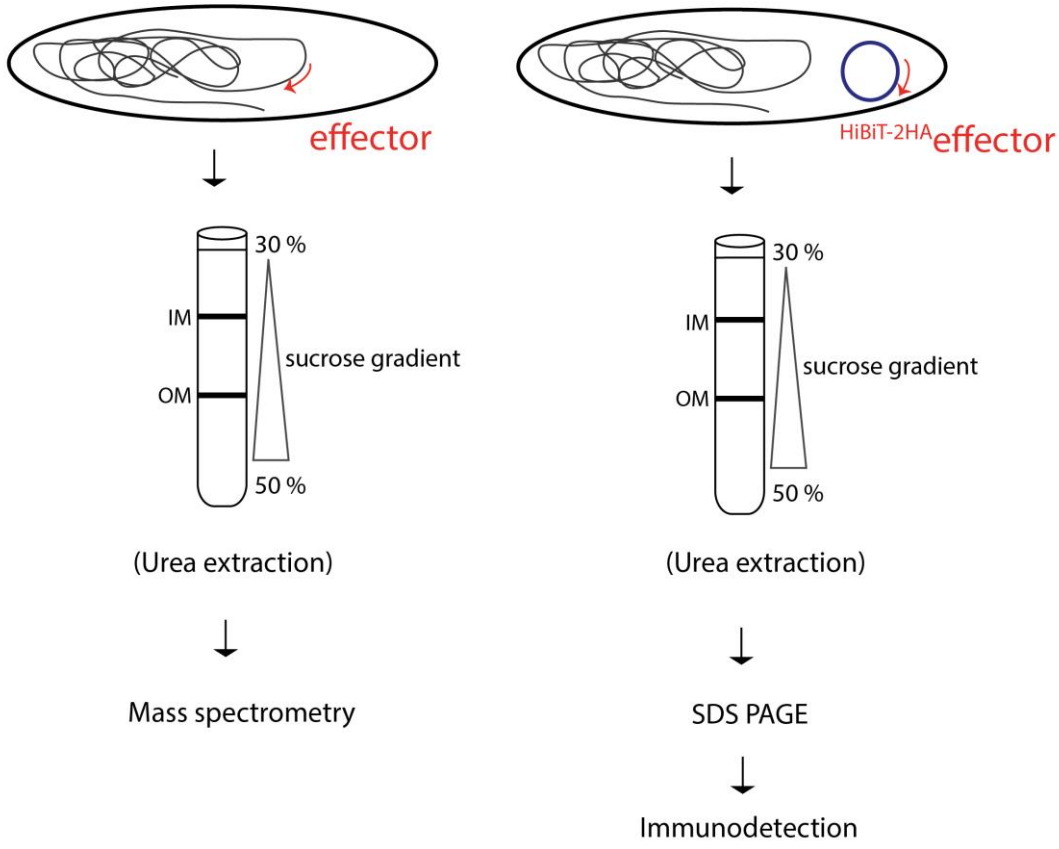
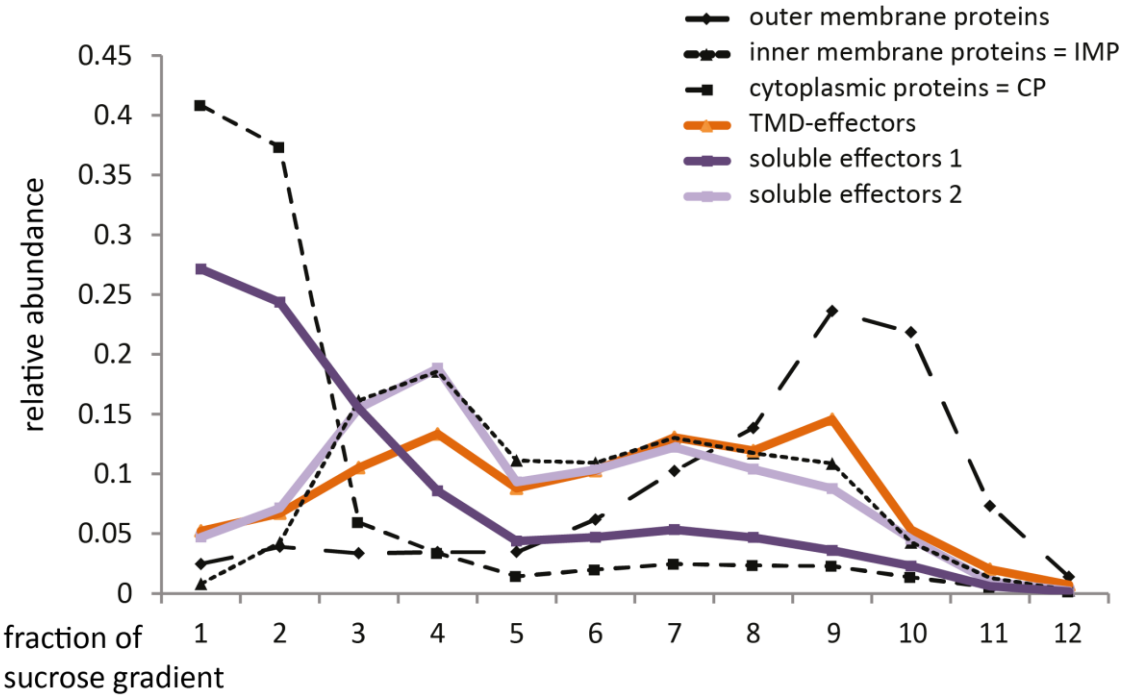


Fig. 16: Overview of membrane fractionation.
L. pneumophila wild type expressing effector protein endogenously from the chromosome (left panel) or from a low-copy number plasmid (right panel) were subjected to membrane fractionation. Crude membranes were separated using a 30-50 % sucrose gradient. Sucrose fractions were washed with 8 M urea or left untreated. Proteins were detected by mass spectrometry or Western blotting.

Based on the results obtained by mass spectrometry analysis, mapping of proteins of known subcellular localization to the different fractions showed that cytosolic proteins peak in fractions 1 and 2, inner membrane proteins in fractions 3 and 4 with a long tail until fraction 9 and outer membrane proteins in fractions 10 and 11 (Fig. 17A). Interestingly, components of the Dot/Icm core complex that span both inner and outer membrane peaked distinctly in fractions 9 and 10 (Fig. 17B upper panel). Members of the T4 coupling complex, independent whether they are membrane proteins (DotL and DotM) or soluble proteins (IcmS, IcmW, LvgA and DotN), run like inner membrane proteins with a peak in fractions 3 and 4 as well as 9 and 10 (Fig. 17B middle panel). The ATPase DotO, which was shown to be tightly associated to the T4 channel (Chetrit *et al.*, 2018), runs like an inner membrane protein, whereas DotB resembles a soluble protein (Fig. 17B lower panel). This fits fine with the observation of Chetrit *et al.* concluding that DotB “travels” between cytoplasm and Dot/Icm system.

By matching the proteins identified by mass spectrometry to the Secret4 database, I was also able to identify 48 reported substrates of the Dot/Icm system. Predicted TMD-effectors peaked in fraction 4 with a long tail until fraction 9 (Fig. 17A), resembling inner membrane proteins with a possible co-fractionation tendency with the Dot/Icm coupling complex. Unexpectedly, soluble effectors could be divided into two subgroups (Fig. 17A). Many of them were found in fractions 1 and 2. The second group, however, showed peaks in fractions 3 and 4 and had a similar tail towards higher fraction numbers as inner membrane proteins, suggesting possible interactions with inner membrane complexes.

A.


outer membrane proteins
inner membrane proteins = IMP
cytoplasmic proteins = CP
TMD-effectors
soluble effectors 1
soluble effectors 2
relative abundance
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
fraction of sucrose gradient
1
2
3
4
5
6
7
8
9
10
11
12

B.

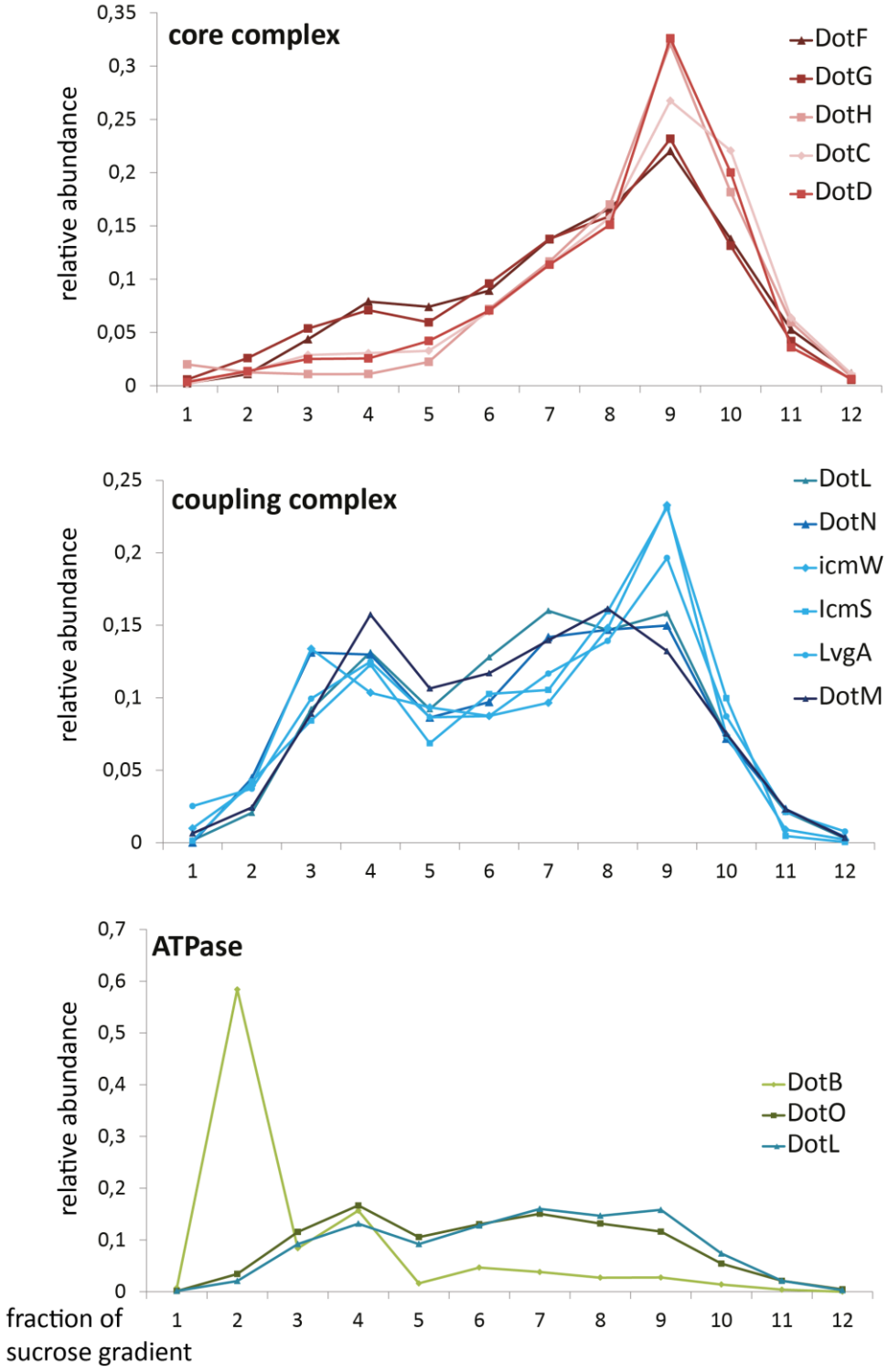


Fig. 17: Analysis of subcellular localization of TMD-effectors and components of the Dot/Icm system.
The subcellular localization of TMD-effectors as well as components of the Dot/Icm system was analyzed using a sucrose density gradient centrifugation protocol and mass spectrometry. The data points represent the mean of at least three independent experiments.
A) The relative abundance of soluble as well as TMD-effectors in 12 fractions of a sucrose gradient are shown as

identified and quantified by mass spectrometry. For comparison, well-known cytoplasmic, inner membrane and outer membrane are plotted.
B) As in A: Relative abundance of indicated protein of the core complex (upper panel), the coupling complex (middle panel) and ATPases (lower panel) of the Dot/Icm system.

The same analysis was conducted with a T4SS deletion strain. However, no change in running behavior could be observed (SidM, LegC2 and LegC3 shown as examples in Fig. 18A). This indicates that the presence of the Dot/Icm system did not influence the subcellular localization of effector proteins. As already observed in crude membranes, when LegC2 and LegC3 are overexpressed from a low-copy number plasmid a small amount of protein can be detected in the "soluble fraction" 1 (Fig. 18B), whereas neither protein could be detected in the same fraction when expressed endogenously from the chromosome (Fig. 18A). These results suggest that endogenously expressed TMD-effector are localized within the bacterial inner membrane. Effector overexpression, however, may overwhelm the SRP targeting system and thus increasing the pool of inner membrane inserted TMD-effectors but at the same time resulting in the presence of TMD-effectors in the cytoplasm.

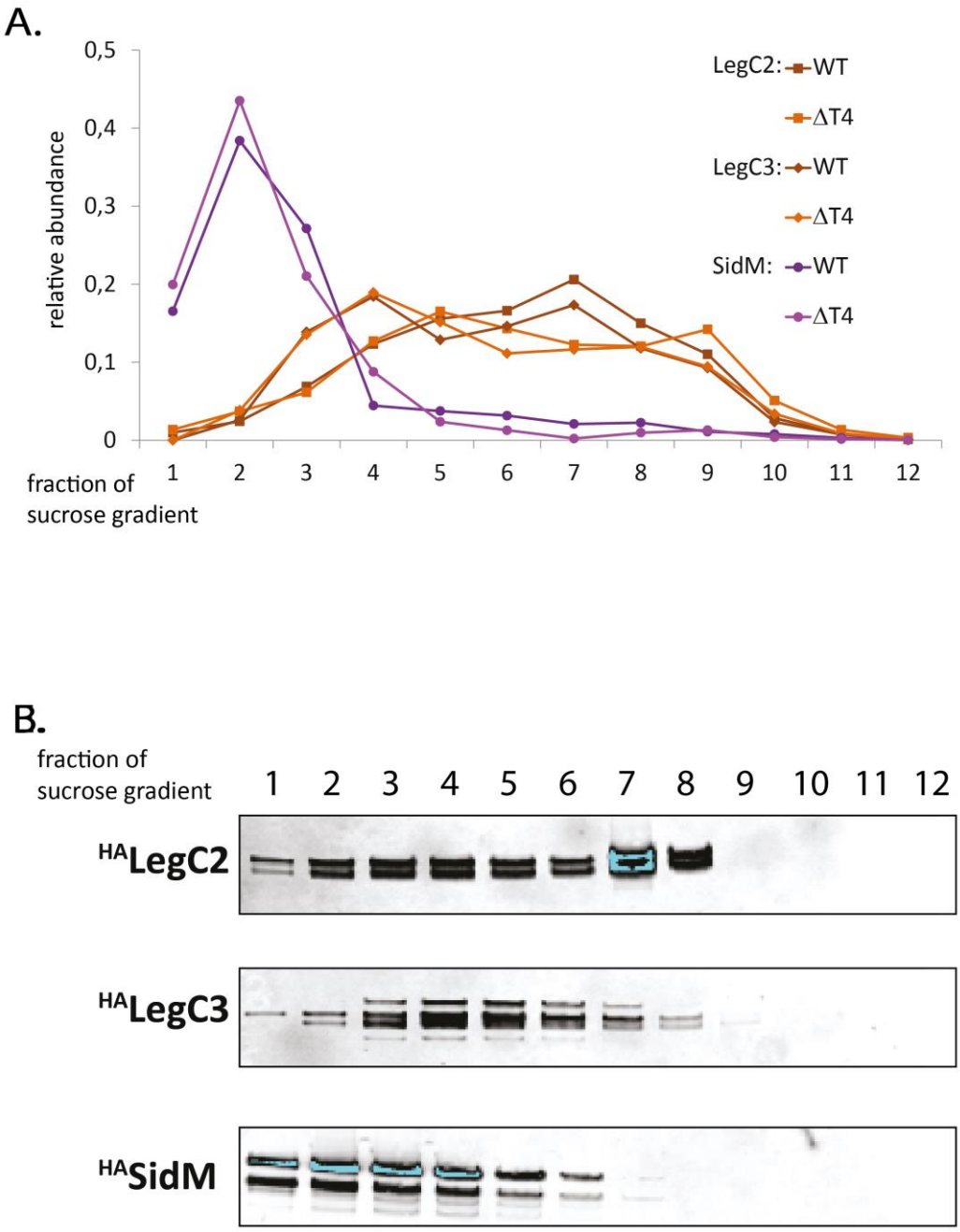


Fig. 18: Analysis of subcellular localization of the TMD-effectors LegC2, LegC3 and the soluble effector SidM.
The relative abundance of selected effectors in 12 fractions of a sucrose gradient was analyzed using a sucrose gradient centrifugation protocol and mass spectrometry (A) or Western blot (B).
A representative result of at least three independent experiments is shown.

To distinguish peripherally membrane-associated proteins from proteins properly integrated into the bacterial inner membrane, I performed urea extraction of pooled membrane fractions (fractions 2-6) followed by SDS-PAGE and Western blotting or mass spectrometry. While soluble effectors of the Dot/Icm system were almost completely extracted from the membranes, the hydrophobic TMD-effectors showed to be true integral membrane proteins. This observation was made for the plasmid-

expressed effectors SidM, LegC2 and LegC3 as well as for endogenous proteins expressed from the chromosome (Fig. 19).

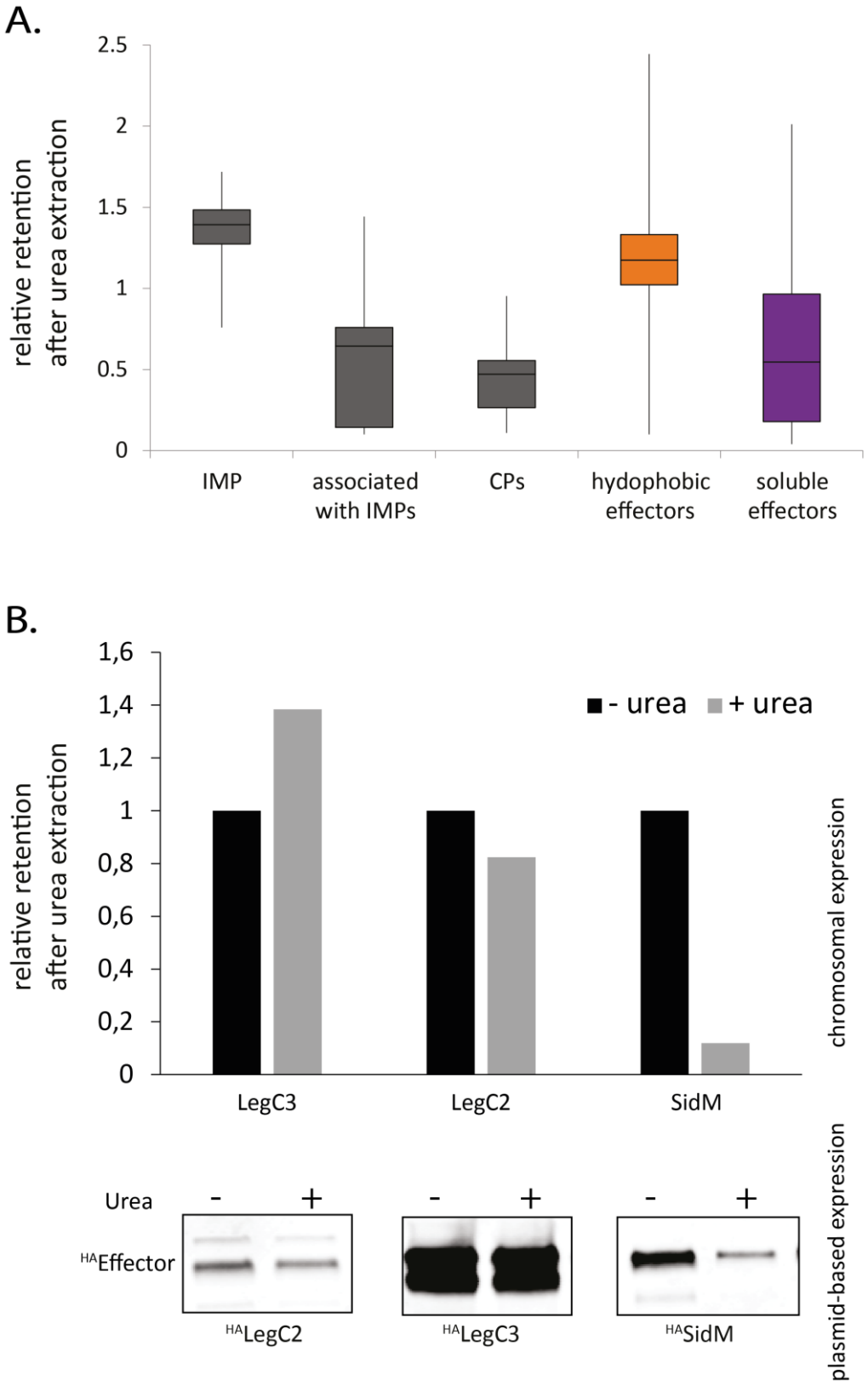


Fig. 19: Analysis of the relative retention of proteins in the bacterial inner membrane after urea extraction. Membrane integration of TMD-effectors (endogenous **(A, B upper panel)** or expressed from a low copy number plasmid **(B lower panel)** was analyzed using a sucrose gradient centrifugation protocol followed by mass spectrometry. Fractions 2-6 were pooled and integral membrane status was assessed by washing the samples with 8M urea. The relative retention in the bacterial inner membrane of the indicated protein groups is shown. A representative result of at least three independent experiments is shown.

In summary, membrane fractionation of *L. pneumophila* showed that TMD-effectors can clearly be found properly integrated into the bacterial inner membrane. This suggests that either TMD-effectors follow a two-step secretion pathway with an inner membrane intermediate or have a strong potential of dead end mistargeting into the bacterial inner membrane.

4.4.3 Can chaperone binding prevent membrane targeting?

As already shown in chapter 4.3.3, selected TMD-effectors depend on the chaperones IcmSW for their efficient translocation into host cells. Moreover, it has been suggested for TMD-effectors of the T3SS that chaperone binding to the effector protein as soon as it emerges from the ribosome exit tunnel prevents binding of other targeting factors such as SRP (Krampen *et al.*, 2018). To assess if interactions with chaperones of the Dot/Icm system may possibly shield hydrophobic stretches from being recognized by SRP and by this prevent inner membrane targeting, I repeated membrane fractionation with subsequent urea extraction and mass spectrometry analysis of a strain overexpressing IcmSW from a low-copy number plasmid (Fig. 20A). Interestingly, similar to the wild type strain TMD-effectors were still found stably integrated into the bacterial inner membrane (Fig. 20B,C). This shows that increasing the amount of the chaperones IcmSW cannot prevent inner membrane insertion of TMD-effectors.

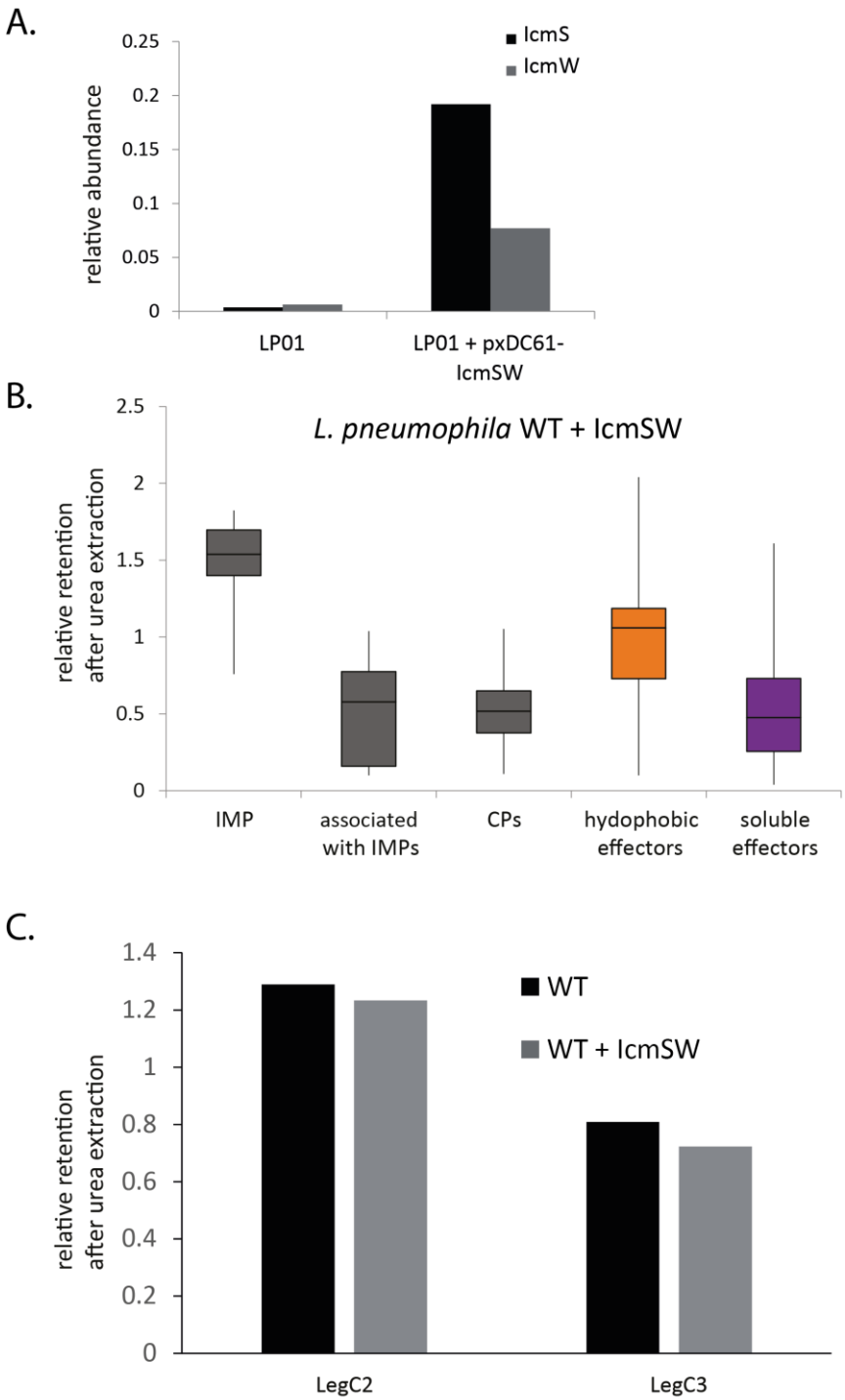


Fig. 20: Analysis of subcellular localization of TMD-effectors in a *Legionella pneumophila* wild type overexpressing *icmS* and *icmW* after urea extraction.

The subcellular localization of TMD-effectors was analyzed in the *Legionella pneumophila* wild type strain as well as the same strain overexpressing the Dot/Icm chaperones *icmS* and *icmW* from the low-copy number plasmid with 0.5 mM IPTG using a sucrose gradient centrifugation protocol. Fractions 2-6 were pooled and washed with 8 M urea to assess integral membrane status.

A) Relative abundance of IcmS and IcmW in the pooled fractions 2-6 of a sucrose gradient as identified and quantified by mass spectrometry.

B) The relative retention in the bacterial inner membrane of the indicated protein groups is shown.
C) The relative retention of the TMD-effectors LegC2 and LegC3 in the bacterial inner membrane is shown.

4.4.4 Investigation of possible IcmSW binding sites in SidF

To further address the question if there is an IcmSW binding site upstream of the TMS of TMD-effectors which may protect SRP targeting, I chose SidF to investigate possible protein-protein interactions using *in vivo* photo-crosslinking. To this end, several amber-stop codons were introduced at the site of the N-terminal domain of SidF facing the subsequent TMS using Quikchange site-directed mutagenesis (Fig. 21A).

As no one has used *in vivo* photo-crosslinking in *L. pneumophila* before, I tested published interactions between IcmS and IcmW (Xu *et al.*, 2017) to establish the method. For this, constructs with amber-mutations in different positions in either IcmS or IcmW were expressed from a low-copy number plasmid in a *L. pneumophila* wild type strain. Additionally, the strain was transformed with the constitutively active pEvol-*p*Bpa plasmid, which provides amber suppressor tRNA and aminoacyl-tRNA synthetase for *p*Bpa incorporation. In contrast to the same method in *Salmonella* or *E. coli*, *L. pneumophila* growth in liquid culture was very slow as soon as *p*Bpa was added (data not shown). As *L. pneumophila* is mostly cultured on agar plates, I tried to add *p*Bpa as well as IPTG for induction of protein expression in the agar medium. With this, I was able to grow *L. pneumophila* overnight in the presence of *p*Bpa resulting in incorporation of the artificial amino acid. After harvesting 1-2 ODU of bacteria from the agar plates, *in vivo* photo-crosslinking was performed as described in Materials and Methods (see chapter 3.6).

The comparison with a negative control without *p*Bpa showed that the addition of *p*Bpa resulted in an incomplete suppression of amber stop codons. Nevertheless, several crosslinks were detected including a crosslink with a size of around 37 kDa which matches an IcmS-IcmW dimer (Fig. 21A). Next, strains with several individual amber positions within the N-terminal domain of SidF were subjected to UV irradiation. Strong crosslinks were detected for positions Q721 and K742, which are 16 and 37 amino acids upstream from the start of the first TMS of SidF, respectively (Fig. 21C). Interestingly, the detected crosslinks show different sizes. As the crosslink at position K742 with approximately 170 kDa appears to be too big to resemble an interaction between SidF and IcmSW, the position Q721 was investigated further in single deletion mutants of *icmS* or *icmW* to analyze if the detected crosslink might be an interaction between IcmSW and SidF. Neither the deletion of *icmS* nor of *icmW* resulted in the disappearance of the detected crosslink, indicating that at this position SidF probably interacts with a yet unknown protein (Fig. 21D).

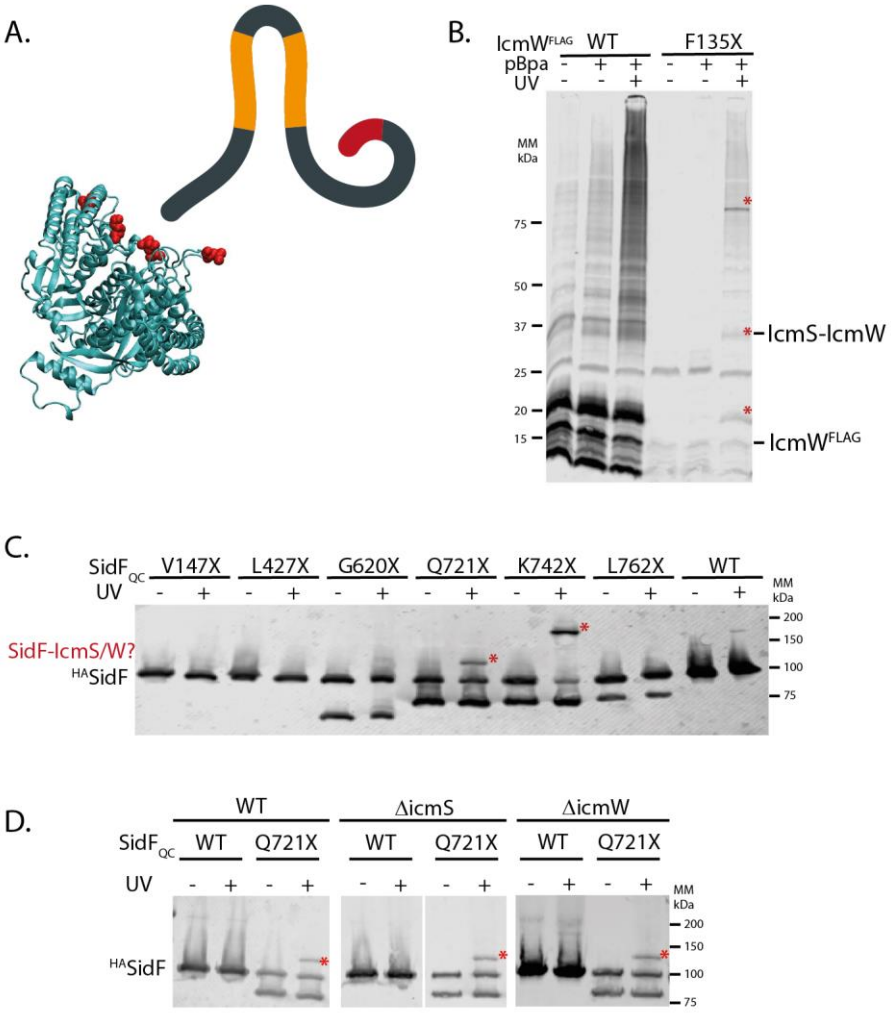


Fig. 21: Protein-protein interaction analysis of SidF by in *vivo* photo-crosslinking.
Legionella pneumophila wild type strain expressing either IcmW^{FLAG} (for proof of concept) or ^{HA}SidF as well as harboring the pEvol-pBpa plasmid were grown over night on agar plates supplemented with 0.5 mM IPTG and 1 mM *p*Bpa. Bacteria were UV$_{365nm}$ irradiated for 30 minutes to induce protein-protein crosslinks. *p*Bpa mutations are denotated as "X". Detected cross-links are labeled with an asterisk.
A representative result of at least three independent experiments is shown.
A) The Structure of the N-terminal domain of SidF as solved in (Hsu *et al.*, 2012) attached to the predicted C-terminus of the protein is shown. TMS are shown in orange. The C-terminal translocation signal is shown in red. Positions 147, 427, 620, 721 and 742 within the N-terminal domain are shown in red.
B) Immunodetection of IcmW^{FLAG} in whole cell with and without UV irradiation. All strains were grown in the presence or absence of the artificial amino acid *p*Bpa.
C, D) Immunodetection of ^{HA}SidF in whole cell lysate with and without UV irradiation.
*p*Bpa *para*-Benzoyl-phenylalanine.

As the translocation of SidF is decreased in *Legionella* strains lacking *IcmSW* (Fig. 13), it is most likely that there is a chaperone binding site within the protein.

To summarize, TMD-effectors are not only predicted to harbor TMS with a high hydrophobicity and thereby are prone to be recognized by SRP, but also can be found in the bacterial inner membrane of *L. pneumophila*. Moreover, although selected TMD-effectors have been shown to be IcmSW dependent for their successful translocation into host cells, no IcmSW-binding site upstream of the TMS of the TMD-effector SidF has been identified so far. Moreover, increased amounts of IcmSW relative to TMD-effectors did not prevent inner membrane integration of the latter. These data strengthen the hypothesis that TMD-effectors are indeed recognized by SRP and targeted to the inner membrane rather than being co-translationally bound and shielded by chaperones of the Dot/Icm system.

To exclude that the presence of TMD-effectors within the bacterial inner membrane of *L. pneumophila* is due to mistargeting, several attempts were made to develop an assay which shows translocation out of the inner membrane. As five effectors listed in the Secret4 database were predicted as TMD-effectors with a cleavable signal sequence at their N-terminus (data not shown), I tried to use one of them to show that the cleavable signal sequence is cleaved off prior to translocation into host cells. Unfortunately, either stable expression of the TMD-effector could not be achieved, or translocation of the effector could not be detected by immunodetection. Another approach was to incorporate cysteine residues within a periplasmic loop of a TMD-effector and label it with fluorescein-maleimide prior to infecting Raw 264.7 macrophages with that strain. Using NanoBret™, a fluorescence signal should only be detectable if the labeling in the periplasm was successful which in turn can only occur after proper membrane integration of the effector. Although the energy transfer from a NanoLuc fusion protein to the fluorescein molecule worked nicely in solution (data not shown), no fluorescence signal could be observed after labeling whole cells. This indicates that the labeling of cysteine residues in the periplasm should be further optimized.

Although I was not able to show effector translocation from the bacterial inner membrane, it is rather unlikely that inner membrane localization of investigated TMD-effectors was solely due to mistargeting (see discussion). Endogenous expression of TMD-effectors did not lead to any detectable presence of the protein in the cytoplasm (fraction 1 of the sucrose gradient), which would have been expected if TMD-effectors could avoid SRP-targeting. Moreover, the same effector localization was observed in a strain overexpressing the effectors IcmSW, which would be most likely involved in an active avoidance mechanism.

In summary, these data indicate that inner membrane integration of TMD-effector is indeed a part of the translocation process and not protein mistargeting.

4.5 Recognition of TMD-effectors by the Dot/Icm system

To gather additional evidence for the hypothesis that TMD-effectors follow a two-step secretion mechanism with an inner membrane intermediate, I next investigated how TMD-effectors are recognized by the Dot/Icm system once integrated into the bacterial inner membrane.

4.5.1 Are TMD-effectors recognized in the cytoplasm or periplasm?

After being integrated into the bacterial inner membrane, TMD-effectors can in principle be recognized by the Dot/Icm system at the cytoplasmic or the periplasmic side of the membrane.
A bioinformatics survey of all predicted TMD-effectors in *L. pneumophila* revealed that the majority of substrates have only very short periplasmic stretches with a length of 10-20 amino acids (Fig. 22). With this, any interaction of such a periplasmic stretch with components of the Dot/Icm system located in the periplasm is rather unlikely. Moreover, the prediction of the transmembrane topology of the more hydrophobic TMD-substrates by TOPCONS (Tsirigos *et al.*, 2015) indicated that the vast majority of substrates possess a C_{in} topology (data not shown). I hypothesized that a cytoplasmic presentation of the C-terminal secretion signal is obligatory for recruitment of most TMD-substrates to the Dot/Icm machinery and their subsequent translocation into host cells.

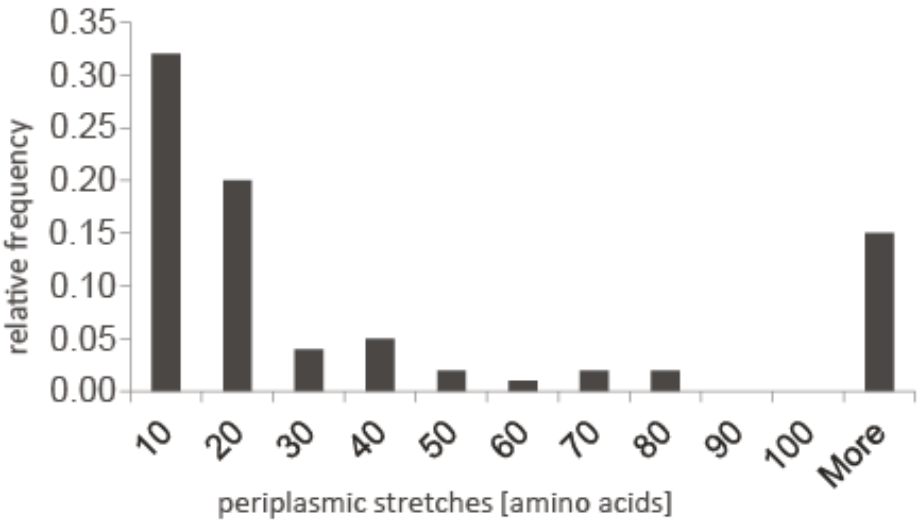


Fig. 22: TMD-effectors have only short stretches located in the bacterial periplasm.
Periplasmic stretches in TMD-effectors were predicted by TOPCONS (Tsirigos *et al.*, 2015). Their size and relative frequency is shown in a histogram.

To confirm that the recognition of TMD-effectors is at the cytoplasmic side of the inner membrane I analyzed a possible recognition of TMD-effectors in the periplasm by following the translocation of TMD-effectors upon deletion of the Dot/Icm components DotF and DotK (Fig. 23A). DotF and DotK have sizable protein parts located in the periplasm and by this might act as putative substrate-accepting proteins in the periplasm (Sutherland *et al.*, 2013). Additionally, I determined the inner membrane topology of selected TMD-effectors *in vivo*.

Using the split NanoLuc-based translocation assay *dotK* and *dotF* deletion strains were investigated with respect to their ability to translocate effector proteins into Raw 264.7 macrophages.
For all tested effectors, the translocation efficiency in a Δ*dotF* strain was reduced to a maximum level of 50 % compared to the wild type strain (Fig. 23B). As this seems to be true for soluble as well as TMD-effectors, an unstable Dot/Icm system is the most likely explanation. This also fits to the observation of a reduced capacity for intracellular growth by Sutherland *et al.* (Sutherland *et al.*, 2013). Interestingly, the translocation of SidF and RalF was nearly completely abolished. Using a two-hybrid system it has already been shown that both proteins interact with a periplasmic region of DotF (Sutherland *et al.*, 2013). Although in principle SidF should have the ability to interact with DotF in the periplasm, it is hard to explain how such an interaction can take place for the soluble effector RalF before it is inserted into the central channel of the Dot/Icm system.
Deleting *dotK* did not seem to influence the translocation of SidM, RalF, LegC2 and LegC3 and resulted in only a slight reduction of translocation efficiency for SidF (Fig. 23B), suggesting no central role of DotK in substrate recognition or translocation. The translocation of Ceg4 was also reduced. However, this might be due to the reduced protein expression in this knock-out strain (appendix. 7.3).

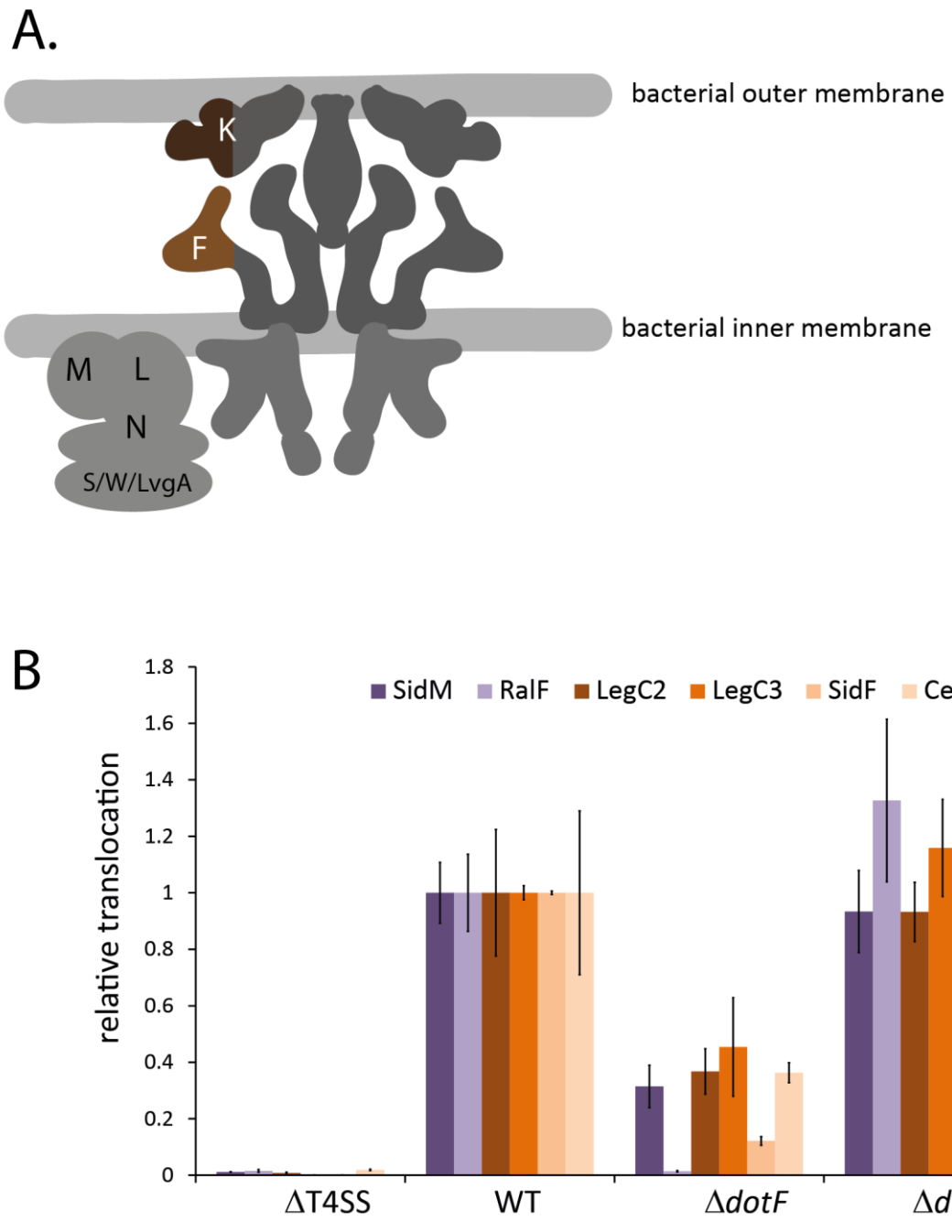


Fig. 23: Analysis of the putative substrate-accepting proteins DotF and DotK

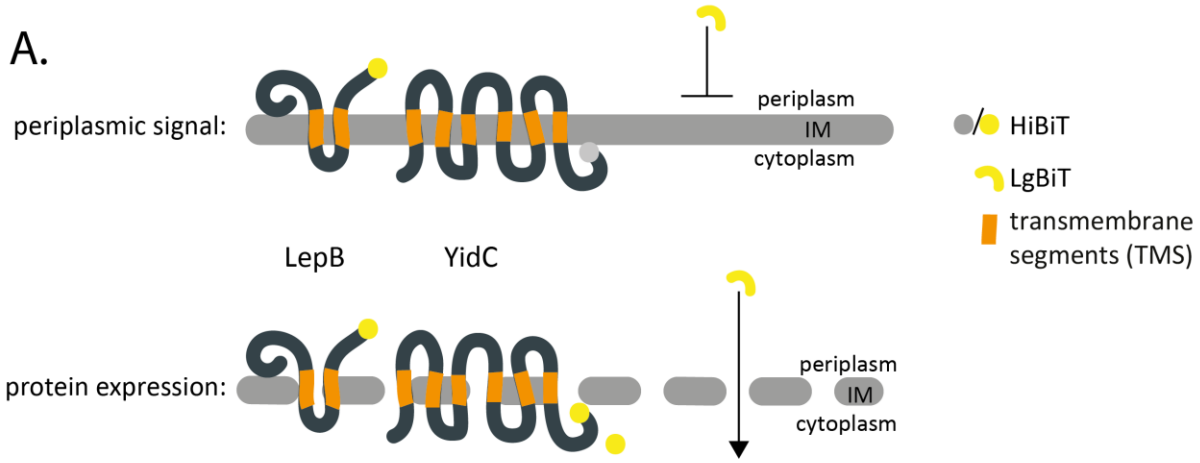
A) Localization of DotF and DotK in the Dot/Icm system.

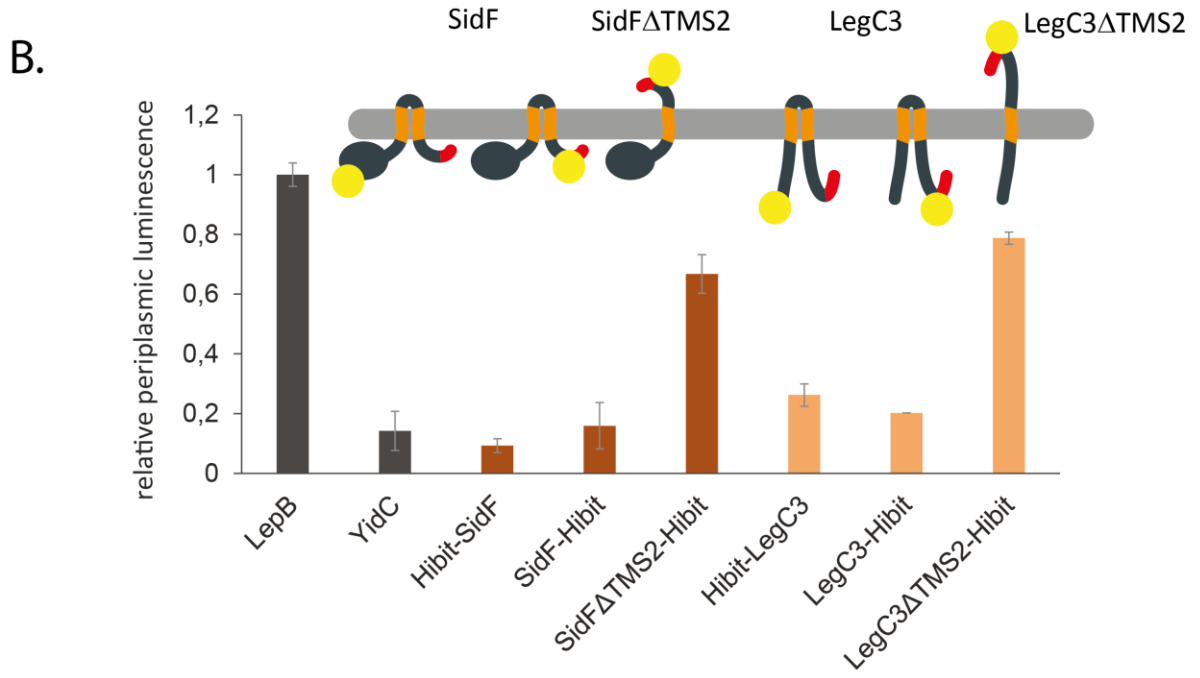
B) *L. pneumophila* wild type, a ΔT4SS mutant strain or strains lacking *dotF* or *dotK*, respectively, expressing the Dot/Icm effector proteins ^{HiBiT}SidM, ^{HiBiT}RalF, ^{HiBiT}LegC2, ^{HiBiT}LegC3, ^{HiBiT}SidF or ^{HiBiT}Ceg4 were used to infect Raw 264.7 macrophages which constitutively express LgBiT. Luminescence was measured over time using a plate reader. Only the luminescence signal 4 hpi is displayed. The data points represent the mean (± standard deviation) of at least three independent experiments.

hpi hours post infection

To validate the C_{in} topology of TMD-effectors within *L. pneumophila*, LegC3 and SidF were subjected to a split NanoLuc- based topology assay. *L. pneumophila* homologs of the well-studied integral membrane proteins LepB and YidC were used as a periplasmic and cytoplasmic control, respectively. LepB has two transmembrane helices and a N_{out}-C_{out} membrane topology (Paetzel, Dalbey and Strynadka, 1998) whereas YidC possesses six transmembrane helices with both termini facing the cytoplasmic side (Oliver and Paetzel, 2007). Both proteins were fused at their C-termini to a HiBiT peptide. The LgBiT protein was added from the outside together with an inner membrane impermeant substrate excluding any luminescence in the cytosol. The periplasmic signal was normalized to the protein expression which was measured after cell lysis (Fig. 24A). To investigate the topology of LegC3 and SidF, both their termini were individually fused to the HiBiT peptide. Only low levels of periplasmic luminescence could be measured for either of these constructs, suggesting a C_{in}-N_{in} topology of LegC3 and SidF (Fig. 24B). To confirm these data, the second TMS of LegC3 and SidF were deleted (ΔTMS2), respectively. By this, the C-terminus should now be located in the periplasm. Indeed, C-terminal fusion of the HiBiT peptide to these constructs showed an increased luminescence up to 70-80 % (Fig. 24B).

To further elucidate the importance of the cytoplasmic localization of the C-terminal secretion signal of TMD-effectors the translocation of SidFΔTMS2 was analyzed. No translocation into host cells could be measured using the Split NanoLuc translocation assay (Fig. 24D) although protein expression of SidFΔTMS2 was three times as high compared to the full-length SidF (Fig. 24C) confirming that the C-terminal T4 signal is indeed recognized in the bacterial cytoplasm.


A.
periplasmic signal:
periplasm
IM
cytoplasm
HiBiT
LgBiT
transmembrane segments (TMS)
LepB
YidC
protein expression:
periplasm
IM
cytoplasm


B.
SidF
SidFΔTMS2
LegC3
LegC3ΔTMS2
relative periplasmic luminescence
1,2
1
0,8
0,6
0,4
0,2
0
LepB
YidC
Hibit-SidF
SidF-Hibit
SidFΔTMS2-Hibit
Hibit-LegC3
LegC3-Hibit
LegC3ΔTMS2-Hibit

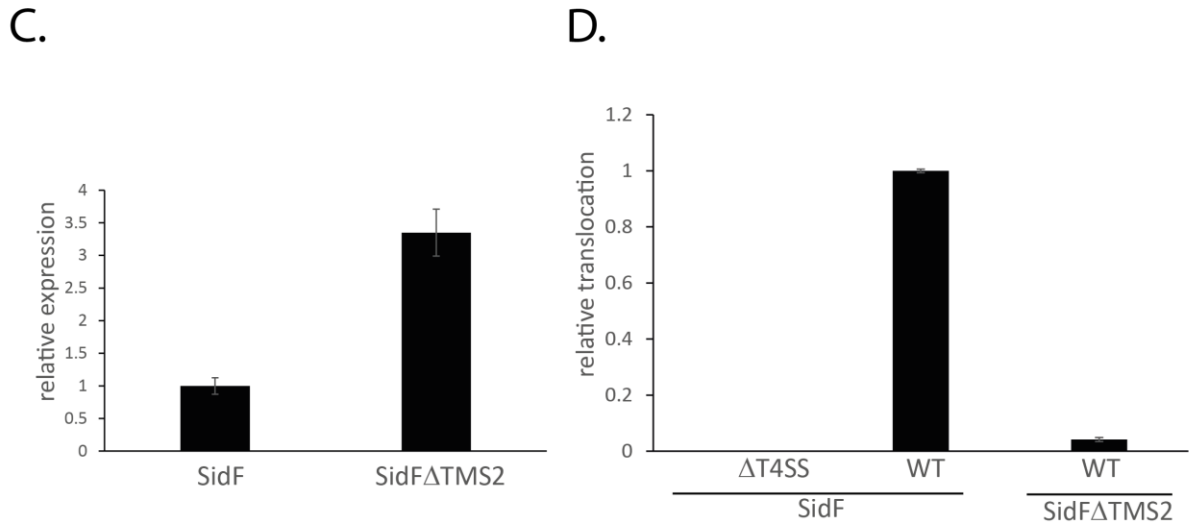

C.
relative expression
4
3.5
3
2.5
2
1.5
1
0.5
0
SidF
SidFΔTMS2
D.
relative translocation
1.2
1
0.8
0.6
0.4
0.2
0
ΔT4SS
WT
WT
SidF
SidFΔTMS2

Fig. 24: Analysis of the membrane topology of TMD-effectors.
Split NanoLuc was used for investigating the membrane topology of the TMD-effectors SidF and LegC3. The integral membrane proteins LepB and YidC were used as periplasmic and cytoplasmic control, respectively. **A)** The principle of the split NanoLuc topology assay is shown. HiBiT was fused to the C-terminus of LepB and YidC. Similar amounts of whole cell samples were resuspended in either PBS or lysis buffer. Luminescence was measured after adding LgBiT and an inner membrane impermeable substrate. Samples resuspended in PBS only give a luminescence signal when the HiBiT peptide is located in the periplasm (upper panel), whereas total protein expression could be measured after cell lysis (lower panel). The periplasmic signal was normalized to the whole cell protein expression.
B) HiBiT was fused to either the N- or C-terminus of the full-length SidF and LegC3, or the C-terminus of SidF or LegC3 lacking their second TMS (SidFΔTMS2, LegC3ΔTMS2). The relative periplasmic luminescence normalized to the whole cell expression levels is shown. The relative luminescence of LepB was set to 1 and all other luminescence values were normalized accordingly.
C) The luminescence of ^{HiBiT}SidF and ^{Hibit}SidFΔTMS2 expressed in the *Legionella pneumophila* wild type strain was measured in bacterial whole cell samples. The data points represent the mean (± standard deviation) of at least three independent experiments.
D) The luminescence of ^{HiBiT}SidF and ^{Hibit}SidFΔTMS2 translocated from either the *Legionella pneumophila* wild type strain or a ΔT4SS deletion mutant into Raw 264.7 macrophages was measured. Experiments were normalized to the wild type strain to 1. The data points represent the mean (± standard deviation) of at least three independent experiments.

Overall, these results support the hypothesis that TMD-effectors are recognized by the Dot/Icm system at the cytoplasmic side of the membrane.

4.5.2 Insertion of additional protein domains into the periplasmic loop of SidF negatively affects with its translocation

As shown above, the majority of TMD-effectors have only small periplasmic loops but sizable cytoplasmic domains. The small membrane-integral part of these proteins seems to function merely as membrane anchors, which can be easily extracted from the membrane. Since the final destination of TMD-effectors is outside of the bacterial cell, extraction towards the periplasm would seem logical. However, recognition of TMD-effectors most likely takes place in the cytoplasm. On top of that, ATPases that could mediate extraction reside on the cytoplasmic side of the bacterial inner membrane, thus extraction of TMD-effectors from the inner membrane towards the cytoplasmic side is more likely.

To assess if bigger protein domains located in the periplasmic space might interfere with protein translocation by preventing extraction from the membrane, protein domains with increasing size were inserted into the periplasmic loop of SidF (Fig. 25A) and the translocation of these protein fusion was investigated.
Using crude membrane preparation, I could show that insertion of the HiBiT peptide as well as

ubiquitinE3GE13G, a folding-impaired ubiquitin variant, into the periplasmic loop of SidF did not impede expression and the proteins were still properly integrated into the inner membrane of *L. pneumophila* (Fig. 25B). Consequently, any decrease in the translocation of SidF-fusions would not be due to poor protein expression or wrong subcellular localization. Upon fusion to wild type ubiquitin, however, no protein expression could be observed. Therefore, SidF-ubiquitin was excluded from the experiment.

Luminescence measurements showed that while the insertion of the 11 residues long HiBiT peptide just resulted in a slight decrease in SidF translocation, translocation was reduced by 90 % for SidF-ubiquitinE3GE13G (Fig. 25B).

However, it must be considered that insertion of an additional protein domain could interfere with protein translocation independently of the effector's subcellular localization. As shown by Amyot *et al.* especially tightly folding proteins like dihydrofolate reductase or ubiquitin indeed resulted in a decrease of the translocation efficiency of the soluble effector RalF (Amyot, DeJesus and Isberg, 2013). This was not the case when ubiquitinE3GE13G was used. In contrast, the ubiquitinE3GE13G-RalF fusion protein showed a slight increase in translocation compared to RalF alone (Amyot, DeJesus and Isberg, 2013).

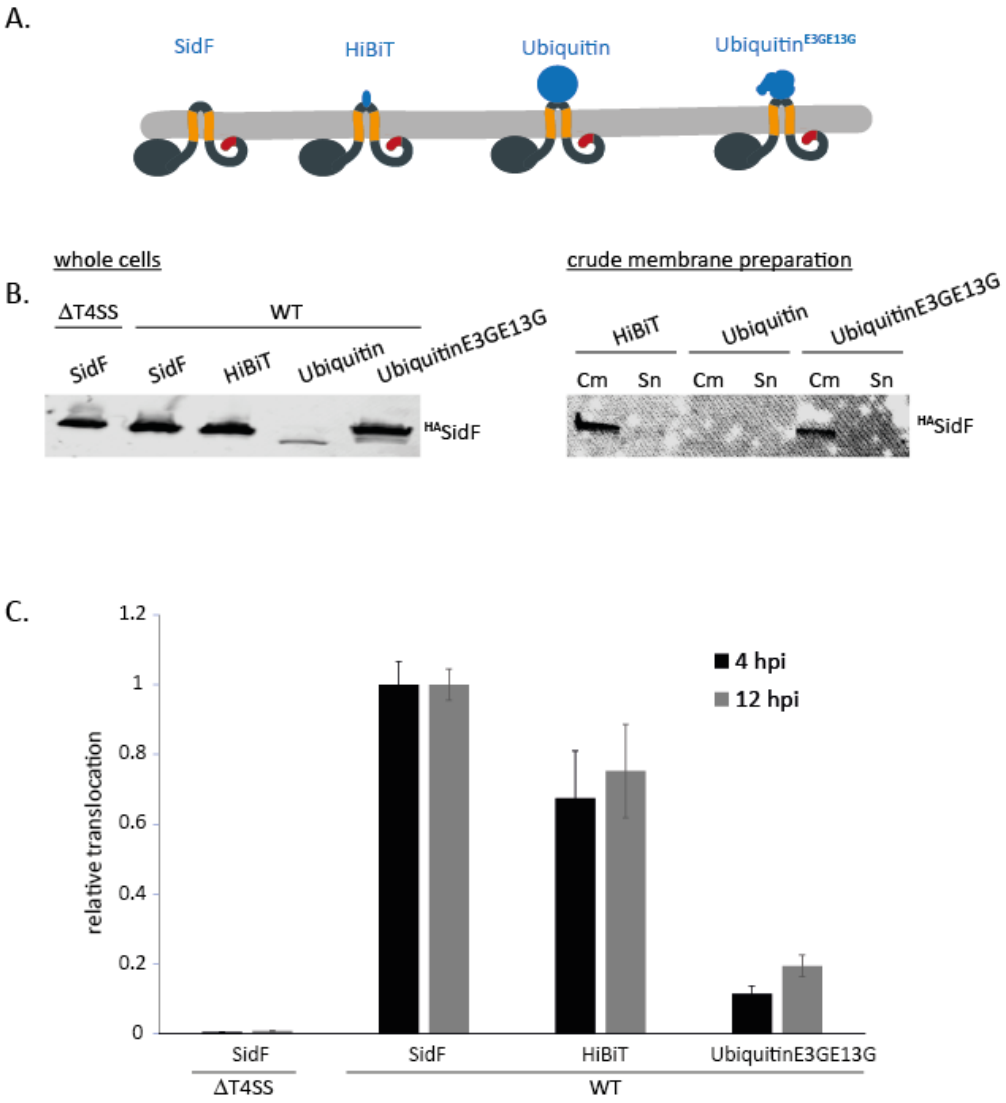


Fig. 25: Insertion of a protein domain in the periplasmic loop of SidF inhibit effector translocation.
A) Overview of SidF periplasmic fusion proteins
B) Protein expression of SidF fusions were analyzed in whole cell and crude membranes with its respective supernatant by SDS-PAGE and Western Blotting. A representative result of at least three independent experiments is shown.
C) *L. pneumophila* wild type or a ΔT4SS mutant strain expressing protein fusion of the Dot/Icm effector ^{HiBiT}SidF with HiBiT or ubiquitinE3GE13G inserted in the periplasmic loop of SidF were used to infect Raw 264.7 macrophages which constitutively express LgBiT. Luminescence was measured over time using a TECAN plate reader. The luminescence signal 4 hpi and 12 hpi is displayed. The data points represent the mean (± standard deviation) of at least three independent experiments.
Cm crude membrane; Sn supernatant; hpi hours post infection

Overall, insertion of protein domains such as a folding-impaired ubiquitin variant interferes with the translocation of TMD-effectors like SidF but not with soluble effectors like RalF. This suggests that translocation of SidF fusion proteins is abolished due to steric hindrance during the extraction of the

effector from the bacterial inner membrane and not due to general problems in translocation through the Dot/Icm system.

4.6 Unraveling the mechanism of TMD-effector recognition in the cytoplasm after inner membrane insertion

As the results above clearly favor that the recognition of TMD-effectors takes place in the cytoplasm, I next investigated which components of the Dot/Icm system as well as which substrate intrinsic properties are essential for a successful translocation.

4.6.1 TMD-effectors LegC2 and SidF can be identified in complex with the coupling complex of the Dot/Icm system

Although the exact mechanism of effector recognition overall is not yet well understood, there is strong evidence that the coupling complex of the Dot/Icm system plays a central role (Meir *et al.*, 2020). To elucidate if TMD-effectors can be found in complex with components of the coupling complex crude membrane samples were solubilized in 1 % (w/v) LMNG and subjected to Blue Native PAGE and Western blotting. Protein expression was analyzed by SDS-PAGE and Western blotting. The TMD-effectors LegC2 and SidF as well as the soluble control RalF and the inner membrane control Lep-inv were expressed from a low-copy number plasmid, each with an N-terminal HA-epitope tag to facilitate detection. DotM, DotN, DotY and DotZ were either C- or N-terminally fused to a 3xFLAG-epitope tag and expressed from the chromosome. Presumably due to insufficient expression, the immunodetection of the FLAG-tagged components of the Dot/Icm coupling complex was not successful (data not shown). Therefore, crude membrane samples separated by Blue Native PAGE were cut in 20 gel pieces and sent for mass spectrometry in order to identify in which band components of the coupling complex can be found (Fig. 26).

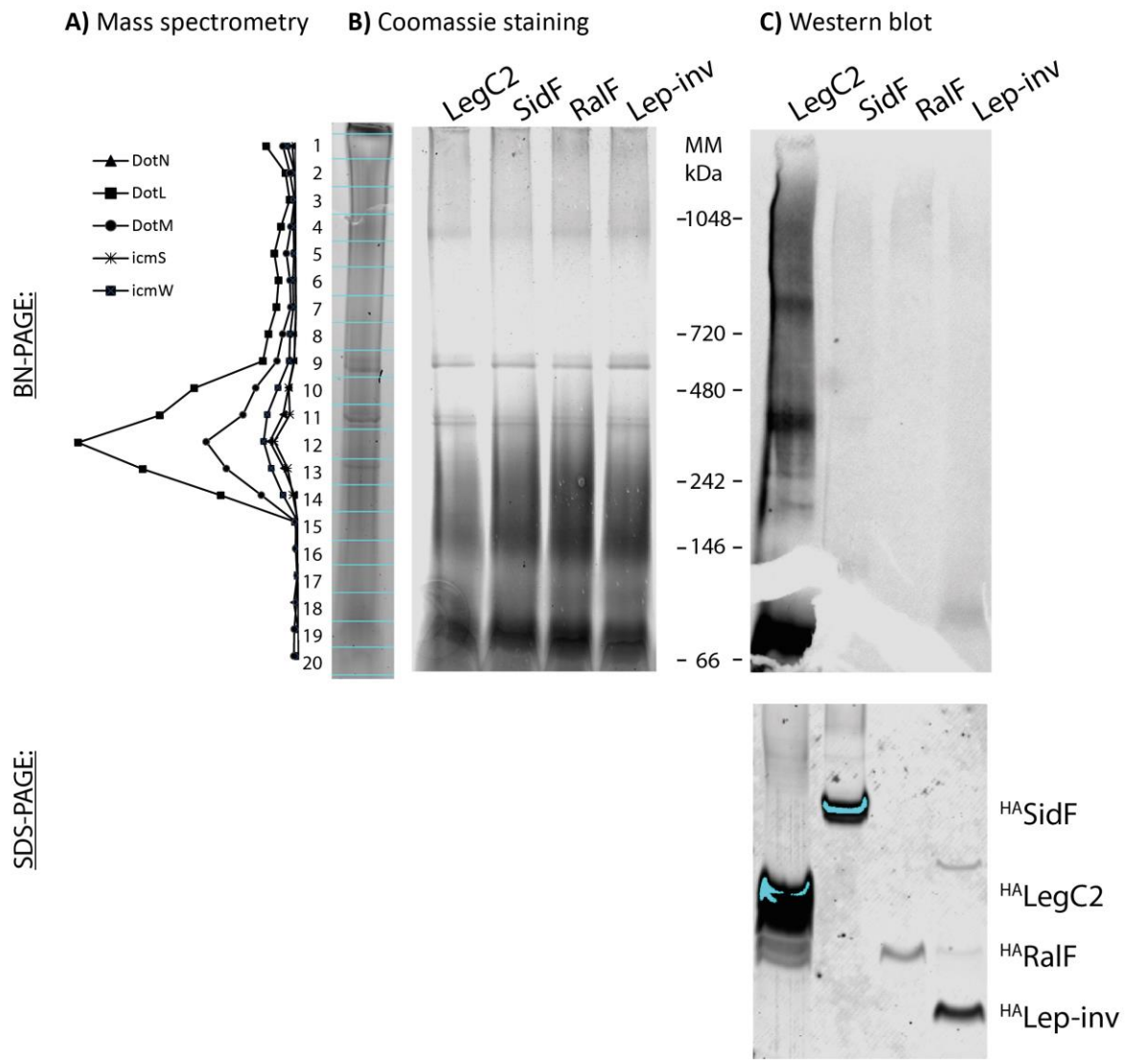


Fig. 26: TMD-effector are in complex with the Dot/Icm coupling complex.
Crude membranes of *L. pneumophila* wild type strain expressing the Dot/Icm effector proteins ^{HA}RalF, ^{HA}LegC2, ^{HA}SidF or the inner membrane protein ^{HA}Lep-inv from low copy number plasmids were separated by BN-PAGE or SDS-PAGE. Proteins were identified by mass spectrometry **(A)** visualized by Coomassie staining **(B)** or by immunodetection against the HA epitope tag after Western blotting **(C)**. A representative result of at least three independent experiments is shown.

The five coupling complex components DotL, DotM, DotN, IcmS and IcmW were found primarily in gel pieces 11-13 which correspond a complex size of approximately 450 kDa. This fits the approximate size of 300 kDa of the coupling complex with one copy of each participating component (DotL, DotM, DotN, DotY, DotZ, IcmS, IcmW and LvgA) surrounded by detergents. At the same height a strong band was observed for samples overexpressing HA-LegC2, suggesting that LegC2 is bound to the coupling complex. Only a faint band was observed for SidF and none at all for RalF and Lep-inv. As the expression of all three proteins is significantly lower that of LegC2, one cannot be certain if SidF and Lep-inv do

not interact with the coupling complex or if their expression is simply too weak to be properly detectable by Western blotting. With respect to RalF, crude membrane preparation and solubilization with 1 % (w/v) LMNG may result in loosening any interaction of the soluble effector with components of the coupling complex.

Overall, these results indicate that LegC2 and SidF are in complex with proteins of the coupling complex, strengthening the hypothesis that proteins of the coupling complex recognize TMD-effectors as substrates for the Dot/Icm system. Unfortunately, control experiments strains lacking single coupling complex components were not possible due to the instability of the complex when not all components are present.

4.6.2 TMD-effectors seem to be recognized by IcmSW and not DotM of the coupling complex

Although TMD-effectors seem to be integrated into the bacterial inner membrane prior to their secretion through the Dot/Icm system, they so far showed a remarkably similar behavior in their translocation to soluble effectors. In respect of the recognition of soluble effectors by the coupling complex of the Dot/Icm system, it is believed that they are either recognized by DotL with the help of IcmSW or have an E-block motif within their C-terminal signal and thus can bind directly to DotM and be translocated without any help from T4 chaperones.
In contrast to this hypothesis and as previously shown in chapter 4.3, all six investigated soluble as well as TMD-effector proteins did not only show an IcmSW dependence (Fig. 13) but also harbor an E-block motif within their C-terminal signal (Fig. 9). This seems also to be true for 65 % of the 98 analyzed TMD-effectors (Fig. 27A).
Furthermore, despite the presence of an E-block motif only a small decrease in translocation efficiency was observed for the TMD-effectors SidF and Ceg4 in a DotM-mutant (Fig. 27B) which was shown to be unable to bind Glu-rich peptides (Meir *et al.*, 2018). SidM, RalF, LegC2 and LegC3 were secreted at levels comparable to the wild type strain. This result is most noticeable for LegC3 which actually has two strong E-block motifs, EESDE and EEE. Only the translocation of SidF and Ceg4 in a DotM-mutant were reduced to 60 % and 50 % of the wild type strain. However, the expression levels of Ceg4 were reduced as well (appendix. 7.3).

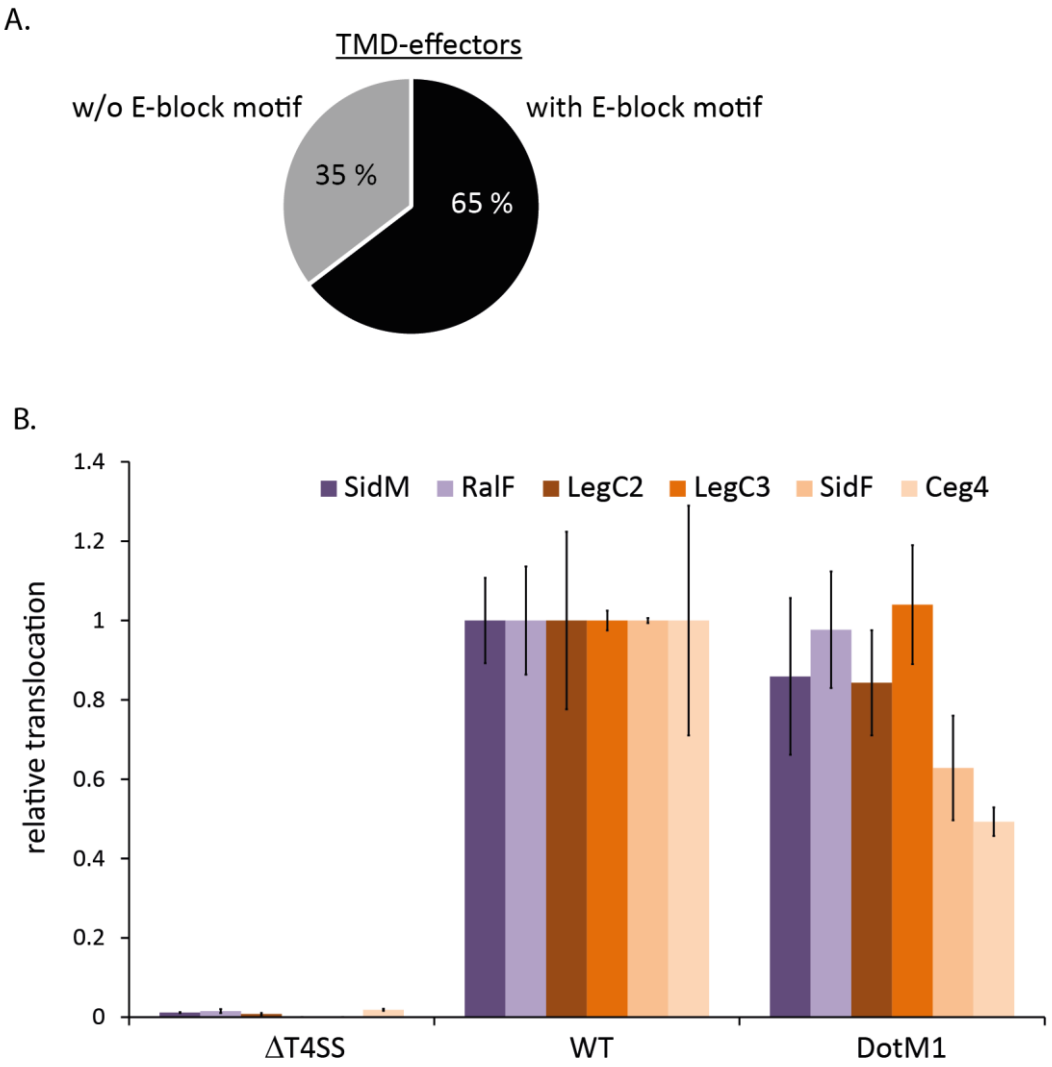


Fig. 27: Possible recognition of TMD-effectors due to their C-terminal E-block motif.
A) The C-terminal 30 amino acids of all predicted TMD-effectors were manually analyzed if they harbor an E-block motif. The classification of E-block motif is according to Huang *et al.* (Huang *et al.*, 2011).
B) *L. pneumophila* wild type, a ΔT4SS mutant strain or a DotM mutant (DotM1) with decreased ability to bind Glutamine-rich peptides (Meir *et al.*, 2018) expressing the Dot/Icm effector proteins ^{HiBiT}SidM, ^{HiBiT}RalF, ^{HiBiT}LegC2, ^{HiBiT}LegC3, ^{HiBiT}SidF or ^{HiBiT}Ceg4, respectively were used to infect Raw 264.7 macrophages which constitutively express LgBiT. Luminescence was measured over time using a plate reader. Only the luminescence signal 4 hours post infection is displayed. The data points represent the mean (± standard deviation) of at least three independent experiments.

This suggests that effectors of the Dot/Icm system cannot be classified as carriers of an E-block motif and by this IcmSW-independent on the one hand and on the other hand as IcmSW-dependent effectors without an E-block motif.

Overall, these results indicate that, similar to soluble effectors, TMD-effectors are probably recognized by the chaperones or adaptor proteins IcmSW, extracted from the bacterial membrane and hands over to the Dot/Icm system.

4.7 Is a C-terminal signal located in the cytoplasm and the presence of all Dot/Icm components sufficient for translocation of an inner membrane protein?

So far two essential properties have been identified for a successful translocation of TMD-effectors into host cells: the presence of the Dot/Icm chaperones and a C-terminal T4 signal within the effector facing the cytoplasmic side. To assess, if the presence of all Dot/Icm components and a T4 secretion signal at the C-terminus located in the cytoplasm is sufficient to translocate an inner membrane protein into host cells through the Dot/Icm system, the inverted leader peptidase, a well-studied inner membrane protein from *E. coli* with a N_{in}-C_{in} topology, was fused to the C-terminal 30 amino acids of SidM. The E-block motif of SidM was investigated by Huang *et al.* and could successfully translocate adenylate cyclase into U937 cells (Huang *et al.*, 2011). The inner membrane protein Lep-inv was chosen because it is not expressed endogenously in *L. pneumophila* and thus is not involved in any complex formation which could hinder targeting to or translocation through the Dot/Icm system.

Although a stable expression of the Lep-inv-SidMc30 fusion protein could be obtained, no translocation into Raw 264.7 macrophages could be measured (Fig. 28A). The N_{in}-C_{in} topology for this fusion protein was validated using TOPCONS prediction and the split NanoLuc-based topology assay (Fig. 28B). This indicates that the presence of a fully-assembled Dot/Icm system and a C-terminal secretion signal located in the cytoplasm is not sufficient to extract or secrete an unrelated inner membrane protein.

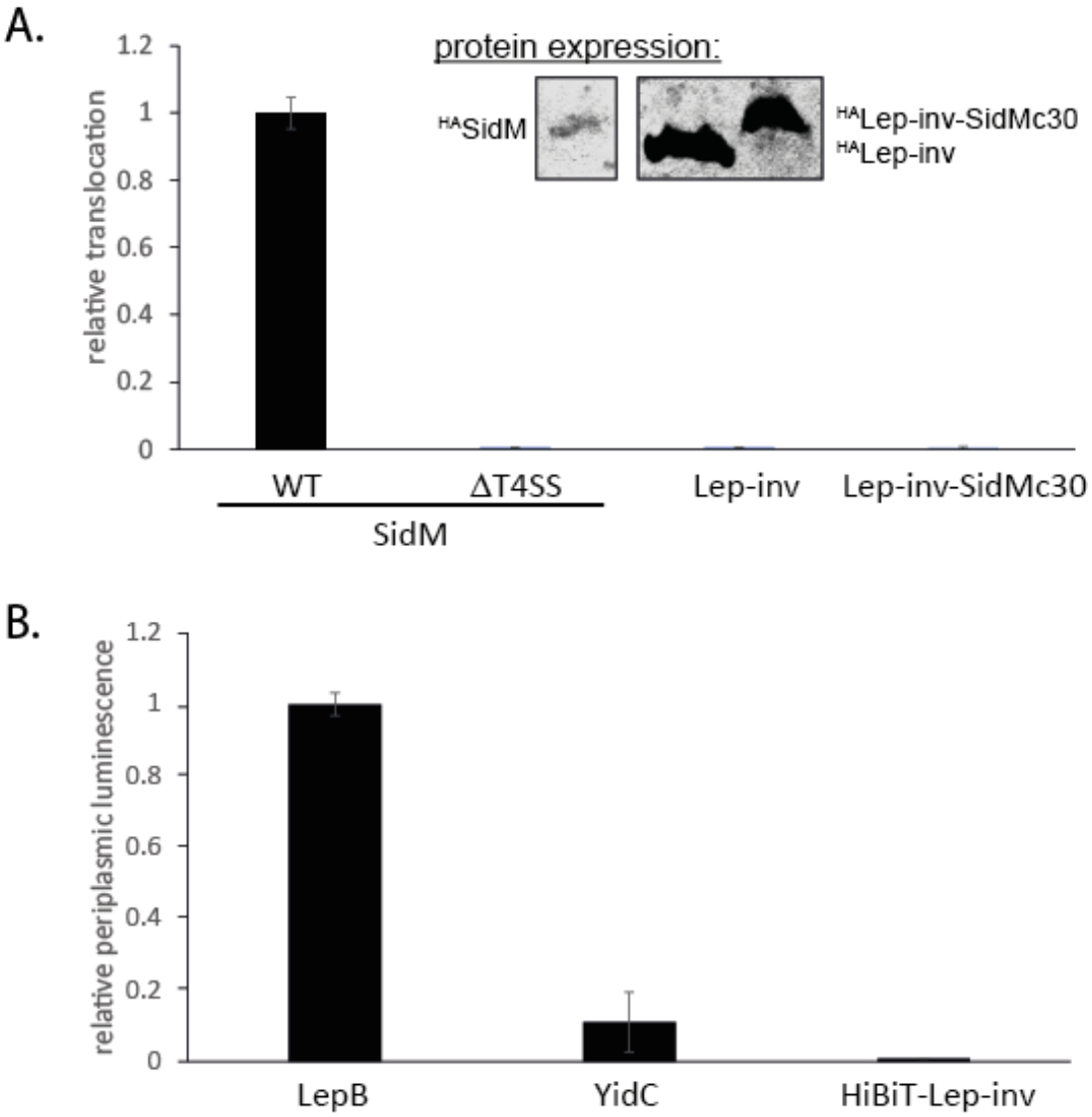


Fig. 28: Analysis of the artificial TMD-effector Lep-inv-SidMc30.

A) *Legionella pneumophila* wild type expressing ^{HiBiT}Lep-inv-SidMc30 was used to infect Raw 264.7 macrophages which constitutively express LgBiT. As controls *Legionella pneumophila* wild type expressing ^{HiBiT}Lep-inv, or ^{HiBiT}SidM as well as a ΔT4SS mutant expressing ^{HiBiT}SidM were used. Luminescence was measured over time using a plate reader. Only the luminescence signal 4 hours post infection is displayed. The data points represent the mean (± standard deviation) of at least three independent experiments. Protein expression was analyzed using SDS-PAGE and Western blotting.

B) The topology of ^{HiBiT}Lep-inv was analyzed using a split NanoLuc topology assay. The data points represent the mean (± standard deviation) of at least three independent experiments.

4.8 Which additional substrate-intrinsic properties are important for successful translocation?

In the previous section, I showed that the selected TMD-effectors harbor an E-block motif which is necessary but not sufficient for their successful translocation into host cells. Also, they depended on the presence of Dot/Icm chaperones IcmSW rather than interacting with the coupling complex component DotM. Further, I explored if IcmSW interacts with the C-terminal signal by deleting *icmS* or *icmW* in addition to the C-terminus of the TMD-effectors.

Surprisingly, three different effects were observed (Fig. 29A). First, deletion of *icmS* or *icmW* and the C-terminal translocation signal resulted in a lower translocation efficiency of the effector proteins SidM, LegC2 and Ceg4, respectively, compared to deleting the C-terminal signal or any of the chaperones individually. This indicates an additive effect and suggests a chaperone binding site which is not located within the C-terminal signal.

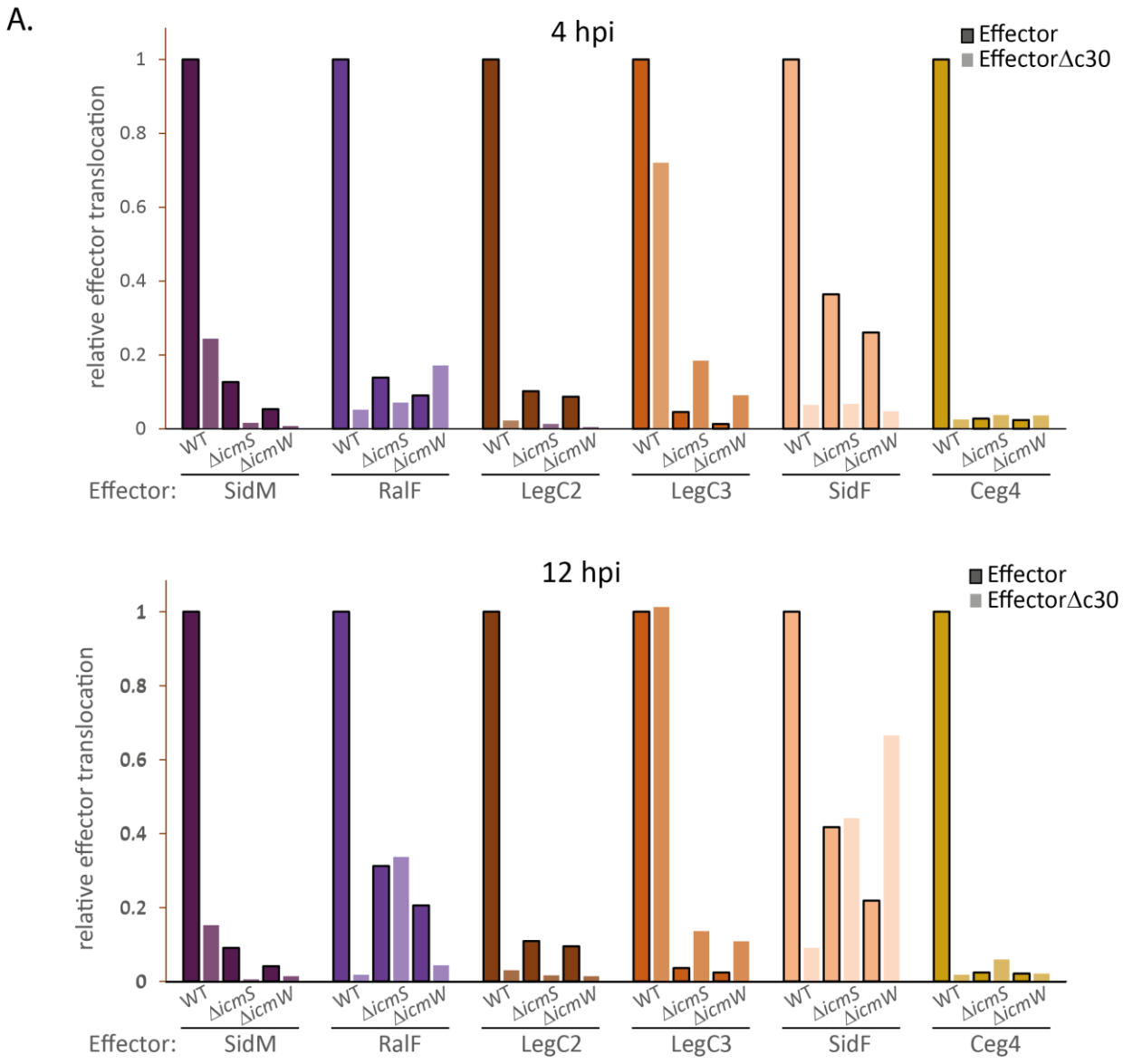
As shown above, LegC3 could be translocated at levels similar to wild type even without its C-terminal signal. The additional deletion of *icmS* or *icmW*, however, reduced the translocation efficiency of LegC3 to less than 20 %, clearly showing an IcmSW dependence even without the C-terminal signal present. Interestingly, full-length LegC3 can only be translocated into host cells to less than 5 % upon deletion of the T4 chaperones, showing a stronger dependence on IcmSW with the C-terminal signal present.

In contrast to these results, an increased translocation efficiency was observed for RalF as well as SidF when the C-terminal signal as well as *icmS* or *icmW* were deleted. This effect could however only be observed 12 hours post infection and was much more pronounced for SidF than for RalF.

It is worth noting that the same effect was observed when the C-terminus of SidF is located in the periplasm (Fig. 29B). Although periplasmic localization of the C-terminal signal completely attenuates the translocation of SidF, the additional deletion of *icmS* or *icmW* resulted in a translocation up to 20 % of wild type levels at 12 hours post infection. This effect was even more pronounced when on top of that the C-terminal secretion signal of SidF was deleted: SidFΔTMS2Δc30 could be translocated with an efficiency of 25 % when all Dot/Icm components were present or in the absence of only IcmW. In contrast, a translocation up to around 60 % was observed when *icmS* was deleted. These results indicate that there might be an internal signal within the N-terminal domain of SidF which facilitate slow but steady translocation of the protein. Interestingly, no enhanced translocation in comparison SidFΔTMS2 could be observed when 20 amino acids upstream to the C-terminal signal were additionally deleted (Fig. 29B).

Together, these results indicate that the C-terminus and the chaperones IcmSW influence each other. To validate the notion that IcmSW do not bind the C-terminal T4 signal, protein-protein interactions between IcmSW and the C-terminal signal of SidF were analyzed using *in vivo* photo-crosslinking.

Individual amber-positions for four sequential residues within the last 30 amino acids of SidF were introduced and subjected to UV irradiation. Although several crosslinks were observed (Fig. 29C), they are very likely not to IcmS or IcmW as they still can be detected in single deletion mutants of these chaperones (Fig. 29D).



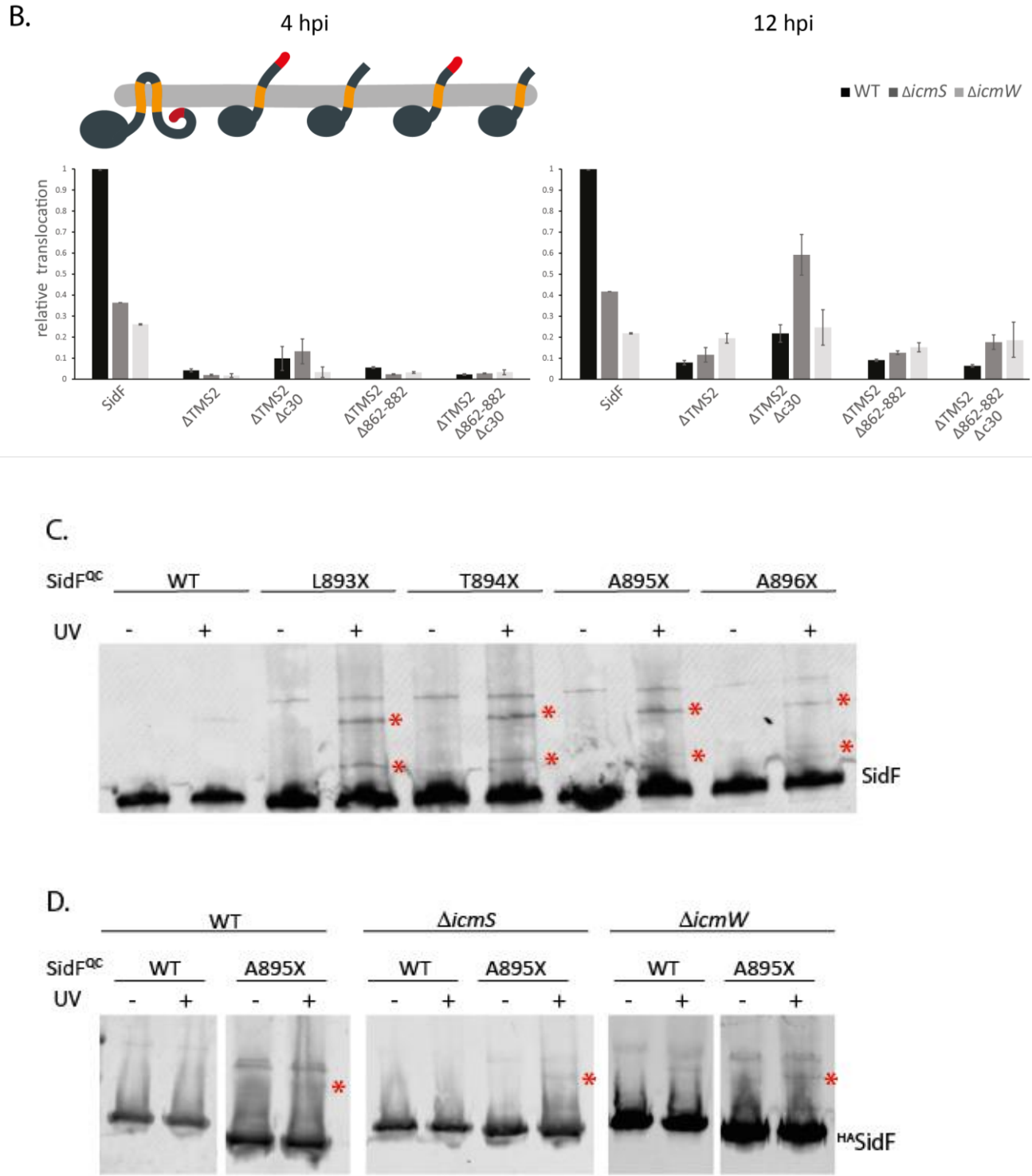


Fig. 29: A region exclusive of the C-terminal translocation signal interferes with the IcmSW dependence of Dot/Icm effectors.

L. pneumophila wild type strain or strains lacking *icmS* or *icmW*, respectively, expressing the Dot/Icm effector proteins ^{HiBiT}SidM, ^{HiBiT}RalF, ^{HiBiT}LegC2, ^{HiBiT}LegC3, ^{HiBiT}SidF or ^{HiBiT}Ceg4, respectively or the same effector proteins lacking their C-terminal translocation signal **(A)** or SidF truncations **(B)**, were used to infect Raw 264.7 macrophages which constitutively express LgBiT. Luminescence was measured over time using a plate reader. The luminescence signal 4 hpi (A, upper panel; B, left panel) or 12 hpi (A, lower panel; B right panel) is displayed. The data points represent the mean of at least three independent experiments.

C,D) *In vivo* photo-crosslinking of *L. pneumophila* wild type **(C)** or strains lacking *icmS* (dark grey) or *icmW* (light grey), respectively **(D)**. Strains expressing ^{HA}SidF as well as harboring the pEvol-pBpa plasmid were grown over night on agar plates supplemented with 0.5 mM IPTG and 1mM *p*Bpa. Bacteria were UV_{365nm} irradiated for 30 minutes. *p*Bpa mutations are denotated as "X". Immunodetection of ^{HA}SidF in whole cell lysate with and without UV irradiation is shown. A representative result of at least three independent experiments is shown.
*p*Bpa *para*-(benzoyl)-phenylalanine

These results suggest that although the chaperones IcmSW do not seem to bind the C-terminal T4 signal itself, they clearly still have an influence on the translocation process of soluble as well as TMD-effectors even if the C-terminal signal is not present, indicating another recognition mechanism which works independently of the C-terminal T4 signal. Unexpectedly, IcmSW do not only seem to support translocation but can also result in its reduction. Why this is only the case when the C-terminal signal is absent has yet to be elucidated.

4.8.1 Which substrate-intrinsic properties mediate successful translocation of the TMD-effector SidF

To further determine which internal regions of a TMD-effector are crucial for its successful translocation into host cells, I selected one TMD-effector to study in more depth. As LegC3 already showed an independence of its C-terminal T4 signal, it would have been most interesting to investigate which internal signal is recognized for its subsequent translocation. Unfortunately, LegC3 turned out to be rather unstable when more than the C-terminal 30 amino acids were deleted. This might be due to its coiled-coil structure which is possibly disturbed when interacting regions are deleted (Yao, Cherney and Cygler, 2014). On these grounds SidF was selected for further investigations (Fig. 30A).

As shown previously, deletion of the C-terminal translocation signal and a region located 20 amino acids upstream resulted in a different translocation behavior of SidF (Fig. 30B). Based on that, various truncation mutants for the C-terminus of SidF were analyzed. The region between the second TMS and the 30 residue-long C-terminal signal was shortened in 20 amino acid steps (Fig. 30A,B). Already the deletion of the 20 amino acids directly adjacent to the secretion signal abrogated any translocation of the protein into host cells. Only a very small amount of about 10 % could be translocated when *icmS* was additionally deleted. Deletion of additional parts of SidF led to a complete loss of translocation. As the C-terminal secretion signal is still present in all of these constructs, either an additional internal signal may be important, or the C-terminal signal is longer than 30 amino acids and is disrupted by the deletion of parts upstream to the last 30 residues. Surprisingly, when the C-terminal signal and its upstream 20 amino acids are absent, an increased translocation of SidFΔ822-882Δc30 can be observed

in strains lacking *icmS* or *icmW* (Fig. 30B). This was not the case for SidFΔ822-882Δc30, when more than 20 residues upstream of the C-terminal were missing. Except for SidFΔ862-882Δc30 all constructs exhibit a protein stability similar to the full-length SidF (Fig. 30C).

As SidF can unexpectedly be translocated without its C-terminal T4 signal when additionally, one of the chaperones IcmS or IcmW is deleted (Fig. 30B), I was interested to know whether that behavior is connected to the inner membrane localization of SidF. Thus, both TMSs of SidF were deleted, which should result in a cytoplasmic localization of the protein. As the protein expression was very different for these constructs (Fig. 30D), the relative translocation of the truncation mutants of the soluble effector SidFΔTMS1ΔTMS2 was normalized to the relative protein expression. Interestingly, a 2.5-fold higher translocation rate was observed for the soluble SidF (SidFΔTMS1ΔTMS2) in a wild type strain (Fig. 30E). However, the observed translocation still seems to be IcmSW-dependent. Additional deletion of the C-terminal 30 amino acids resulted in complete abolishment of translocation independent of the presence or absence of the T4 chaperones. To elucidate a possible internal signal in the region between the second TMS and the last 30 amino acids, the last 30 amino acids were directly fused to the soluble N-terminal domain of SidF. This resulted in a 50 % reduction of translocation 4 hours post infection in comparison to the full-length soluble SidF (SidFΔTMS1ΔTMS2), which is almost completely abrogated in the absence of IcmS or IcmW. When only the N-terminal domain of SidF without its C-terminal secretion signal was investigated, a translocation efficiency of about 30 % of the full-length soluble SidF (SidFΔTMS1ΔTMD2) was observed. These are levels similar to the inner-membrane integrated SidF. Again, almost no translocation was measured when IcmS or IcmW were absent.

In summary, SidFΔTMS1ΔTMD2 show a very similar translocation behavior to other soluble effectors such as SidM. Similar to the full-length inner-membrane integrated SidF, deletion of the region upstream of the C-terminal T4 signal diminished translocation to a certain degree. Intriguingly, the soluble N-terminal domain of SidF could still be translocated even without C-terminal signal and showed an IcmSW dependence similar to soluble effectors.

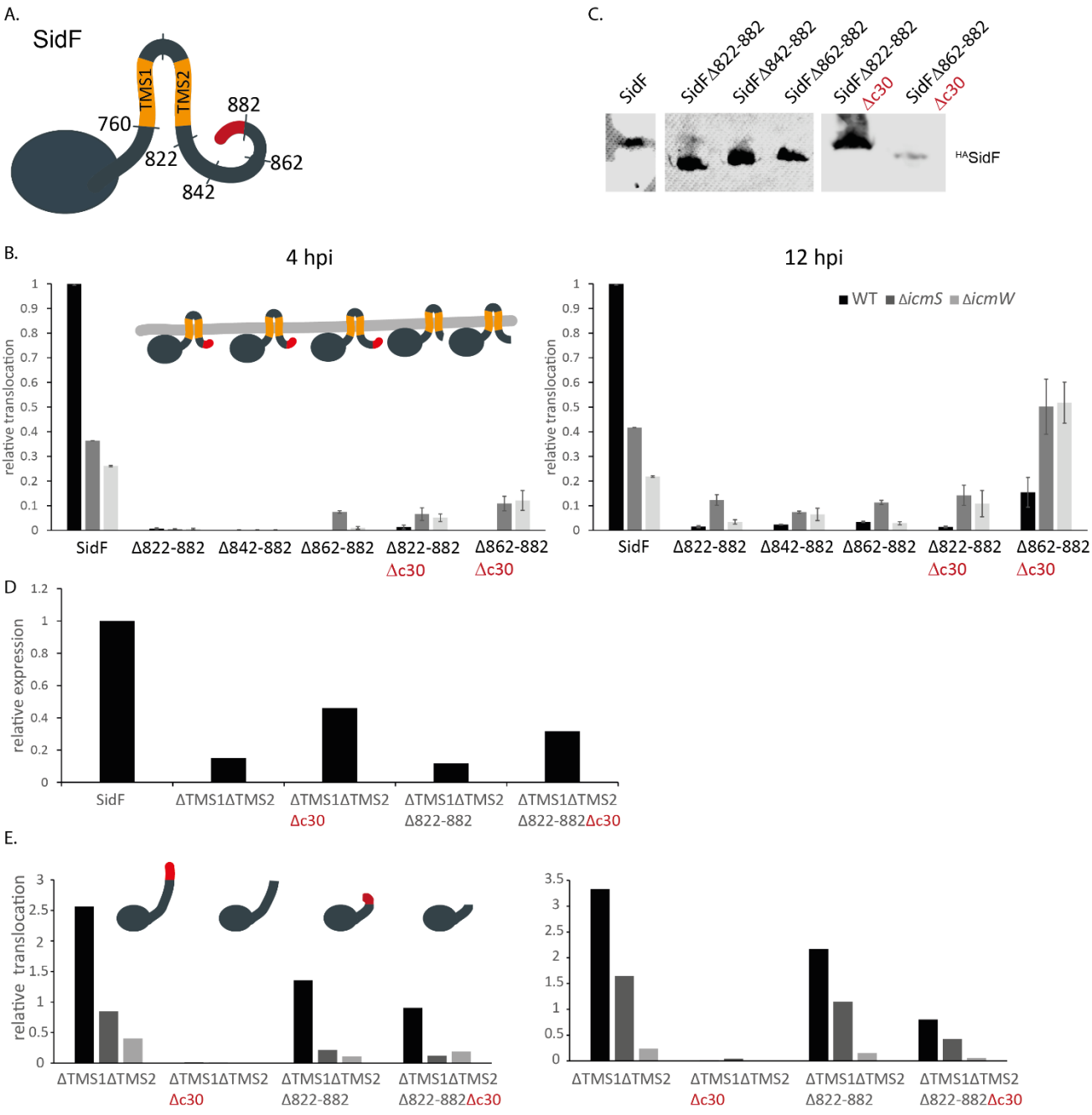


Fig. 30: Analysis of SidF truncation mutants

A) Overview of truncation sites within the TMD-effector SidF

B, D) *L. pneumophila* wild type or strains lacking *icmS* or *icmW*, respectively, expressing the Dot/Icm TMD-effector ^{HiBiT}SidF **(B)** or the soluble effector SidFΔTMS1ΔTMS2 **(D)** and truncation mutants of both were used to infect Raw 264.7 macrophages which constitutively express LgBiT. Luminescence was measured over time using a plate reader. The luminescence signal 4 hpi (left panel) or 12 hpi (right panel) is displayed. The data points represent the mean (± standard deviation) of at least three independent experiments.

C) The expression of ^{HiBiT}SidF or truncation mutants were analyzed using SDS-PAGE and Western blotting. A representative result of at least three independent experiments is shown.

D) The expression of ^{HiBiT}SidF or truncation mutants of the soluble effector SidFΔTMS1ΔTMS2 was analyzed by luminescence measurement.

E) The relative translocation of the truncation mutants of the soluble effector SidFΔTMS1ΔTMS2 was normalized

to the relative protein expression.

Together, these results suggest that the C-terminal signal of SidF is longer than its last 30 amino acids and the whole signal is necessary for its successful translocation into host cells. Moreover, distinct differences in the translocation behavior of SidF can be observed dependent on its subcellular localization. Although the membrane-embedded full-length SidF needs its C-terminal secretion signal for successful translocation, a loss of the signal can be partially overcome by additional deletion of one of the T4 chaperones *icmSW*, which is not the case for the soluble SidF SidFΔTMS1ΔTMS2. It is worth noting that apparently this also resulted in a different translocation kinetic as this effect could only be measured 12 hours post infection. In contrast, as soon as SidF cannot be integrated into the bacterial inner membrane due to the deletion of its TMSs, the effector shows an IcmSW-dependence very similar to proper soluble effectors like SidM. This strongly suggests an IcmSW binding site within the N-terminal domain of SidF.

Overall, these results suggest that there is more than one mechanism by which TMD-effectors are recognized leading to their subsequent translocation.

5 Discussion

Many gram-negative as well as gram-positive bacteria depend on functional secretion systems for their pathogenicity. The T3SS and T4SS are used, among others, by *Salmonella* Typhimurium and *Legionella pneumophila*, respectively, to translocate effector proteins in one step into host cells in order to hijack and subvert host cell functions. Protein export is especially challenging for gram-negative bacteria as secreted proteins have to cross at least three membranes, including the host cell plasma membrane.

One important subgroup of effectors which promote intracellular survival of bacteria such as *Salmonella* Typhimurium or *Legionella pneumophila* are transmembrane proteins (TMD-effectors). After being translocated from *L. pneumophila* into host cells such as amoeba or lung macrophages, TMD-effectors are mostly found integrated into the *L*CV (De Felipe *et al.*, 2008; Toulabi *et al.*, 2013). Here, they alter the host cell vesicle trafficking in order to ensure *L*CV biogenesis. Furthermore, they change the appearance of the *L*CV by recruiting ribosomes and by the modification of phospholipids (Tilney *et al.*, 2001; Toulabi *et al.*, 2013).
In contrast to their function in host cells, it has not yet been elucidated how these substrates are targeted to and recognized by the secretion system within the bacterium.

In this study, I investigated the targeting and recognition mechanism of TMD-effectors in *Legionella pneumophila* strain Philadelphia-1, focusing on their similarities as well as differences to well-known soluble effectors. Selected TMD-effectors seem to behave very similarly to well-known soluble effectors such as SidM and RalF in their translocation into host cells, depending for instance on the presence of a C-terminal translocation signal as well as the Dot/Icm chaperones IcmSW. Nevertheless, the strong hydrophobicity of their TMS makes the co-translational targeting by SRP and a subsequent inner membrane insertion very likely. I analyzed if TMD-effectors can be found in the bacterial inner membrane and if membrane insertion can be prevented by chaperone binding. Moreover, I focused on how TMD-effectors are recognized by the Dot/Icm system and which components of the machinery and which substrate-intrinsic properties such as their membrane topology, IcmSW binding sites or additional internal signals are important for their successful translocation into host cells.

5.1 Targeting of TMD-effectors to the Dot/Icm system in *L. pneumophila*

Proteins synthesized in the cytosol must be targeted to the right destination (e.g. inner or outer membrane, periplasm or a specific host cell compartment) in order to fulfil their specific function. Mis-targeting is often toxic for the cell and associated with diseases. Targeting information is often

contained in N-, C-, or internal signals. Once the nascent chain emerges from the ribosome, competition for the protein between different targeting factors begins.

With respect to TMD-effectors of the Dot/Icm system in *L. pneumophila*, it is of great importance to understand if effector targeting happens in a co- or posttranslational manner. Unlike soluble effectors, TMD-effectors harbor hydrophobic stretches which are needed for successful integration into the lipid bilayer of their target membrane. As hydrophobic stretches can be co-translationally recognized by SRP, inner membrane targeting cannot be avoided if targeting to the Dot/Icm system depends only on the C-terminal secretion signal. With this there are several questions to answer: Are TMD-effectors actually recognized by SRP? Are there any signals within TMD-effectors that prevent SRP recognition? If TMD-effectors are indeed inserted into the bacterial inner membrane, how are they recognized in and extracted from the inner membrane prior to their translocation into host cells by the Dot/Icm system?

As shown for TMD-effectors in T3SS (Krampen *et al.*, 2018), effector-intrinsic properties such as a balanced hydrophobicity of the TMS or chaperone-binding can result in a passive or active avoidance of SRP targeting. In contrast, T4SS have evolved to also accept substrates from the periplasm or from within the bacterial membrane. The pertussis toxin in *B. pertussis* for instance is translocated across the bacterial inner membrane using the Sec system. After proper folding and oligomerization, the active form of the toxin is translocated across the outer membrane in a T4SS-dependent manner (Burns *et al.*, 2003; Locht, Coutte and Mielcarek, 2011). Effector proteins of the VirB T4SS of *Brucella* spp. are translocated in a similar fashion, using the Sec-machinery in a first step and then T4SS to cross the bacterial outer membrane (de Jong *et al.*, 2008; De Barsy *et al.*, 2011; Marchesini *et al.*, 2011).

Based on this, TMD-effectors of Dot/Icm system of *L. pneumophila* may take a similar pathway, including inner membrane targeting and insertion as a part of their secretion mechanism instead of trying to avoid it.

In order to determine if TMD-effectors of the Dot/Icm system are avoiding or embracing SRP-targeting, I analyzed the hydrophobic nature of the TMS present in a TMD-effector. Due to the thickness of the lipid bilayer a hydrophobic stretch of 18 to 40 amino acids is required for proper membrane integration. Moreover, TMS are recognized by the Sec translocon based on their thermodynamic partitioning into the lipid bilayer which can be expressed as apparent free energy of insertion (ΔG_{app}). A ΔG_{app} value of 0 corresponds to an insertion of 50 % of the molecules (Hessa, Nadja M Meindl-Beinker, *et al.*, 2007). Furthermore, Hessa *et al.* could show that not only the residue type but their position within the TMS as well as the length of the TMS and flanking amino acids influence the ΔG_{app} value. The hydrophobic nature of TMS is, however, not only important for membrane insertion but also for the co-translational targeting of inner membrane proteins by SRP. Schibich *et al.* could show

that SRP binds to a limited hydrophobic stretch of 12-17 amino acids containing at least four hydrophobic residues (Schibich *et al.*, 2016). Moreover, for membrane proteins in *E. coli* it has been described that a lower ΔG_{app} is needed for SRP targeting than for membrane integration itself (using a window of 19-23 amino acids) (Lee and Bernstein, 2001; Schibich *et al.*, 2016). If the hydrophobicity of the first TMS within protein is too low, it is often skipped by SRP. Similar to membrane integration, flanking basic amino acids facilitate SRP binding (von Heijne, 1985). With this, one can say that SRP-targeting needs a high hydrophobic density whereas membrane integration can be achieved with a lower hydrophobicity but sufficient length of the hydrophobic segment.

Based on this, I used the ΔG predictor (Hessa, Nadja M Meindl-Beinker, *et al.*, 2007) to analyze the membrane integration potential as well as SRP-targeting propensity of TMD-effectors of the Dot/Icm system in *L. pneumophila*.

Intriguingly, almost 30 % of all effector proteins known so far in *L. pneumophila* harbor one or more TMS. An intermediate hydrophobicity between -1.5 and 0.5 kcal/mol was predicted using the ΔG predictor program and the "SRP"-window of 12-17 amino acids. Furthermore, preliminary data based on *E. coli* model systems showed no membrane integration for TMS of three of four TMD-effector proteins of the Dot/Icm system (Krampen *et al.*, 2018). These data suggest that, similar to TMD-effectors of the T3SS in *Salmonella*, a subgroup of TMD-effectors seems to be sufficiently hydrophobic for membrane integration but not for being recognized and targeted by SRP. This balanced hydrophobicity can be used as a passive avoidance mechanism in order to prevent inner membrane targeting and facilitate direct targeting to the secretion system (Krampen *et al.*, 2018). In contrast to TMD-effectors of the T3SS, there are also TMD-effectors within *L. pneumophila* that exhibit a rather high hydrophobicity within their TMS which makes their SRP targeting very likely.

TMD-effectors harboring TMS with higher hydrophobicity can in principle take one of three possible routes after their translation in the cytoplasm: misfolding and protein degradation by bacterial quality control systems, inner membrane insertion or direct targeting to the Dot/Icm system protected by bound chaperones.

Hydrophobic patches of TMD-effectors can be recognized as signals for the heat shock response system and subsequently degraded, as soon as they emerge from the ribosome, to prevent mis-folding and aggregation (Young *et al.*, 2004). It is, however, known that expression and assembly of the Dot/Icm system as well as effector expression takes place prior to host cell contact at a different point of their bi-phasic life cycle (Aurass *et al.*, 2016) making it necessary to keep effector proteins in a secretion competent state until the secretion machinery switches from an inactive to an active state. For most soluble effectors this most probably happens by chaperone binding. For the TMD-effector SidF I could show a stability for at least 1-2 hours in the bacterial cytoplasm suggesting that they are

not directly degraded (data not shown.). This can in principle be achieved by chaperone binding or inner membrane integration.

In principle, TMS of TMD-effectors can be co-translationally recognized by SRP and thereby targeted to and integrated into the bacterial inner membrane where they are recognized by the Dot/Icm machinery and translocated into host cells. Using membrane fractionation by a sucrose gradient centrifugation protocol as well as urea extraction, I could show that TMD-effectors in *L. pneumophila* are not only predicted to be targeted by SRP but can indeed be found properly integrated into the bacterial inner membrane (Fig. 19,20). Surprisingly, this was also true for TMD-effectors with lower hydrophobicity. This indicates that the semi-hydrophobic nature of some TMS does not prevent targeting to and integration into bacterial inner membrane. Since I was not yet able to show that effector proteins can be translocated into host cells after being integrated into the bacterial inner membrane, mistargeting to the inner membrane cannot be excluded.

To protect their hydrophobic patches and avoid aggregation or recognition by SRP, TMD-effectors could be bound by chaperones. For TMD-effectors of the T3SS it was shown that cognate chaperones bind co-translationally to a chaperone binding domain at the N-terminus of effector proteins and by this protect TMS from being recognized and bound by SRP (Krampen *et al.*, 2018). This was even the case, when the hydrophobicity of TMS was increased. Although chaperones of the Dot/Icm system show very similar characteristics to chaperones of the T3SS, it is unclear if this active avoidance mechanism also applies for T4SS TMD-effectors since binding sites for the Dot/Icm chaperones IcmSW are very ill defined. To determine the general involvement of Dot/Icm chaperones in the translocation process of TMD-effectors, I investigated if the selected TMD-effectors LegC2, LegC3, SidF and Ceg4 depend on IcmSW in order to get successfully translocated into host cells. All investigated TMD-effectors showed an IcmSW dependence similar to the soluble effectors SidM and RalF (Fig. 13), suggesting that IcmSW may bind to TMD-effectors and by this possibly prevent SRP targeting. To investigate if IcmSW are involved in a potential active avoidance mechanism of SRP-binding, I tried to identify a chaperone-binding site within SidF. Unfortunately, no binding site upstream of the TMS of SidF was found using *in vivo* photo-crosslinking (Fig. 21). It is, however, important to note that there is no evidence so far that binding of chaperones must occur directly upstream or at the TMS in order to protect the hydrophobic stretches from SRP recognition. It is conceivable that chaperone binding induces conformational changes of the protein thus hiding hydrophobic stretches. This might be important especially when the second TMS possesses a higher hydrophobicity as it is the case for SidF. Based on this, the existence of an IcmSW binding site within SidF cannot be excluded. Now that in vivo photo-crosslinking in *L. pneumophila* has been successfully established many more possible crosslinking sites as well as other TMD-effectors can be investigated.

Nevertheless, if IcmSW binding would result in the avoidance of SRP targeting and inner membrane integration, it should be possible to find TMD-effectors bound to chaperones in the cytoplasmic fraction of the cell. This applies for effector proteins bound to chaperones of the Dot/Icm system as well as to other proteins such as the TF which can bind to the translating ribosome and avoid misfolding of proteins with its unfoldase activity until they might be handed over to chaperones of the Dot/Icm system (Hartl and Hayer-Hartl, 2002; Hoffmann *et al.*, 2012). After subcellular fractionation, TMD-effectors, in contrast to soluble effectors, could not be found in the cytoplasmic fraction of a sucrose gradient. Moreover, I analyzed if the subcellular localization of TMD-effectors changes in a strain overexpressing IcmSW. Interestingly, an increased concentration of IcmSW could not prevent inner membrane insertion of TMD-effectors (Fig. 20). These data support the notion that IcmSW are important for the successful translocation of TMD-effectors into host cells but presumably act more like adaptor proteins, bound to the coupling complex of the secretion machinery and involved in post-translational recognition of effector proteins rather than being involved in a co-translational targeting process.

In summary, these data suggest that TMD-effectors of the Dot/Icm system are recognized by SRP and subsequently intergraded into the bacterial membrane following a two-step targeting mechanism with an inner membrane intermediate.
It is important to note, that some TMD-effectors of the T3SS are toxic for the bacterial cell when mis-targeted to the inner membrane (Krampen *et al.*, 2018). One example are hydrophobic translocators which are secreted by the T3SS in order to form pores in the host cell membrane to inject T3SS effector proteins directly into the host cell cytoplasm (Büttner, 2012). In contrast to the T3SS, the Dot/Icm system in *L. pneumophila* does not have a needle to inject effector proteins but deliver their effector proteins most likely by membrane fusion. Consistent with this, so far, no pore forming substrates of the Dot/Icm system are known. As there are several times more substrate proteins known for the Dot/Icm system than for T3SS, different targeting pathway to the secretion machinery for different effector subgroups could be more efficient for the bacterium.

5.2 Recognition of TMD-effectors by the Dot/Icm system

In addition to the targeting process of TMD-effectors to the Dot/Icm machinery, their proper recognition by the secretion machinery plays a central role for a successful translocation into host cells. Soluble effectors have to be recognized by components located in the cytoplasm whereas TMD-effectors can in principle be recognized at the cytoplasmic or periplasmic side of the IM after their presumed integration into the bacterial inner membrane.

To investigate a possible recognition site for TMD-effectors, I determined the membrane topology of TMD-effectors.

The prediction of the transmembrane topology of the more hydrophobic TMD-substrates by TOPCONS (Tsirigos *et al.*, 2015) indicated that the vast majority of substrates possesses only very short periplasmic stretches of a length of 10-20 amino acids (Fig. 21). Moreover, in most cases the C-terminal end is predicted to be located in the cytoplasm. This was confirmed for the TMD-effectors SidF and LegC3 using a split NanoLuc topology assay. Only a small luminescence signal similar to the cytoplasmic control YidC was observed when HiBiT was individually fused to the N- or C-terminus of the proteins of interest. Once the second TMS of both proteins was deleted a strong luminescence signal showed the presence of the C-terminus in the periplasm, confirming the predicted N_{in}-C_{in} topology for full-length SidF and LegC3 (Fig. 24). With this, only very few amino acids between the first and second TMS of these TMD-effectors could possibly interact with other proteins in the periplasm upon their inner membrane integration.

DotF and DotK are two components of the Dot/Icm system located in the periplasm which might act as putative substrate receptors (Sutherland *et al.*, 2013). However, no significant reduction of neither TMD-effector nor soluble effector translocation could be observed upon DotK deletion, suggesting no central role of DotK in effector recognition. Deleting DotF resulted in a reduced translocation of not only the TMD-effectors LegC2, LegC3 and Ceg4 but also the soluble effectors SidM, suggesting an overall instability of the system (Fig. 23). Interestingly, the absence of DotF resulted in a complete abrogation of translocation of SidF and RalF, which implies an important role of DotF in either the recognition or export of the effector proteins. Intriguingly, there have been controversial reports about a possible receptor function of DotF. Using a two-hybrid assay, interactions of the effector protein RalF with DotF were shown by Sutherland *et al.* in 2013. Based on that, more effector proteins have been identified, including SidF (Sutherland *et al.*, 2013). The investigation of several truncation mutants of DotF by Sutherland *et al.* revealed that the interaction site is found within the periplasmic domain of the protein, supporting the hypothesis that DotF does not act as receptor in the cytoplasm but interacts with RalF during its export. Although, in principle SidF could interact with DotF in the periplasm after being integrated into the bacterial inner membrane, it seems rather unlikely with only four amino acids as possible interaction site. Furthermore, the cytoplasmic presentation of the C-terminal signal seems to be crucial as its periplasmic localization of SidFΔTMS2 resulted in the complete abrogation of translocation (Fig. 24).

Together with the fact that the selected TMD-effectors depend on the presence of IcmSW which is located in the cytoplasm either as a free cytoplasmic form or bound to DotL, recognition of TMD-effectors at the cytoplasmic side of the bacterial membrane seems very likely.

In literature it is described that soluble effectors are recognized by components of the coupling complex in two different ways: Effector proteins are either bound by IcmSW and by this targeted to components of the coupling complex, or they are IcmSW-independent by harboring a negatively charged E-block motif within their C-terminal T4 signal which interacts with the positively charged surface of DotM. Lifshitz *et al.* characterized the properties of the C-terminal signal of Dot/Icm effectors and stated that a strong secretion signal consists of a large glutamic acid stretch and several hydrophobic residues at the C-terminal end of the effector (Lifshitz *et al.*, 2013). Moreover, they could show that effectors with such a strong signal are less IcmSW-dependent than effectors with a weak signal. An increasing number of acidic residues within the C-terminal signal and with this a higher independence of the chaperones IcmSW was also observed for several Dot/Icm effectors in *Coxiella burnetii* (Larson, Beare and Heinzen, 2019). In contrast, I could show that all six investigated soluble as well as TMD-effector proteins did not only show an IcmSW dependence (Fig. 13) but also harbor an E-block motif within their C-terminal signal (Fig. 9) on which they, except for LegC3, depend for successful translocation into host cells. Interestingly, disrupting identified interaction sites in DotM by point mutations (Meir *et al.*, 2018), did not interfere with translocation of LegC2, LegC3, SidF or Ceg4 nor with the soluble effectors SidM and RalF (Fig. 27), indicating that the presence of an E-block motif does not necessarily result in the recognition of the effector protein by the coupling complex component DotM.

With this, one can say that in spite of being presumably integrated into the bacterial inner membrane, TMD-effectors behave similar to the soluble effectors SidM and RalF in their translocation depending on the presence of a C-terminal signal and the help of the chaperones IcmSW.

However, one must keep in mind that deletion of *icmS* or *icmW* does not only prevent any direct interaction between effector proteins and the respective chaperone but also results in a destabilization of the entire coupling complex. Neither DotL nor DotM can be detected anymore in the stationary phase of *Legionella* growth when IcmSW are absent (Vincent *et al.*, 2012). Therefore, deletion of the full-length *icmS* or *icmW* protein does not reveal if there are direct interactions between the chaperones and investigated effector proteins or if other proteins of the coupling complex are involved in the recognition process. Interestingly, the absence of the Dot/Icm chaperones does not result in any severe growth defect or attenuation of the contact dependent cytotoxicity (Coers *et al.*, 2000), which implies that enough effectors can still be successfully translocated into host cells. It is thinkable, that effector proteins also have a sufficient affinity for the Dot/Icm system to be secreted without the presence of the main effector receptor DotL. Another possibility is that there are Dot/Icm components independent of the coupling complex which may act as receptors.

Although TMD-effectors are probably recognized in a similar fashion by the coupling complex of the Dot/Icm system as soluble effectors, they may need to be extracted from the bacterial inner

membrane prior to being translocated into host cells. As TMD-effectors tend to exhibit small periplasmic loops but sizable cytoplasmic domains, extraction towards the cytoplasmic face of the membrane seems logical.

In order to corroborate effector recognition and subsequent extraction to the cytoplasmic side of the inner membrane, protein domains with increasing sizes were incorporated into the periplasmic loop of the TMD-effector SidF. Already the insertion of the 11 amino acid long HiBiT peptide resulted in a decrease of translocation by 30 % and UbiquitinE3GE13G-SidF could only be translocated up to 20 % in comparison to the wild type SidF (Fig. 25). These results not only strengthen the hypothesis of an extraction from the inner membrane towards the cytoplasmic side but also the two-step translocation with an inner membrane intermediate itself as it was shown by Amyot *et al.* that UbiquitinE3GE13G fused to the soluble effector RalF did not hinder its translocation (Amyot, DeJesus and Isberg, 2013).

Although I could show that a C-terminal translocation signal located in the cytoplasm is important for the translocation of 3 of the 4 selected TMD-effectors (Fig. 12), additional substrate-intrinsic properties have to be important as the fusion of the C-terminal 30 amino acid of SidM to the inner membrane protein Lep-inv was not sufficient for its translocation into host cells (Fig. 28). Moreover, all selected TMD-effectors show an IcmSW dependence in their translocation, suggesting a possible IcmSW binding site within the protein. The simultaneous deletion of the C-terminal signal and either IcmS or IcmW resulted in the total loss of translocation of the effectors SidM, LegC2, Ceg4 and the soluble form of SidF (SidFΔTMS1ΔTMS2) whereas upon the individual deletion of their C-terminal signal or the chaperones a small but detectable translocation into host cells could be observed (Fig. 29 and 30). This additive effect suggests a chaperone binding site outside the C-terminal signal. The same seems to be true for SidF. Surprisingly, the simultaneous deletion of the C-terminal signal and either *icmS* or *icmW* resulted in an increase instead of a decrease of translocation in comparison to the individual deletion effects (Fig. 29 and 30). Until now, the repression of translocation by IcmS was only shown for some effectors in *Coxiella burnetii*. In 2019, Larson *et al.* identified three different profiles of IcmS dependence for effector proteins in *Coxiella burnetii* (Larson, Beare and Heinzen, 2019): IcmS-independent, IcmSW-dependent and IcmSW-inhibited. SidF is not a clear IcmSW-inhibited effector as the absence of IcmS or IcmW resulted in a decrease of translocation efficiency by more than 60 % compared to the wild type strain. IcmSW seems to repress the translocation of SidF only after deletion of its C-terminal signal. Moreover, this effect was only observed after 12 hours post infection, whereas the IcmS dependence of effectors in *Coxiella burnetii* seems to be constant throughout infection (Larson, Beare and Heinzen, 2019). Interestingly, the same observation was obtained for the soluble effector RalF albeit to a much lower extent. This indicates that there might be sites within the effector proteins which can change IcmSW dependence. With respect to the localization of such sites, deletion of the 20 amino acids upstream of the C-terminal signal eliminates the repression of IcmSW for SidF

lacking its C-terminal signal. Similarly, Larson *et al.* could show that “residues encompassed within the region from -50 to -40 control IcmS dependency of” effectors of *Coxiella burnetii.* Another profile of IcmSW dependence could be observed for the TMD-effector LegC3. At first glance, it showed a similar IcmSW dependence as SidM, LegC2 and Ceg4. It could, however, be successfully translocated without the presence of its C-terminal signal, strongly suggesting an additional internal signal. Moreover, LegC3Δc30 seems to be less dependent on IcmSW than the full-length protein suggesting some kind of regulatory site within the C-terminal 30 amino acids of LegC3 (Fig. 29).

Overall, next to the C-terminal translocation signal there seem to be other signals and regulatory sites within effector protein which influence the translocation process. First of all, and comparable to many soluble effectors, the selected TMD-effectors very likely have IcmSW binding sites, although their exact localization has not been identified yet. Secondly, at least in SidF, RalF and LegC3 there appear to be sites which modulate IcmSW dependence. Based on the indication stated above that such sites reside within or directly upstream of the C-terminal signal, the subcellular localization of the same seems to be of no greater importance as the same effect could be observed for SidF with its C-terminus in the cytoplasm or periplasm (Fig. 29). On top of that, some effectors such as SidM and LegC3 presumably have additional internal signals which, however, seem to be important at different time points in translocation. LegC3Δc30 can be successfully translocated into host cells at levels comparable to the full-length protein 12 hours post infection whereas SidMΔc30 reaches a maximum of translocation of 20 % at 4 hours post infection compared to the full-length SidM protein. It has already been reported that two independent signals in the soluble effector SidJ are used to secret the effector protein at different time points during host cell infection (Jeong, Sutherland and Vogel, 2015). This suggests that the variety of translocation signals within and between effector proteins might be connected to the timing of translocation of different sets of effectors.

It is important to note that all different translocation profiles described above always include at least one TMD-effector and one soluble effector meaning that so far except for the difference in subcellular localization there does not appear to be a specific translocation mechanism for TMD-effectors. It is however obvious that effector recognition and translocation by the Dot/Icm system in *L. pneumophila* has a complexity which was so far unknown.

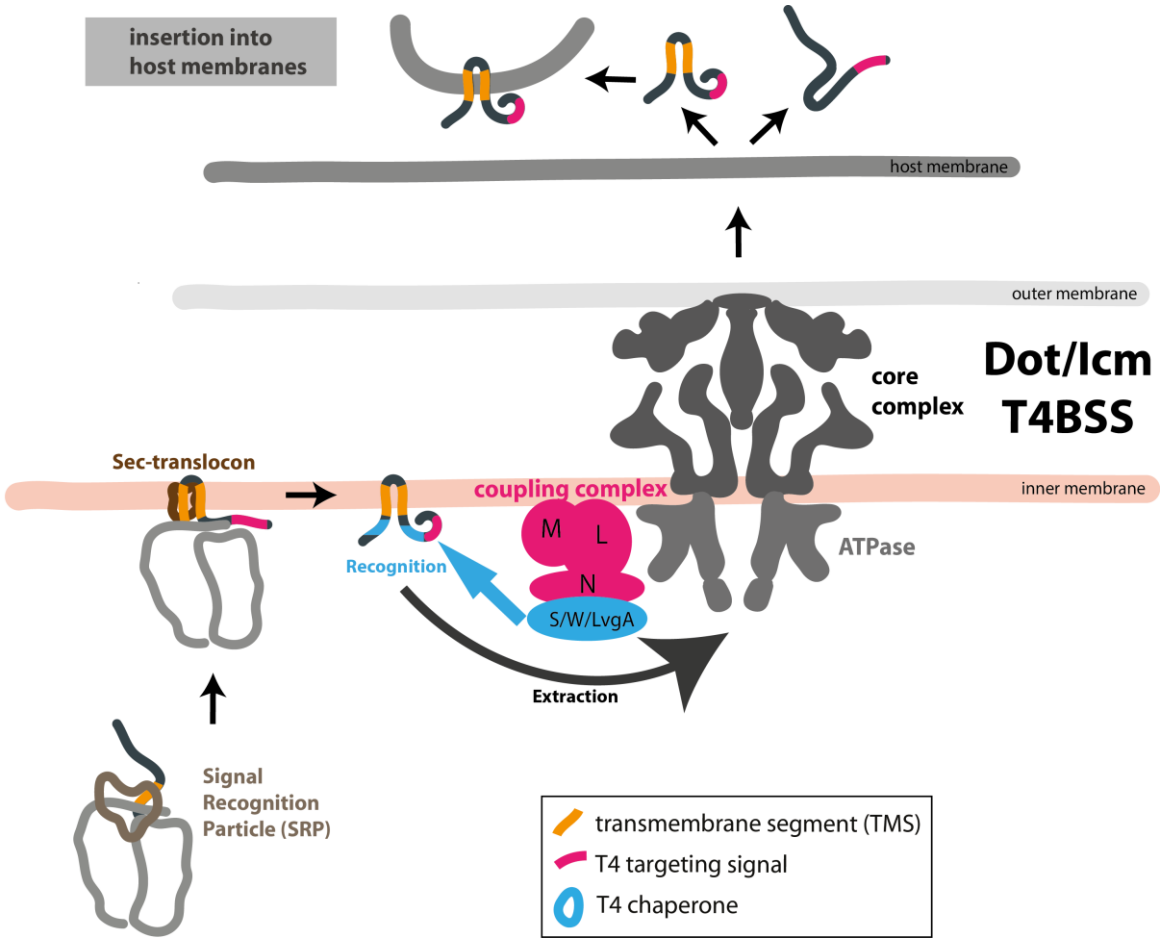


Fig. 31: Model of a two-step secretion mechanism of TMD-effectors.
TMD-effectors are pos-translationally recognized by SRP and subsequently targeted to and "anchored" into the bacterial inner membrane. The adaptor proteins IcmSW bound to the coupling protein DotL recognizes them as substrates of the Dot/Icm system. Next, TMD-effectors are extracted from the bacterial inner membrane towards the cytoplasmic side and handed over to the Dot/Icm system for translocation into host cells.

5.3 Outlook

This study could give first insights into a novel targeting and recognition mechanism of TMD-effectors. Based on the presented data, TMD-effector seems to embrace SRP-targeting as a part of their translocation mechanism and by this circumvent a possible targeting conflict. To confirm this two-step secretion mechanism with an inner membrane intermediate, translocation from within the bacterial inner membrane must be demonstrated. The different experimental approaches described in this study can be optimized and then used to determine which substrate-intrinsic information leads to the translocation of TMD-effectors. Ultimately, the aim would be to be able to transform any transmembrane protein into a substrate of the Dot/Icm system given the information obtained from these analyses.

Moreover, there is still much unknown regarding the role of the chaperones of the Dot/Icm system. In this study, I mostly focused on selected TMD-effectors with just two soluble effectors as comparison. It would, however, be very interesting to know at what time in effector expression, targeting and recognition in general the chaperones IcmSW are involved. The newly established *in vivo* photo-crosslinking procedure in *L. pneumophila* will be helpful to detect effector-chaperone complexes at different stages during effector targeting, recognition and translocation.

After leaving the bacteria due to successful translocation, it still remains unclear how effector proteins find their way inside the host cell. Future studies must pose question like: What is the targeting pathway of effector proteins within host cells? Are there any specific targeting signals or chaperones involved? How are TMD-effectors specifically protected from mis-targeting and degradation? How are TMD-effectors inserted into their destination membrane?
Until now, there is much known about the destination of effector proteins, their function and even about their localization within the host cell, and remarkably little of the way they have to take to get there.

6 Literature

Adams, H. *et al.* (2002) 'The presence of a helix breaker in the hydrophobic core of signal sequences of secretory proteins prevents recognition by the signal-recognition particle in Escherichia coli', *European Journal of Biochemistry*, 269(22), pp. 5564–5571. doi: 10.1046/j.1432-1033.2002.03262.x.

Agilent (2018) *QuikChange II Site-Directed Mutagenesis Kit*. Available at: http://www.chem.agilent.com/%0Alibrary/usermanuals/Public/200523.pdf.

Akeda, Y. and Galán, J. E. (2005) 'Chaperone release and unfolding of substrates in type III secretion', *Nature*, 437(7060), pp. 911–915. doi: 10.1038/nature03992.

Altman, E. and Segal, G. (2008) 'The response regulator CpxR directly regulates expression of several Legionella pneumophila icm/dot components as well as new translocated substrates', *Journal of Bacteriology*, 190(6), pp. 1985–1996. doi: 10.1128/JB.01493-07.

Alvarez-Martinez, C. E. and Christie, P. J. (2009) 'Biological Diversity of Prokaryotic Type IV Secretion Systems', *Microbiology and Molecular Biology Reviews*, 73(4), pp. 775–808. doi: 10.1128/mmbr.00023-09.

Amyot, W. M., DeJesus, D. and Isberg, R. R. (2013) 'Poison domains block transit of translocated substrates via the Legionella pneumophila Icm/Dot system', *Infection and Immunity*, 81(9), pp. 3239–3252. doi: 10.1128/IAI.00552-13.

Ariosa, A. *et al.* (2015) 'Regulation by a chaperone improves substrate selectivity during cotranslational protein targeting', *Proceedings of the National Academy of Sciences of the United States of America*, 112(25), pp. E3169–E3178. doi: 10.1073/pnas.1422594112.

Arnold, R., Jehl, A. and Rattei, T. (2010) 'Targeting effectors: the molecular recognition of Type III secreted proteins', *Microbes and Infection*. Elsevier Masson SAS, 12(5), pp. 346–358. doi: 10.1016/j.micinf.2010.02.003.

Aurass, P. *et al.* (2016) 'Life stage-specific proteomes of Legionella pneumophila reveal a highly differential abundance of virulence-associated Dot/Icm effectors', *Molecular and Cellular Proteomics*, 15(1), pp. 177–200. doi: 10.1074/mcp.M115.053579.

Ayers, K. L. *et al.* (2010) 'The Long-Range Activity of Hedgehog Is Regulated in the Apical Extracellular Space by the Glypican Dally and the Hydrolase Notum', *Developmental Cell*. Elsevier Ltd, 18(4), pp. 605–620. doi: 10.1016/j.devcel.2010.02.015.

Baram, D. *et al.* (2005) 'Structure of trigger factor binding domain in biologically homologous complex with eubacterial ribosome reveals its chaperone action', *Proceedings of the National Academy of Sciences of the United States of America*, 102(34), pp. 12017–12022. doi: 10.1073/pnas.0505581102.

Bardill, J. P., Miller, J. L. and Vogel, J. P. (2005) 'IcmS-dependent translocation of SdeA into macrophages by the Legionella pneumophila type IV secretion system', *Molecular Microbiology*, 56(1), pp. 90–103. doi: 10.1111/j.1365-2958.2005.04539.x.

De Barsy, M. *et al.* (2011) 'Identification of a Brucella spp. secreted effector specifically interacting with human small GTPase Rab2', *Cellular Microbiology*, 13(7), pp. 1044–1058. doi: 10.1111/j.1462-5822.2011.01601.x.

Beck, K. *et al.* (2000) 'Discrimination between SRP- and SecA/SecB-dependent substrates involves

selective recognition of nascent chains by SRP and trigger factor', *EMBO Journal*, 19(1), pp. 134–143. doi: 10.1093/emboj/19.1.134.

Belyi, Y. *et al.* (2008) 'Lgt: A family of cytotoxic glucosyltransferases produced by Legionella pneumophila', *Journal of Bacteriology*, 190(8), pp. 3026–3035. doi: 10.1128/JB.01798-07.

Bennett, J. C. Q. and Hughes, C. (2000) 'From flagellum assembly to virulence: The extended family of type III export chaperones', *Trends in Microbiology*, 8(5), pp. 202–204. doi: 10.1016/S0966-842X(00)01751-0.

Bennett, T. L. *et al.* (2013a) 'LegC3, an Effector Protein from Legionella pneumophila, Inhibits Homotypic Yeast Vacuole Fusion In Vivo and In Vitro', *PLoS ONE*, 8(2). doi: 10.1371/journal.pone.0056798.

Berger, K. H., Merriam, J. J. and Isberg, R. R. (1994) 'Altered intracellular targeting properties associated with mutations in the Legionella pneumophila dotA gene', *Molecular Microbiology*, 14(4), pp. 809–822. doi: 10.1111/j.1365-2958.1994.tb01317.x.

Bi, D. *et al.* (2013) 'SecReT4: A web-based bacterial type IV secretion system resource', *Nucleic Acids Research*, 41(D1), pp. 660–665. doi: 10.1093/nar/gks1248.

Bleves, S., Llosa, M. and Haven, N. (2021) 'functionally convergent', 22(5), pp. 1–17. doi: 10.1111/cmi.13157.Bacterial.

Bocian-Ostrzycka, K. M. *et al.* (2017) 'Bacterial thiol oxidoreductases — from basic research to new antibacterial strategies', *Applied Microbiology and Biotechnology*, 101(10), pp. 3977–3989. doi: 10.1007/s00253-017-8291-8.

Bornemann, T., Holtkamp, W. and Wintermeyer, W. (2014) 'Interplay between trigger factor and other protein biogenesis factors on the ribosome', *Nature Communications*. Nature Publishing Group, 5(May). doi: 10.1038/ncomms5180.

Brassinga, A. K. C. *et al.* (2003) 'A 65-kilobase pathogenicity island is unique to Philadelphia-1 strains of Legionella pneumophila', *Journal of Bacteriology*, 185(15), pp. 4630–4637. doi: 10.1128/JB.185.15.4630-4637.2003.

Burns, V. C. *et al.* (2003) 'Role of Bordetella O antigen in respiratory tract infection', *Infection and Immunity*, 71(1), pp. 86–94. doi: 10.1128/IAI.71.1.86-94.2003.

Burstein, D. *et al.* (2009) 'Genome-scale identification of Legionella pneumophila effectors using a machine learning approach', *PLoS Pathogens*, 5(7). doi: 10.1371/journal.ppat.1000508.

Burstein, D. *et al.* (2016) 'Genomic analysis of 38 Legionella species identifies large and diverse effector repertoires', *Nature Genetics*. Nature Publishing Group, 48(2), pp. 167–175. doi: 10.1038/ng.3481.

Büttner, D. (2012) 'Protein Export According to Schedule: Architecture, Assembly, and Regulation of Type III Secretion Systems from Plant- and Animal-Pathogenic Bacteria', *Microbiology and Molecular Biology Reviews*, 76(2), pp. 262–310. doi: 10.1128/mmbr.05017-11.

Cabezón, E., de la Cruz, F. and Arechaga, I. (2017) 'Conjugation inhibitors and their potential use to prevent dissemination of antibiotic resistance genes in bacteria', *Frontiers in Microbiology*, 8(NOV), pp. 1–7. doi: 10.3389/fmicb.2017.02329.

Cambronne, E. D. and Roy, C. R. (2007) 'The Legionella pneumophila IcmSW complex interacts with multiple Dot/Icm effectors to facilitate type IV translocation', *PLoS Pathogens*, 3(12), pp. 1837–1848.

doi: 10.1371/journal.ppat.0030188.

Cascales, E. and Christie, P. J. (2003) 'The versatile bacterial type IV secretion systems', *Nature Reviews Microbiology*, 1(2), pp. 137–149. doi: 10.1038/nrmicro753.

Cazalet, C. *et al.* (2004) 'Evidence in the Legionella pneumophila genome for exploitation of host cell functions and high genome plasticity', *Nature Genetics*, 36(11), pp. 1165–1173. doi: 10.1038/ng1447.

Chauhan, D. and Shames, S. R. (2021) 'Pathogenicity and Virulence of Legionella: Intracellular replication and host response', *Virulence*. Taylor & Francis, 12(1), pp. 1122–1144. doi: 10.1080/21505594.2021.1903199.

Chen, J. *et al.* (2007) 'Host cell-dependent secretion and translocation of the LepA and LepB effectors of Legionella pneumophila', *Cellular Microbiology*, 9(7), pp. 1660–1671. doi: 10.1111/j.1462-5822.2007.00899.x.

Chetrit, D. *et al.* (2018) 'A unique cytoplasmic ATPase complex defines the Legionella pneumophila type IV secretion channel', *Nature Microbiology*, 3(6), pp. 678–686. doi: 10.1038/s41564-018-0165-z.

Chong, A. *et al.* (2009) 'The purified and recombinant Legionella pneumophila chaperonin alters mitochondrial trafficking and microfilament organization', *Infection and Immunity*, 77(11), pp. 4724–4739. doi: 10.1128/IAI.00150-09.

Christie, P. J. (2016) 'Mosiac Type IV Secretion Systems', *EcoSal Plus*, 7(1), pp. 1–34. doi: 10.1128/ecosalplus.ESP-0020-2015.The.

Coburn, B., Sekirov, I. and Finlay, B. B. (2007) 'Type III secretion systems and disease', *Clinical Microbiology Reviews*, 20(4), pp. 535–549. doi: 10.1128/CMR.00013-07.

Coers, J. *et al.* (2000) 'Identification of Icm protein complexes that play distinct roles in the biogenesis of an organelle permissive for Legionella pneumophila intracellular growth', *Molecular Microbiology*, 38(4), pp. 719–736. doi: 10.1046/j.1365-2958.2000.02176.x.

Correia, A. M. *et al.* (2016) 'Probable Person-to-Person Transmission of Legionnaires' Disease', *New England Journal of Medicine*, 374(5), pp. 497–498. doi: 10.1056/nejmc1505356.

Costa, T. R. D. *et al.* (2015) 'Secretion systems in Gram-negative bacteria: structural and mechanistic insights', *Nature Reviews Microbiology*. Nature Publishing Group, 13(6), pp. 343–359. doi: 10.1038/nrmicro3456.

Couturier, M. R. *et al.* (2006) 'Interaction with CagF is required for translocation of CagA into the host via the Helicobacter pylori type IV secretion system', *Infection and Immunity*, 74(1), pp. 273–281. doi: 10.1128/IAI.74.1.273-281.2006.

Craig, E. A., Eisenman, H. C. and Hundley, H. A. (2003) 'Ribosome-tethered molecular chaperones: The first line of defense against protein misfolding?', *Current Opinion in Microbiology*, 6(2), pp. 157–162. doi: 10.1016/S1369-5274(03)00030-4.

Cunha, B. A., Burillo, A. and Bouza, E. (2016) 'Legionnaires' disease', *The Lancet*. Elsevier Ltd, 387(10016), pp. 376–385. doi: 10.1016/S0140-6736(15)60078-2.

Dalebroux, Z. D. *et al.* (2010) 'ppGpp Conjures Bacterial Virulence', *Microbiology and Molecular Biology Reviews*, 74(2), pp. 171–199. doi: 10.1128/mmbr.00046-09.

Dalebroux, Z. D., Edwards, R. L. and Swanson, M. S. (2009) 'SpoT governs Legionella pneumophila differentiation in host macrophages', *Molecular Microbiology*, 71(3), pp. 640–658. doi:

10.1111/j.1365-2958.2008.06555.x.

Dolezal, P. *et al.* (2012) 'Legionella pneumophila secretes a mitochondrial carrier protein during infection', *PLoS Pathogens*, 8(1). doi: 10.1371/journal.ppat.1002459.

Driessen, A. J. M. and Nouwen, N. (2008) 'Protein translocation across the bacterial cytoplasmic membrane', *Annual Review of Biochemistry*, 77, pp. 643–667. doi: 10.1146/annurev.biochem.77.061606.160747.

Dym, O. *et al.* (2008) 'Crystal structure of the Agrobacterium virulence complex VirE1-VirE2 reveals a flexible protein that can accommodate different partners', *Proceedings of the National Academy of Sciences of the United States of America*, 105(32), pp. 11170–11175. doi: 10.1073/pnas.0801525105.

Edwards, M. T., Fry, N. K. and Harrison, T. G. (2008) 'Clonal population structure of Legionella pneumophila inferred from allelic profiling', *Microbiology*, 154(3), pp. 852–864. doi: 10.1099/mic.0.2007/012336-0.

Elad, N. *et al.* (2007) 'Topologies of a Substrate Protein Bound to the Chaperonin GroEL', *Molecular Cell*. Elsevier Inc., 26(3), pp. 415–426. doi: 10.1016/j.molcel.2007.04.004.

Elofsson, A. and Von Heijne, G. (2007) 'Membrane protein structure: Prediction versus reality', *Annual Review of Biochemistry*, 76, pp. 125–140. doi: 10.1146/annurev.biochem.76.052705.163539.

Ensminger, A. W. and Isberg, R. R. (2009) 'Legionella pneumophila Dot/Icm translocated substrates: a sum of parts', *Current Opinion in Microbiology*, 12(1), pp. 67–73. doi: 10.1016/j.mib.2008.12.004.

Ensminger, A. W. and Isberg, R. R. (2010) 'E3 ubiquitin ligase activity and targeting of BAT3 by multiple Legionella pneumophila translocated substrates', *Infection and Immunity*, 78(9), pp. 3905–3919. doi: 10.1128/IAI.00344-10.

Farell, I. S. *et al.* (2005) 'Photocross- linking interacting proteins with a genetically encoded benzophenone.', *Nature methods*, 2, pp. 377–384.

Faucher, S. P., Mueller, C. A. and Shuman, H. A. (2011) 'Legionella pneumophila transcriptome during intracellular multiplication in human macrophages', *Frontiers in Microbiology*, 2(APR), pp. 1–18. doi: 10.3389/fmicb.2011.00060.

De Felipe, K. S. *et al.* (2005) 'Evidence for acquisition of Legionella type IV secretion substrates via interdomain horizontal gene transfer', *Journal of Bacteriology*, 187(22), pp. 7716–7726. doi: 10.1128/JB.187.22.7716-7726.2005.

De Felipe, K. S., Glover, R. T., Charpentier, X., Anderson, O. R., Reyes, M., Pericone, C. D. and Shuman, H. A. (2008) 'Legionella eukaryotic-like type IV substrates interfere with organelle trafficking', *PLoS Pathogens*, 4(8). doi: 10.1371/journal.ppat.1000117.

De Felipe, K. S., Glover, R. T., Charpentier, X., Anderson, O. R., Reyes, M., Pericone, C. D. and Shuman, H. a. (2008) 'Legionella eukaryotic-like type IV substrates interfere with organelle trafficking', *PLoS Pathogens*, 4(8). doi: 10.1371/journal.ppat.1000117.

Flannagan, R. S., Cosío, G. and Grinstein, S. (2009) 'Antimicrobial mechanisms of phagocytes and bacterial evasion strategies', *Nature Reviews Microbiology*, 7(5), pp. 355–366. doi: 10.1038/nrmicro2128.

Franco, I. S., Shohdy, N. and Shuman, H. A. (2012) 'The Legionella pneumophila effector VipA is an actin nucleator that alters host cell organelle trafficking', *PLoS Pathogens*, 8(2). doi: 10.1371/journal.ppat.1002546.

Franco, I. S., Shuman, H. A. and Charpentier, X. (2009) 'The perplexing functions and surprising origins of Legionella pneumophila type IV secretion effectors', *Cellular Microbiology*, 11(10), pp. 1435–1443. doi: 10.1111/j.1462-5822.2009.01351.x.

Fraser, D. *et al.* (1977) 'Legionnaires' disease: description of an epidemic of pneumonia.', *N Engl J Med*, 297, pp. 1189–1197.

Galan, J. E. and Waksman, G. (2018) 'Protein Injection Machines in Bacteria', *Cell*, 172(6), pp. 1306–1318. doi: 10.1016/j.cell.2018.01.034.

Ghosal, D. *et al.* (2017) ' In situ structure of the Legionella Dot/Icm type IV secretion system by electron cryotomography ', *EMBO reports*, 18(5), pp. 726–732. doi: 10.15252/embr.201643598.

Ghosal, D. *et al.* (2018) 'Molecular architecture of the Legionella Dot/Icm type IV secretion system', *bioRxiv*. doi: 10.1101/312009.

Ghosal, D. *et al.* (2019) 'Molecular architecture, polar targeting and biogenesis of the Legionella Dot/Icm T4SS', *Nature Microbiology*. Springer US, 4(7), pp. 1173–1182. doi: 10.1038/s41564-019-0427-4.

Gibson, D. G. *et al.* (2010) 'Creation of a bacterial cell controlled by a chemically synthesized genome.', *Science (New York, N.Y.)*, 329(5987), pp. 52–56.

Gomez-Valero, L. *et al.* (2019) 'More than 18,000 effectors in the Legionella genus genome provide multiple, independent combinations for replication in human cells', *Proceedings of the National Academy of Sciences of the United States of America*, 116(6), pp. 2265–2273. doi: 10.1073/pnas.1808016116.

Gomis-Rüth, F. X. *et al.* (2001) 'The bacterial conjugation protein TrwB resembles ring helicases and F1-atpase', *Nature*, 409(6820), pp. 637–641. doi: 10.1038/35054586.

Gonzalez-Rivera, C., Bhatty, M. and Christie, P. J. (2016) 'Mechanism and function of type IV secretion during infection of the human host', *Virulence Mechanisms of Bacterial Pathogens*, (i), pp. 265–303. doi: 10.1128/9781555819286.ch10.

Green, E. R. and Mecsas, J. (2016) 'Bacterial Secretion Systems: An Overview', *Microbiology Spectrum*, 4(1), pp. 1–32. doi: 10.1128/microbiolspec.vmbf-0012-2015.

Grohmann, E. *et al.* (2018) 'Type IV secretion in Gram-negative and Gram-positive bacteria', *Mol Microbiol.*, 107(4), pp. 455–471. doi: 10.1111/mmi.13896.

Hammer, B. K., Tateda, E. S. and Swanson, M. S. (2002) 'A two-component regulator induces the transmission phenotype of stationary-phase Legionella pneumophila', *Molecular Microbiology*, 44(1), pp. 107–118. doi: 10.1046/j.1365-2958.2002.02884.x.

Hartl, F. U. and Hayer-Hartl, M. (2002) 'Protein folding. Molecular chaperones in the cytosol: From nascent chain to folded protein', *Science*, 295(5561), pp. 1852–1858. doi: 10.1126/science.1068408.

Hatakeyama, M. (2017) 'Structure and function of helicobacter pylori caga, the first-identified bacterial protein involved in human cancer', *Proceedings of the Japan Academy Series B: Physical and Biological Sciences*, 93(4), pp. 196–219. doi: 10.2183/pjab.93.013.

von Heijne, G. (1985) 'Signal sequences. The limits of variation', *Journal of Molecular Biology*, 184(1), pp. 99–105. doi: 10.1016/0022-2836(85)90046-4.

von Heljne, G. (1989) 'Control of topology and mode of assembly of a polytopic membrane protein by

positively charged residues', *Nature*, 341(5 October), pp. 456–458.

Hessa, T., Meindl-Beinker, Nadja M, *et al.* (2007) 'Molecular code for transmembrane-helix recognition by the Sec61 translocon.', *Nature*, 450(7172), pp. 1026–1030. doi: 10.1038/nature06387.

Hessa, T., Meindl-Beinker, Nadja M., *et al.* (2007) 'Molecular code for transmembrane-helix recognition by the Sec61 translocon', *Nature*, 450(7172), pp. 1026–1030. doi: 10.1038/nature06387.

Hoffmann, A. *et al.* (2012) 'Concerted Action of the Ribosome and the Associated Chaperone Trigger Factor Confines Nascent Polypeptide Folding', *Molecular Cell*, 48(1), pp. 63–74. doi: 10.1016/j.molcel.2012.07.018.

Hohlfeld, S. *et al.* (2006) 'A C-terminal translocation signal is necessary, but not sufficient for type IV secretion of the Helicobacter pylori CagA protein', *Molecular Microbiology*, 59(5), pp. 1624–1637. doi: 10.1111/j.1365-2958.2006.05050.x.

Holtkamp, W. *et al.* (2012) 'Dynamic switch of the signal recognition particle from scanning to targeting', *Nature Structural and Molecular Biology*. Nature Publishing Group, 19(12), pp. 1332–1337. doi: 10.1038/nsmb.2421.

Horwitz, M. A. *et al.* (1983) 'Formation of a novel phagosome b the Legionnaires' disease bacteium (Legionella pneumophila) in human monoctes', Journal of Experimental Medicine, 158 (4), pp. 1319-1331.doi:10.1084/jem.158.4.1319

Hsu, F. S. *et al.* (2012) 'Structural basis for substrate recognition by a unique Legionella phosphoinositide phosphatase', *Proceedings of the National Academy of Sciences of the United States of America*, 109(34), pp. 13567–13572. doi: 10.1073/pnas.1207903109.

Huang, L. *et al.* (2011) 'The E Block motif is associated with Legionella pneumophila translocated substrates', *Cellular Microbiology*, 13(2), pp. 227–245. doi: 10.1111/j.1462-5822.2010.01531.x.

Hubber, A. and Roy, C. R. (2010) 'Modulation of host cell function by legionella pneumophila type IV effectors', *Annual Review of Cell and Developmental Biology*, 26, pp. 261–283. doi: 10.1146/annurev-cellbio-100109-104034.

Huddleston, J. R. (2014) 'Horizontal gene transfer in the human gastrointestinal tract: Potential spread of antibiotic resistance genes', *Infection and Drug Resistance*, 7, pp. 167–176. doi: 10.2147/IDR.S48820.

Ingmundson, A. *et al.* (2007) 'Legionella pneumophila proteins that regulate Rab1 membrane cycling', *Nature*, 450(7168), pp. 365–369. doi: 10.1038/nature06336.

Isaac, D. T. *et al.* (2015) 'MavN is a Legionella pneumophila vacuole-associated protein required for efficient iron acquisition during intracellular growth', *Proceedings of the National Academy of Sciences of the United States of America*, 112(37), pp. E5208–E5217. doi: 10.1073/pnas.1511389112.

Ito, K. and Akiyama, Y. (2005) 'Cellular functions, mechanism of action, and regulation of FtsH protease', *Annual Review of Microbiology*, 59, pp. 211–231. doi: 10.1146/annurev.micro.59.030804.121316.

Jeong, K. C. C. *et al.* (2018) 'Polar targeting and assembly of the Legionella Dot/Icm type IV secretion system (T4SS) by T6SS-related components', *bioRxiv*, 1. doi: 10.1101/315721.

Jeong, K. C., Sexton, J. A. and Vogel, J. P. (2015) 'Spatiotemporal Regulation of a Legionella pneumophila T4SS Substrate by the Metaeffector SidJ', *PLoS Pathogens*, 11(3), pp. 1–22. doi: 10.1371/journal.ppat.1004695.

Jeong, K. C., Sutherland, M. C. and Vogel, J. P. (2015) 'Novel export control of a LegionellaDot/Icm substrate is mediated by dual, independent signal sequences', *Molecular Microbiology*, 96(1), pp. 175–188. doi: 10.1111/mmi.12928.

de Jong, M. F. *et al.* (2008) 'Identification of VceA and VceC, two members of the VjbR regulon that are translocated into macrophages by the Brucella Type IV secretion system', *Mol Microbiol.*, 70(6), pp. 1378–1396. doi: 10.1111/j.1365-2958.2008.06487.x.

Kamen, L. A., Levinsohn, J. and Swanson, J. A. (2007) 'Diacylglycerol Is Required for the Formation of COPI Vesicles in the Golgi-to-ER Transport Pathway', *Molecular Biology of the Cell*, 18(July), pp. 2463–2472. doi: 10.1091/mbc.e07-01-0061.

Kessler, A. *et al.* (2013) 'The Legionella pneumophila orphan sensor kinase LqsT regulates competence and pathogen-host interactions as a component of the LAI-1 circuit', *Environmental Microbiology*, 15(2), pp. 646–662. doi: 10.1111/j.1462-2920.2012.02889.x.

Kihara, A., Akiyama, Y. and Koreaki, I. (1999) 'Dislocation of membrane proteins in FtsH-mediated proteolysis', *EMBO Journal*, 18(11), pp. 2970–2981. doi: 10.1093/emboj/18.11.2970.

Kol, S., Nouwen, N. and Driessen, A. J. M. (2008) 'Mechanisms of YidC-mediated insertion and assembly of multimeric membrane protein complexes', *Journal of Biological Chemistry*, 283(46), pp. 31269–31273. doi: 10.1074/jbc.R800029200.

Koraimann, G. (2018) 'Spread and Persistence of Virulence and Antibiotic Resistance Genes: A Ride on the F Plasmid Conjugation Module', *EcoSal Plus*, 8(1). doi: 10.1128/ecosalplus.esp-0003-2018.

Kramer, G. *et al.* (2002) 'L23 protein functions as a chaperone docking site on the ribosome', *Nature*, 419(6903), pp. 171–174. doi: 10.1038/nature01047.

Krampen, L. *et al.* (2018) 'Revealing the mechanisms of membrane protein export by virulence-associated bacterial secretion systems', *Nature Communications*. Springer US, 9(1). doi: 10.1038/s41467-018-05969-w.

Kubori, T. *et al.* (2014) 'Native structure of a type IV secretion system core complex essential for Legionella pathogenesis', *Proceedings of the National Academy of Sciences of the United States of America*, 111(32), pp. 11804–11809. doi: 10.1073/pnas.1404506111.

Kubori, T., Hyakutake, A. and Nagai, H. (2008) 'Legionella translocates an E3 ubiquitin ligase that has multiple U-boxes with distinct functions', *Molecular Microbiology*, 67(6), pp. 1307–1319. doi: 10.1111/j.1365-2958.2008.06124.x.

Kuroda, T. *et al.* (2015) 'Molecular and structural analysis of Legionella DotI gives insights into an inner membrane complex essential for type IV secretion', *Scientific Reports*. Nature Publishing Group, 5(February), pp. 1–14. doi: 10.1038/srep10912.

Kwak, M. J. *et al.* (2017) 'Architecture of the type IV coupling protein complex of Legionella pneumophila', *Nature Microbiology*, 2(July). doi: 10.1038/nmicrobiol.2017.114.

Lara-Tejero, M. *et al.* (2011) 'A sorting platform determines the order of protein secretion in bacterial type III systems.', *Science (New York, N.Y.)*, 331(6021), pp. 1188–1191. doi: 10.1126/science.1201476.

Larson, C. L., Beare, P. A. and Heinzen, R. A. (2019) 'Dependency of coxiella burnetii type 4B secretion on the chaperone IcmS', *Journal of Bacteriology*, 201(23), pp. 1–14. doi: 10.1128/JB.00431-19.

Lee, H. C. and Bernstein, H. D. (2001) 'The targeting pathway of Escherichia coli presecretory and integral membrane proteins is specified by the hydrophobicity of the targeting signal', *Proceedings of*

the National Academy of Sciences of the United States of America, 98(6), pp. 3471–3476. doi: 10.1073/pnas.051484198.

Liberek, K., Lewandowska, A. and Ziętkiewicz, S. (2008) 'Chaperones in control of protein disaggregation', *EMBO Journal*, 27(2), pp. 328–335. doi: 10.1038/sj.emboj.7601970.

Lifshitz, Z. *et al.* (2013) '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*, 110(8). doi: 10.1073/pnas.1215278110.

Liu, W. *et al.* (2007) 'Structural basis for the recognition of para-benzoyl-L-phenylalanine by evolved aminoacyl-tRNA synthetases', *Angew. Chem. Int. Ed. Engl*, 46, pp. 6073–6075.

Locht, C., Coutte, L. and Mielcarek, N. (2011) 'The ins and outs of pertussis toxin', *FEBS Journal*, 278(23), pp. 4668–4682. doi: 10.1111/j.1742-4658.2011.08237.x.

Luirink, J. *et al.* (1994) 'An alternative protein targeting pathway in Escherichia coli: Studies on the role of FtsY', *EMBO Journal*, 13(10), pp. 2289–2296. doi: 10.1002/j.1460-2075.1994.tb06511.x.

Luirink, J. and Sinning, I. (2004) 'SRP-mediated protein targeting: Structure and function revisited', *Biochimica et Biophysica Acta - Molecular Cell Research*, 1694(1-3 SPEC.ISS.), pp. 17–35. doi: 10.1016/j.bbamcr.2004.03.013.

Luo, Z. Q. (2012) 'Legionella secreted effectors and innate immune responses', *Cellular Microbiology*, 14(1), pp. 19–27. doi: 10.1111/j.1462-5822.2011.01713.x.

Luo, Z. Q. and Isberg, R. R. (2004) '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*, 101(3), pp. 841–846. doi: 10.1073/pnas.0304916101.

Lycklama a Nijeholt, J. A. and Driessen, A. J. M. (2012) 'The bacterial Sec-translocase: Structure and mechanism', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1592), pp. 1016–1028. doi: 10.1098/rstb.2011.0201.

Machner, M. P. and Isberg, R. R. (2006) 'Targeting of Host Rab GTPase Function by the Intravacuolar Pathogen Legionella pneumophila', *Developmental Cell*, 11(1), pp. 47–56. doi: 10.1016/j.devcel.2006.05.013.

Marchesini, M. I. *et al.* (2011) 'In search of Brucella abortus type iv secretion substrates: Screening and identification of four proteins translocated into host cells through virb system', *Cellular Microbiology*, 13(8), pp. 1261–1274. doi: 10.1111/j.1462-5822.2011.01618.x.

De Matteis, M. A. and Godi, A. (2004) 'PI-loting membrane traffic', *Nature Cell Biology*, 6(6), pp. 487–492. doi: 10.1038/ncb0604-487.

Mayer, M. P. and Bukau, B. (2005) 'Hsp70 chaperones: Cellular functions and molecular mechanism', *Cellular and Molecular Life Sciences*, 62(6), pp. 670–684. doi: 10.1007/s00018-004-4464-6.

McDade, J. *et al.* (1977) 'Legionnaires' disease: isolation of a bacterium and demonstration of its role in other respiratory disease.', *N Engl J Med*, 297, pp. 1197–1203.

Meir, A. *et al.* (2018) 'Legionella DotM structure reveals a role in effector recruiting to the Type 4B secretion system', *Nature Communications*, 9(1), pp. 1–12. doi: 10.1038/s41467-017-02578-x.

Meir, A. *et al.* (2020) 'Mechanism of effector capture and delivery by the type IV secretion system from Legionella pneumophila', *Nature Communications*. Springer US, 11(1), pp. 1–11. doi:

10.1038/s41467-020-16681-z.

Nagai, H. *et al.* (2002) 'A bacterial guanine nucleotide exchange factor activates ARF on Legionella phagosomes', *Science*, 295(5555), pp. 679–682. doi: 10.1126/science.1067025.

Nagai, H. *et al.* (2005) '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 of the United States of America*, 102(3), pp. 826–831. doi: 10.1073/pnas.0406239101.

Nagai, H. and Kubori, T. (2011) 'Type IVB secretion systems of Legionella and other gram-negative bacteria', *Frontiers in Microbiology*, 2(JUNE), pp. 1–12. doi: 10.3389/fmicb.2011.00136.

Neunuebel, M. R. *et al.* (2011) 'De-AMPylation of hte small GTPase Rab1 by the pahtogen *Legionella pneumophila*.', *Science*, 333(6041), pp. 453–6. doi: 10.1126/science.1207193.De-AMPylation.

Ninio, S. *et al.* (2005) 'The Legionella IcmS-IcmW protein complex is important for Dot/Icm-mediated protein translocation', *Molecular Microbiology*, 55(3), pp. 912–926. doi: 10.1111/j.1365-2958.2004.04435.x.

Ninio, S. and Roy, C. R. (2007) 'Effector proteins translocated by Legionella pneumophila: strength in numbers', *Trends in Microbiology*, 15(8), pp. 372–380. doi: 10.1016/j.tim.2007.06.006.

Notti, R. Q. and Stebbins, C. E. (2016) 'The structure and function of type iii secretion systems', *Virulence Mechanisms of Bacterial Pathogens*, pp. 241–264. doi: 10.1128/9781555819286.ch9.

O'Connor, T. J. *et al.* (2011) 'Erratum: Minimization of the Legionella pneumophila genome reveals chromosomal regions involved in host range expansion', *Proceedings of the National Academy of Sciences of the United States of America*, 108(43), p. 17856. doi: 10.1073/pnas.1115233108.

Odorizzi, G., Babst, M. and Emr, S. D. (2000) 'Phosphoinositide signaling and the regulation of membrane trafficking in yeast', *Trends in Biochemical Sciences*, 25(5), pp. 229–235. doi: 10.1016/S0968-0004(00)01543-7.

Oliver, D. and Paetzel, M. (2007) 'Crystal structure of the periplasmic domain of the Escherichia Coli YIDC', *J. Biol. Chem.*, 283, pp. 5208–5216.

Paetzel, M. *et al.* (2002) 'Signal peptidases', *Chemical Reviews*, 102(12), pp. 4549–4579. doi: 10.1021/cr010166y.

Paetzel, M., Dalbey, R. E. and Strynadka, N. C. J. (1998) 'Crystal structure of a bacterial signal peptidase in complex with a β-lactam inhibitor', *Nature*, 396, pp. 186–190.

Di Paolo, G. and De Camilli, P. (2006) 'Phosphoinositides in cell regulation and membrane dynamics', *Nature*, 443(7112), pp. 651–657. doi: 10.1038/nature05185.

Papaloukas, C. *et al.* (2008) 'Estimating the length of transmembrane helices using Z-coordinate predictions', *Protein Science*, 17(2), pp. 271–278. doi: 10.1110/ps.073036108.

Papanikou, E., Karamanou, S. and Economou, A. (2007) 'Bacterial protein secretion through the translocase nanomachine', *Nature Reviews Microbiology*, 5(11), pp. 839–851. doi: 10.1038/nrmicro1771.

Park, D. *et al.* (2020) 'Analysis of DOT/ICM type IVB secretion system subassemblies by cryoelectron tomography reveals conformational changes induced by DotB binding', *mBio*, 11(1), pp. 1–11. doi: 10.1128/mBio.03328-19.

Parsot, C., Hamiaux, C. and Page, A. L. (2003) 'The various and varying roles of specific chaperones in

type III secretion systems', *Current Opinion in Microbiology*, 6(1), pp. 7–14. doi: 10.1016/S1369-5274(02)00002-4.

Pattis, I. *et al.* (2007) 'The Helicobacter pylori CagF protein is a type IV secretion chaperone-like molecule that binds close to the C-terminal secretion signal of the CagA effector protein', *Microbiology*, 153(9), pp. 2896–2909. doi: 10.1099/mic.0.2007/007385-0.

Payne, N. R. and Horwitz, M. A. (1987) 'Phagocytosis of Legionetta pneumophiia is mediated by human monocyte compiement receptors.', *J Exp Med*, 166, pp. 1377–1389.

Pena, R. T. *et al.* (2019) 'Relationship between quorum sensing and secretion systems', *Frontiers in Microbiology*, 10(JUN). doi: 10.3389/fmicb.2019.01100.

Purcell, M. and Shuman, H. A. (1998) 'The Legionella pneumophila icmGCDJBF genes are required for killing of human macrophages', *Infection and Immunity*, 66(5), pp. 2245–2255. doi: 10.1128/iai.66.5.2245-2255.1998.

Rasis, M. and Segal, G. (2009) 'The LetA-RsmYZ-CsrA regulatory cascade, together with RpoS and PmrA, post-transcriptionally regulates stationary phase activation of Legionella pneumophila Icm/Dot effectors', *Molecular Microbiology*, 72(4), pp. 995–1010. doi: 10.1111/j.1365-2958.2009.06705.x.

Sahr, T. *et al.* (2009) 'Two small ncRNAs jointly govern virulence and transmission in Legionella pneumophila', *Molecular Microbiology*, 72(3), pp. 741–762. doi: 10.1111/j.1365-2958.2009.06677.x.

Sahr, T. *et al.* (2012) 'Deep sequencing defines the transcriptional map of L. pneumophila and identifies growth phase-dependent regulated ncRNAs implicated in virulence', *RNA Biology*, 9(4), pp. 503–519. doi: 10.4161/rna.20270.

Sahr, T. *et al.* (2017) 'The Legionella pneumophila genome evolved to accommodate multiple regulatory mechanisms controlled by the CsrA-system', PLoS Genetics,13(2, pp.1-38. doi: 10.1371/journal.pgen.1006629.

Saibil, H. R. (2008) 'Chaperone machines in action', *Current Opinion in Structural Biology*, 18(1), pp. 35–42. doi: 10.1016/j.sbi.2007.11.006.

Sandikci, A. *et al.* (2013) 'Dynamic enzyme docking to the ribosome coordinates N-terminal processing with polypeptide folding', *Nature Structural and Molecular Biology*. Nature Publishing Group, 20(7), pp. 843–850. doi: 10.1038/nsmb.2615.

Schibich, D. *et al.* (2016) 'Global profiling of SRP interaction with nascent polypeptides', *Nature*. Nature Publishing Group, 536(7615), pp. 219–223. doi: 10.1038/nature19070.

Schoebel, S. *et al.* (2009) 'RabGDI Displacement by DrrA from Legionella Is a Consequence of Its Guanine Nucleotide Exchange Activity', *Molecular Cell*. Elsevier Ltd, 36(6), pp. 1060–1072. doi: 10.1016/j.molcel.2009.11.014.

Schulein, R. *et al.* (2005) 'A bipartite signal mediates the transfer of type IV secretion substrates of Bartonella henselae into human cells', *Proceedings of the National Academy of Sciences of the United States of America*, 102(3), pp. 856–861. doi: 10.1073/pnas.0406796102.

Segal, G., Purcell, M. and Shuman, H. A. (1998) 'Host cell killing and bacterial conjugation require overlapping sets of genes within a 22-kb region of the Legionella pneumophila genome', *Proceedings of the National Academy of Sciences of the United States of America*, 95(4), pp. 1669–1674. doi: 10.1073/pnas.95.4.1669.

Segal, G. and Shuman, H. A. (1997) 'Characterization of a new region required for macrophage killing

by Legionella pneumophila', *Infection and Immunity*, 65(12), pp. 5057–5066. doi: 10.1128/iai.65.12.5057-5066.1997.

Segal, G. and Shuman, H. A. (1999) 'Legionella pneumophila utilizes the same genes to multiply within Acanthamoeba castellanii and human macrophages', *Infection and Immunity*, 67(5), pp. 2117–2124. doi: 10.1128/iai.67.5.2117-2124.1999.

Shi, X. *et al.* (2016) 'Direct targeting of membrane fusion by SNARE mimicry: Convergent evolution of Legionella effectors', *Proceedings of the National Academy of Sciences of the United States of America*, 113(31), pp. 8807–8812. doi: 10.1073/pnas.1608755113.

Singh, N. and Wagner, S. (2019) 'Investigating the assembly of the bacterial type III secretion system injectisome by in vivo photocrosslinking', *International Journal of Medical Microbiology*, 309(6).

Sutherland, M. C. *et al.* (2013) 'Reassessing the Role of DotF in the Legionella pneumophila Type IV Secretion System', *PLoS ONE*, 8(6). doi: 10.1371/journal.pone.0065529.

Tiaden, A. *et al.* (2007) 'The Legionella pneumophila response regulator LqsR promotes host cell interactions as an element of the virulence regulatory network controlled by RpoS and LetA', *Cellular Microbiology*, 9(12), pp. 2903–2920. doi: 10.1111/j.1462-5822.2007.01005.x.

Tiaden, A. *et al.* (2008) 'Synergistic contribution of the Legionella pneumophila lqs genes to pathogen-host interactions', *Journal of Bacteriology*, 190(22), pp. 7532–7547. doi: 10.1128/JB.01002-08.

Tilney, L. G. *et al.* (2001) 'How the parasitic bacterium Legionella pneumophila modifies its phagosome and transforms it into rough ER: Implications for conversion of plasma membrane to the ER membrane', *Journal of Cell Science*, 114(24), pp. 4637–4650. doi: 10.1242/jcs.114.24.4637.

Toulabi, L. *et al.* (2013) 'Identification and structural characterization of a Legionella phosphoinositide phosphatase', *Journal of Biological Chemistry*. Â© 2013 ASBMB. Currently published by Elsevier Inc; originally published by American Society for Biochemistry and Molecular Biology., 288(34), pp. 24518–24527. doi: 10.1074/jbc.M113.474239.

Tsirigos, K. D. *et al.* (2015) 'The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides', *Nucleic Acids Research*, 43(W1), pp. W401–W407. doi: 10.1093/nar/gkv485.

Valent, Q. A. *et al.* (1995) 'Early events in preprotein recognition in E. coli: interaction of SRP and trigger factor with nascent polypeptides', *EMBO Journal*, 14(22), pp. 5494–5505. doi: 10.1002/j.1460-2075.1995.tb00236.x.

Vergne, I., Chua, J. and Deretic, V. (2003) 'Tuberculosis toxin blocking phagosome maturation inhibits a novel Ca 2+/calmodulin-PI3K hVPS34 cascade', *Journal of Experimental Medicine*, 198(4), pp. 653–659. doi: 10.1084/jem.20030527.

Vergunst, A. C. *et al.* (2000) 'VirB/D4-dependent protein translocation from Agrobacterium into plant cells', *Science*, 290(5493), pp. 979–982. doi: 10.1126/science.290.5493.979.

Vergunst, A. C. *et al.* (2005) 'Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of Agrobacterium', *Proceedings of the National Academy of Sciences of the United States of America*, 102(3), pp. 832–837. doi: 10.1073/pnas.0406241102.

Vieira, O. V. *et al.* (2001) 'Distinct roles of class I and class III phosphatidylinositol 3-kinases in phagosome formation and maturation', *Journal of Cell Biology*, 155(1), pp. 19–25. doi: 10.1083/jcb.200107069.

Vincent, C. D. *et al.* (2006) 'Identification of the core transmembrane complex of the Legionella Dot/Icm type IV secretion system', *Molecular Microbiology*, 62(5), pp. 1278–1291. doi: 10.1111/j.1365-2958.2006.05446.x.

Vincent, C. D. *et al.* (2012) 'Identification of the DotL coupling protein subcomplex of the Legionella Dot/Icm type IV secretion system', *Molecular Microbiology*, 85(2), pp. 378–391. doi: 10.1111/j.1365-2958.2012.08118.x.

Vincent, C. D. and Vogel, J. P. (2006) 'The Legionella pneumophila IcmS-LvgA protein complex is important for Dot/Icm-dependent intracellular growth', *Molecular Microbiology*, 61(3), pp. 596–613. doi: 10.1111/j.1365-2958.2006.05243.x.

Vogel, J. P. *et al.* (1998) 'Conjugative transfer by the virulence system of Legionella pneumophila', *Science*, 279(5352), pp. 873–876. doi: 10.1126/science.279.5352.873.

Wallden, K., Rivera-Calzada, A. and Waksman, G. (2010) 'Type IV secretion systems: Versatility and diversity in function', *Cellular Microbiology*, 12(9), pp. 1203–1212. doi: 10.1111/j.1462-5822.2010.01499.x.

Wallin, E. and von Heijne, G. (1998) 'Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms', *Protein Science*, 7(4), pp. 1029–1038. doi: 10.1002/pro.5560070420.

Weber, G. F. and Menko, A. S. (2006) 'Phosphatidylinositol 3-kinase is necessary for lens fiber cell differentiation and survival', *Investigative Ophthalmology and Visual Science*, 47(10), pp. 4490–4499. doi: 10.1167/iovs.06-0401.

Westerhausen, S. *et al.* (2020) 'A NanoLuc luciferase-based assay enabling the real-time analysis of protein secretion and injection by bacterial type III secretion systems', *Molecular Microbiology*, 113(6), pp. 1240–1254. doi: 10.1111/mmi.14490.

White, S. H. and Heijne, G. Von (2004) 'The machinery of membrane protein assembly', *Current Opinion in Structural Biology*, 14(4), pp. 397–404. doi: 10.1016/j.sbi.2004.07.003.

Xu, J. *et al.* (2017) '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*, 114(51), pp. 13543–13548. doi: 10.1073/pnas.1706883115.

Yamamoto, M. *et al.* (2019) 'Cell cell and virus cell fusion assay based analyses of alanine insertion mutants in the distal 9 portion of the JRFL gp41 subunit from HIV-1', *Journal of Biological Chemistry*, 294(14), pp. 5677–5687. doi: 10.1074/jbc.RA118.004579.

Yao, D., Cherney, M. and Cygler, M. (2014), *Acta Cryst*, 70(D), pp. 436–441. doi: 10.1107/S139900471302991X.

Yoshida, Y. *et al.* (2014) 'Functional characterization of the type III secretion ATPase SsaN encoded by Salmonella pathogenicity island 2', *PLoS ONE*, 9(4). doi: 10.1371/journal.pone.0094347.

Young, J. C. *et al.* (2004) 'Pathways of chaperone-mediated protein folding in the cytosol', *Nature Reviews Molecular Cell Biology*, 5(10), pp. 781–791. doi: 10.1038/nrm1492.

Zechner, E. L., Lang, S. and Schildbach, J. F. (2012) 'Assembly and mechanisms of bacterial type IV secretion machines', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1592), pp. 1073–1087. doi: 10.1098/rstb.2011.0207.

Zhao, Z. *et al.* (2001) 'Activities of virE1 and the VirE1 secretion chaperone in export of the

multifunctional VirE2 effector via an Agrobacterium type IV secretion pathway', *Journal of Bacteriology*, 183(13), pp. 3855–3865. doi: 10.1128/JB.183.13.3855-3865.2001.

Zuckman, D. M., Hung, J. B. and Roy, C. R. (1999) 'Pore-forming activity is not sufficient for Legionella pneumophila phagosome trafficking and intracellular growth', *Molecular Microbiology*, 32(5), pp. 990–1001. doi: 10.1046/j.1365-2958.1999.01410.x.

Zusman, T. *et al.* (2007) 'The response regulator PmrA is a major regulator of the icm/dot type IV secretion system in Legionella pneumophila and Coxiella burnetii', *Molecular Microbiology*, 63(5), pp. 1508–1523. doi: 10.1111/j.1365-2958.2007.05604.x.

7 Appendix

7.1 List of TMD-containing T3SS substrates with ΔG_{app} predictions (window 10-35; length correction ON) and their amino acid sequence

Protein name	Protein length/ aa	ΔGpred (window 18-35 length corr. ON) # TMS	ΔGpred (most hydro-phobic)	number (most hydro-phobic)	length (most hydro-phobic)	sequence (most hydrophobic)
CegC1	525	1	1.283	1	22	FTPWSIRVYILGHAIALRYARL
Lpg0041	748	1	-1.936	1	29	MILLNLMGFTMIRTFLTAFLMLLCSLALA
Ceg2	368	3	-2.093	2	34	LSRVMSAGILIAAAAGAIPLIYLGTLFPAMWVYL
Lem1	1324	1	1.471	1	19	VPFGLPASLSVLTIFLTYA
Ceg4	397	2	-1.236	1	32	FRTLKIMGAFAGIGFGFGFALGVTLVATGVFA
VipF	286	1	0.542	1	19	FLYYQNGALIGFLSIYFFY
CegC2	1102	1	0.599	1	28	VFADLASALVTLCTLGGANLYAGRWRLF
RavC	251	5	-2.593	1	32	MIYYSGARLVILLSFIGILLGATVSQTVYALL
LegU1	188	1	-0.804	1	20	VALIEVLTVILVFAAITSLA
Lpg0181	303	1	1.232	1	20	LTGIIFLLPTLITGVFCQNF
Ceg9	241	2	-0.008	2	27	YLEITTVALAMPFIALLSTVGRCIATV
Lpg0257	330	1	-3.377	1	24	AGFPVIILLCVFLFILIYHYSYLF
Lpg0260	132	1	1.311	1	24	YITLRNFAVVYNGLTGLAVLYSLW
SdbA	1116	2	-1.934	2	25	LTYGRFLLAGVVYALLTPLIWLAGW
AnkG/ ankZ/ legA7	514	1	-2.264	1	20	MIYVLICIMLVCGLVVYQLM
LegY	449	1	-0.137	1	23	IMLKRIFFLMIFFVSQTMASVFT
Ceg15	349	1	-1.813	1	29	LLLKPILGYSCFNFSISVFILILFYIPLT
Lpg0518	282	3	0.671	1	23	FWAGLALVGITPIINGLYACYYA
SidA	474	2	-2.762	1	33	IISSILIAAIIALSIAAITLGGLAAAIIGTVIA
Lpg0716	337	3	-2.815	3	35	FLPIGFFLLITTPIPLTIIACVSMLLVHWLITNLI
Ceg18	243	1	0.558	1	19	AVQILGVILAAMGMLMFAS
LidA	729	1	0.966	1	18	LFKRRLMAALFLWYLSKK
Lpg0963	413	1	1.05	1	28	LLMLGTTGPMGVALFSQIIASKLVTAFF
RavM	686	1	-0.416	1	24	LVGSVRFIFTFSLFCFRAIAGLLT
Lpg1148	503	1	-2.118	1	35	LMGFFLLPGLGVLMGAIAGAFIGVGAGLLISGIIA
RavP	318	1	-0.518	1	27	IRALTLIAIGTSPGIALGLPISTLLLI
Lpg1158	256	1	0.465	1	20	MIIYLSIWLIIGMKELIAFI
RavR	670	1	-2.51	1	24	LGYLGLIALIPISVIALYLVIHWM
RavS	630	1	-2.445	1	35	LIVGVSAALFLTGVLAPIGVALIAVGAALITTGLV
LegT	821	1	-1.391	1	21	YSTRIIITTILLLFTLMGYAG
SidG	974	1	-0.321	1	22	LKVGLTLMALPAMIVLAPLAAI

LegK1	529	1	-2.193	1	25	ALWLVAELSTAILLLGIGYLFAFGI
Lpg1484	269	2	-1.076	2	25	LLVNVGAAIIGLGVLYLLAASINYY
Lpg1578	149	1	-2.968	1	35	YLRLITILYLCTIGLVLYSSLPVLIEGVLIFLIVI
LegC6	672	1	0.977	1	27	LFKVIVNAIVFCVTFGASQGFFATTRA
LegL2	428	1	0.779	1	24	VYRRLGFVGVSPLSLAGLASFFVA
Ceg23	439	1	0.719	1	25	FLLKVGASVVGTSSALITLMGIAIL
Lem23	130	1	0.57	1	24	FFSIKNIGLGIVALAAAATVGLVV
SidB	417	1	1.28	1	22	WFKFIAIRTVFPPVLLWDLIKI
Lpg1661	372	8	-3.124	5	25	FKKFYLLAGIGLLFLLLGWLWNMSF
Lpg1666	467	2	1.453	2	25	FALLLPKGGVFCLVGDALMSANFLL
Lpg1667	463	1	-1.404	1	19	MIRFFIICFIFLVSHLAYA
Lpg1689	208	1	-2.134	1	35	LLLGSIPTFIFNPLFGVIMVSLAVTLLLPSGFYLL
PpeA/legC3	560	2	-4.633	1	35	LLLFGLTFLFSIPLILTISGVIPFFIAPALLYILV
PpeB	546	2	-2.859	2	34	AVLLTLVIAPPALLLLAGLGVGIATIIYAVKAYF
Lpg1751	436	1	1.155	1	18	KQAILFLLGALLHRYFRL
Lpg1776	161	2	-2.564	2	23	LLFVGGALLAAVGICCLPLLLCC
Lpg1803	311	1	-2.045	1	26	FFSTSNLLFFGGIVLAAVIFVPRLLT
YlfB/ legC2	405	1	-2.382	1	35	LWLKITLGIVLVVPTLTLGIAIQVASLIVVSVLTL
Lpg1907	601	2	-1.279	2	35	FGIIPLLTIGTLIGLSTAGIIPAIVSIGTLAVFAI
Lpg1959	664	1	-5.266	1	24	LMPYLLILLFLLLFILVQLGLFLT
PieE	635	2	-1.2	2	25	VFFAIPGALAALGVVSLVIAVVSYC
WipC	366	1	0.141	1	19	RMALFIIFIIHCFFSQSFA
Lpg2223	407	2	-0.811	1	31	VMVVLLHIGGIISAVALSVVGGLLGGILGFV
Lpg2271	216	1	-0.955	1	29	LMFWATAAFFSYVACVSVGIPLTFCDPLL
YlfA/ legC7	425	1	-2.095	1	32	LVLTAPTLAVGLFAHIGVLLVIGGVTGLTYTA
MavE	208	1	0.334	1	22	YVLFPFVAATVAVAATAASVLF
LegA1	368	2	0.683	2	27	LLRLALRLGNQGACALLLSIPSVLALT
Lpg2443	185	1	-0.081	1	25	MIRHFFIVLFFFVGFMNQASALPWF
MavI	204	1	0.701	1	18	ALKWGLFGGAALAITYAV
AnkF/ legA14/ ceg31	921	2	-0.225	2	20	FVRRFIGVIATLAVIPALVV
LepB	1294	2	-2.241	2	25	WKYLLANVTLGVFLLGIGYLAAILI
Lpg2505	295	1	1.106	1	20	CLLLANAYFYLAGFYLSLDL
SdjA	807	1	0.691	1	22	YLYLNIIAEYLLVIQLVIGCYG
Lpg2552	555	5	-2.46	4	27	YGLLRCVLLGMVSTVFLAVYATLTWFT
SidF	912	2	-2.161	2	35	LLFPPAGLAVAALGIGVAAGAAIVAGVTFAATYLF
LegS1	385	1	-0.137	1	25	FSIRVILLLTGFVSLAFSAPKFLTI
Lpg2628	250	2	-4.383	1	28	LLEYLLHPFFLCLLALGICIFLLYRRLA
Lpg2637	403	1	1.492	1	30	LAALLFPAQPLSMLGFFQGVAAAMGRDYFY
Lpg2692	176	2	0.748	1	19	AFTLTTLMAGVHVGIYYIL
LegN	343	1	1.379	1	21	MIWYKLFNLICGSIISLARSL
Lpg2745	655	1	1.368	1	22	LLYRIKTFLVSLFVTPFEVFPF
Lem29	468	4	-2.588	4	23	TFWPVLLFSVIVGLAAFSLYWYV
MavN	683	4	-4.489	1	33	IPYYLFLLLLTVGASLILGFLSFGGMYALWPVL
Lpg2828	418	1	-0.128	1	31	ILCLLRAIFYLIHHCCTIEQLTLIPYLIYFR
SidH	2225	1	1.334	1	24	RTRFVMTTFNALVMNICFSKYYVM

Lpg2879	583	2	1.139	1	19	WYAPYSLFLSGLALTNYLV
Lpg2884	245	3	-1.035	2	30	AISVAFWSSLVVAGAAALTVAFWPAALAAV
Lpg2885	184	2	-2.37	1	30	YLTMLLIGSAVVGLSVTFLALFFPPVLAAF
Lpg2888	637	5	-2.069	5	33	AVVFASFVFLPLGIGFAVLGATIGVALATNLMI
Lpg2912	495	1	-0.104	1	28	FFTALWYAAPTAFDFILGGILIYGRAYY
Lpg3000	612	3	-2.294	2	34	ALLMMGFTASMLVSPPILVVGCFFACTLAVAMYL

7.2 List of ΔG_{app} SRP predictions in TMD-containing T3SS substrates with sliding window 12-17 and length correction OFF

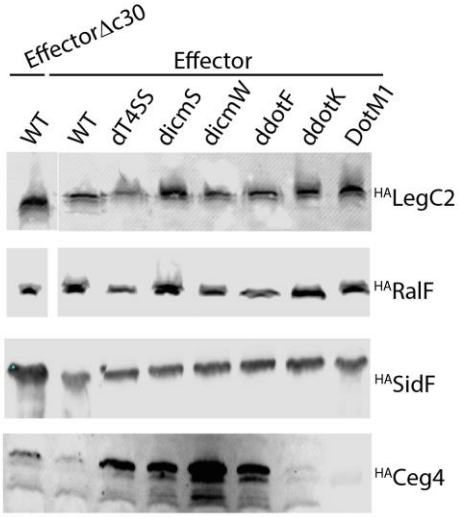
Protein name	Protein length/ aa	ΔGpred (window 18-35 length corr. ON) # TMS	ΔGpred (most hydro-phobic)	number (most hydro-phobic)	length (most hydro-phobic)	sequence (most hydrophobic)
Lpg1959	664	13	-4.239	10	14	YLLILLFLLLFILV
Lpg2628	250	11	-3.611	2	17	FFLCLLALGICIFLLYR
Lpg0257	330	5	-3.308	1	14	VIILLCVFLFILIY
Lpg1661	372	15	-2.993	9	17	FYLLAGIGLLFLLLGWL
Lpg1667	463	5	-2.716	1	12	IRFFIICFIFLV
AnkG/ ankZ/ legA7	514	8	-2.258	1	17	MIYVLICIMLVCGLVVY
MavN	683	18	-2.051	1	12	MLLIIIFIWYKV
Ceg15	349	9	-1.997	4	13	FSISVFILILFYI
LidA	729	3	-1.923	1	12	RRLMAALFLWYL
LegT	821	17	-1.799	2	13	IIITTILLLFTLM
Lpg2443	185	2	-1.777	1	12	FFIVLFFFVGFM
WipC	366	6	-1.73	1	14	RMALFIIFIIHCFF
Lpg0041	748	11	-1.701	2	14	FLTAFLMLLCSLAL
Lpg1803	311	5	-1.625	5	15	LLFFGGIVLAAVIFV
Lpg2552	555	17	-1.609	4	12	MALCMVFCLGLL
SidA	474	10	-1.578	4	16	ILIAAIIALSIAAITL
RavC	251	10	-1.531	7	17	LLLYVYLVSAIFFSFYL
LegK1	529	10	-1.495	9	14	LSTAILLLGIGYLF
Lpg1578	149	3	-1.475	1	17	LRLITILYLCTIGLVLY
LegY	449	9	-1.466	1	12	LKRIFFLMIFFV
LegU1	188	4	-1.427	4	12	VLTVILVFAAIT
PpeA/ legC3	560	5	-1.411	2	17	LLLFGLTFLFSIPLILT
RavR	670	3	-1.408	3	17	YLGLIALIPISVIALYL
Lpg2888	637	15	-1.377	10	14	IAYAVGLMLAFALL

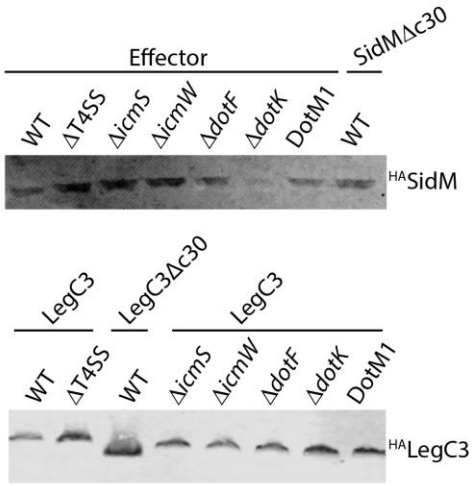
Lpg1484	269	4	-1.322	4	12	AAIIGLGVLYLL
Lem29	468	10	-1.322	3	12	FWLLAIAGTLLA
Lpg1776	161	5	-1.226	5	15	ALLAAVGICCLPLLL
SidF	912	9	-1.216	7	14	LLAAGGFVALTLLF
Lpg1689	208	4	-1.16	4	15	LFGVIMVSLAVTLLL
Lpg0716	337	10	-1.052	8	12	FLPIGFFLLITT
PpeB	546	7	-0.993	7	17	LLLLAGLGVGIATIIYA
Lpg3000	612	13	-0.886	11	12	FFACTLAVAMYL
LepB	1294	13	-0.839	13	12	TLGVFLLGIGYL
SdbA	1116	17	-0.836	7	12	RFLLAGVVYALL
Lpg1158	256	4	-0.763	1	12	IIYLSIWLIIGM
RavM	686	9	-0.752	1	12	RFIFTFSLFCFR
Ceg23	439	7	-0.733	6	12	ALITLMGIAILS
LegS1	385	9	-0.675	1	12	FSIRVILLLTGF
VipF	286	9	-0.659	2	12	ALIGFLSIYFFY
YlfB/ legC2	405	5	-0.593	3	12	VASLIVVSVLTL
Lpg1907	601	11	-0.572	8	12	LIVFGIIPLLTI
Ceg2	368	9	-0.537	6	14	WVYLGIGIMGFGIL
Lpg0181	303	5	-0.528	1	12	LTGIIFLLPTLI
Lpg1751	436	8	-0.523	2	12	AILFLLGALLHR
RavS	630	4	-0.5	3	15	IFMGLIVGVSAALFL
Ceg18	243	3	-0.486	3	14	ILGVILAAMGMLMF
PieE	635	5	-0.447	5	15	ALAALGVVSLVIAVV
Lpg1148	503	8	-0.36	5	12	AIIGGLMGFFLL
Lpg2885	184	5	-0.335	1	12	YLTMLLIGSAVV
Lem23	130	1	-0.334	1	12	IGLGIVALAAAA
Lem1	1324	30	-0.241	1	12	SLSVLTIFLTYA
Lpg2912	495	11	-0.238	7	12	FILGGILIYGRA
SidG	974	10	-0.195	10	13	LTLMALPAMIVLA
SdjA	807	18	-0.193	13	12	YLLVIQLVIGCY
Lpg0518	282	9	-0.165	4	12	LTLSTCLLLMAP
Lpg0963	413	12	-0.152	4	12	ASRVVIAACLYV
Lpg2828	418	7	-0.137	4	12	ILCLLRAIFYLI
AnkF/ legA14/ ceg31	921	13	-0.114	12	12	RRFIGVIATLAV
YlfA/legC7	425	6	-0.089	3	17	LFAHIGVLLVIGGVTGL
Ceg4	397	6	-0.08	3	12	FGFALGVTLVAT
Ceg9	241	6	-0.064	4	12	TVALAMPFIALL
MavI	204	6	-0.058	1	12	VKLCFALFRLYC
Lpg2271	216	4	-0.046	2	13	LMFWATAAFFSYV
Lpg2505	295	6	-0.034	2	12	FYCGYVLYLAAL
Lpg2745	655	12	0.002	5	12	YYLLELFVLINM
RavP	318	6	0.049	3	12	IRALTLIAIGTS
Lpg2884	245	9	0.089	5	14	LVVAGAAALTVAFW
Lpg0260	132	4	0.11	2	12	GLTGLAVLYSLW
MavE	208	5	0.138	1	12	TRFIMLSFVTGY
Lpg2879	583	19	0.277	4	12	LFLSGLALTNYL

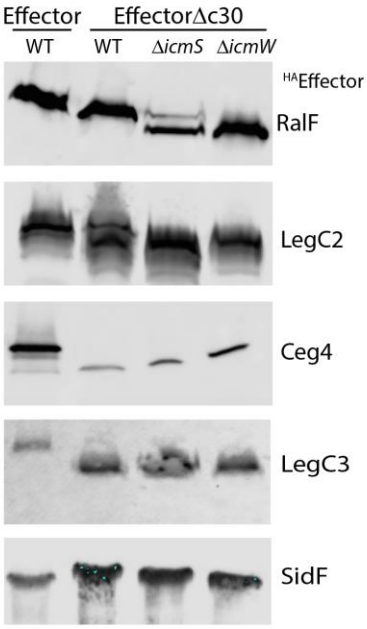
Lpg2692	176	2	0.309	2	12	TICFSMCVGVCL
LegC6	672	7	0.317	6	12	VIVNAIVFCVTF
CegC2	1102	20	0.334	19	12	LASALVTLCTLG
LegA1	368	11	0.474	11	14	ACALLLSIPSVLAL
LegL2	428	12	0.517	12	12	LSLAGLASFFVA
CegC1	525	8	0.573	4	12	VYILGHAIALRY
Lpg1666	467	10	0.609	2	12	AGLASALALLVK
Lpg2223	407	10	0.755	4	13	LLHIGGIISAVAL
LegN	343	9	0.803	9	13	LFNLICGSIISLA
SidB	417	8	1.105	1	12	RTVFPPVLLWDL
Lpg2637	403	9	1.297	4	12	LNFFVECLVAAF
SidH	2225	15	1.471	7	12	TTFNALVMNICF

7.3 Figure S1: Expression levels of effector proteins in *L. pneumophila*

The expression of the soluble Dot/Icm effector proteins ^{HiBiT}SidM and ^{HiBiT}Ralf or the TMD-effector ^{HiBiT}LegC3, ^{HiBiT}LegC2, ^{HiBiT}Ceg4 or ^{HiBiT}SidF, respectively in *Legionella pneumophila* wild type or mutant strains were investigated using SDS-PAGE and Western Blotting.


EffectorΔc30
Effector
WT
WT
dT4SS
dicmS
dicmW
ddotF
ddotK
DotM1
HALegC2
HARalF
HASidF
HACeg4


Effector
SidMΔc30
WT
ΔT4SS
ΔicmS
ΔicmW
ΔdotF
ΔdotK
DotM1
WT
HASidM
LegC3
LegC3Δc30
LegC3
WT
ΔT4SS
WT
ΔicmS
ΔicmW
ΔdotF
ΔdotK
DotM1
HALegC3


Effector
EffectorΔc30
WT
WT
ΔicmS
ΔicmW
HAEffector
RalF
LegC2
Ceg4
LegC3
SidF