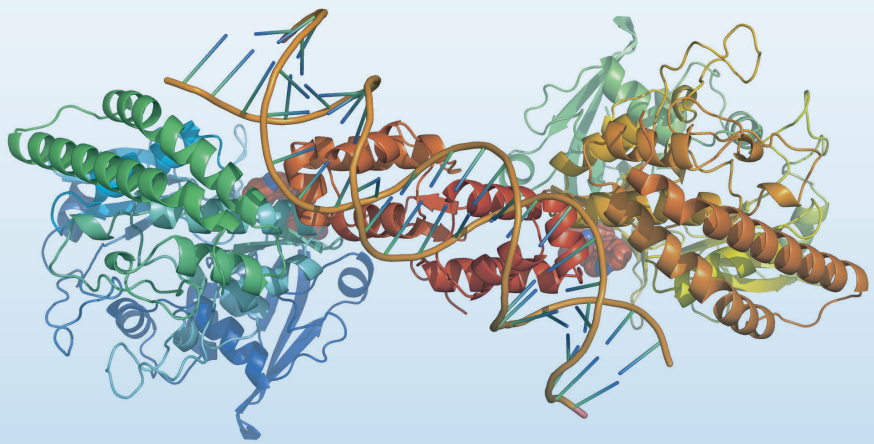
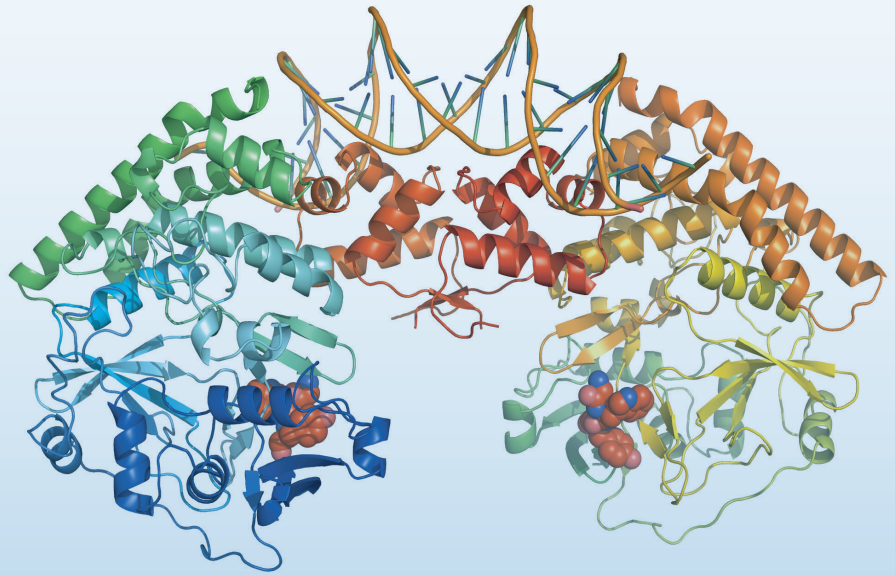


An Integrative Structural Biology Approach in the Study of Bacterial Persistence and Virulence Yurong Wen 2014

ISBN 978-90-5989-685-7

An Integrative Structural Biology Approach in the Study of Bacterial Persistence and Virulence

Yurong Wen



Thesis submitted in fulfillment of the requirements for the degree of
Doctor in Science: Biochemistry and Biotechnology
Gent University 2014

UNIVERSITEIT GENT

FACULTY OF SCIENCES

Department of Biochemistry and Microbiology
Laboratory for Protein Biochemistry and Biomolecular Engineering

An Integrative Structural Biology Approach in the Study of Bacterial Persistence and Virulence

Yurong Wen
Promoter: Prof. Dr. Bart Devreese
Co-Promoter: Prof. Dr. Savvas Savvides
Gent University 2014

Thesis submitted in fulfillment of the requirements for the degree of Doctor in Science: Biochemistry and Biotechnology

Examination Committee:

Prof. Dr. Nico Callewaert (Chairman)

The Inflammation Research Center
Ghent University and Flanders Institute for Biotechnology (VIB)

Prof. Dr. Bart Devreese (Promoter)

Unit for Biological Mass Spectrometry and Proteomics
Laboratory for Protein Biochemistry and Biomolecular Engineering, Ghent University

Prof. Dr. Savvas Savvides (Co-Promoter)

Unit for Structural Biology and Biophysics
Laboratory for Protein Biochemistry and Biomolecular Engineering, Ghent University

Dr. Bjorn Vergauwen (Reading Committee)

Unit for Structural Biology and Biophysics
Laboratory for Protein Biochemistry and Biomolecular Engineering, Ghent University

Prof. Dr. Remy Loris (Reading Committee)

Structural Biology Brussels Laboratory, Department of Structural Biology
Vrije Universiteit Brussel (VUB) and Flanders Institute for Biotechnology (VIB)

Prof. Dr. Laurence Van Melderen (Reading Committee)

Unit for Bacterial Genetics and Physiology
Institute of Molecular Biology and Medicine, Université libre de Bruxelles (ULB)

Prof. Dr. Frank Sobott (Reading Committee)

Biomolecular & Analytical Mass Spectrometry group, University of Antwerp

Dr. Kim Heylen

Laboratory of Microbiology, Ghent University

This project was financially supported by the China Scholarship Council (CSC) and the Gent University Bijzonder Onderzoeksfonds (BOF)

Please refer to this work as below:

Yurong Wen (2014). "An Integrative Structural Biology Approach in the Study of Bacterial Persistence and Virulence". PhD thesis, Ghent University, Belgium.

ISBN: 978-90-5989-685-7

List of Abbreviations

ACN	Acetonitrile
AFM	Atomic Force Microscopy
AMPPNP	Adenosine 5′-(β, γ-imido) Triphosphate
AP	Affinity purification
ATD	Arrival Time Distribution
ATP	Adenosine 5′-Triphosphate
bp	Basepair
Cb	Carbencillin
Cm	Chloramphenicol
$CC_{1/2}$	Pearson correlation coefficient
CCS	Collision Cross Section
CID	Collision Induced Dissociation
eDNA	extracellular DNA
EM	Electron Microscopy
ESI	Electrospray Ionization
EFTu	Elongation factor Tu
ETD	Electron Transfer Dissociation
FDR	False Discovery Rate
FRET	Förster Resonance Energy Transfer
FTICR	Fourier Transform Ion Cyclotron Resonance
GltX	Glutamyl tRNA synthetase
HDXMS	Hydrogen Deuterium Exchange Mass Spectrometry
Hip	High persister factor
HTH	Helix-Turn-Helix
IMS	Ion Mobility Spectrometry
IMMS	Ion Mobility Mass Spectrometry
ISB	Integrative Structural Biology
ITC	Isothermal Titration Calorimetry
IUP	Intrinsically Unstructured Protein
K_D	Dissociation constant
Km	Kanamycin

LB	Lysogeny Broth
LCMS	Liquid Chromatography Mass Spectrometry
Maldi	Matrix assisted laser desorption ionization
MDR	Multidrug Resistance
MDT	Multidrug Tolerance
MIC	Minimal Inhibitory Concentration
MR	Metal Reducing
MS	Mass Spectrometry
MW	Molecular Weight
NMR	Nuclear Magnetic Resonance
OD	Optical Density
PCD	Programmed Cell Death
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PIN	Protein-Protein Interaction Network
PPI	Protein-Protein Interaction
(p)ppGpp	Guanosine pentaphosphate or tetraphosphate
SAXS	Small Angle X-ray Scattering
SDS PAGE	Sodium Dodecylsulfate Polyacrylamide Gel Electrophoresis
SEC	Size Exclusion Chromatography
SO	*Shewanella oneidensis*
SPR	Surface Plasmon Resonance
TA	Toxin-Antitoxin system
TADB	Toxin-Antitoxin system Database
TAP	Tandem Affinity Purification
TOF	Time of Flight
XLMS	Crosslinking Mass Spectrometry

Table of Contents

List of Abbreviation ..... I-II

Table of Contents ..... III-VI

Preface: Objectives and Achievements 1

Chapter 1 Introduction Section A: Toxin-Antitoxin systems: their role in persistence, biofilm formation and pathogenicity 5

1.1 Abstract 6

1.2 Introduction 7

1.3 Function and classification of Toxin-Antitoxin systems 8

1.3.1 Type I TA systems 9

1.3.2 Type II TA systems 11

1.3.3 Type III TA systems 13

1.3.4 Type IV and V TA system 13

1.3.5 Other TA systems 13

1.4 Toxin-Antitoxin systems and persistence 14

1.5 Toxin-Antitoxin systems and biofilm formation 16

1.6 Toxin-Antitoxin systems and virulence 19

1.7 Toxin-Antitoxin systems and antimicrobial drug development 20

1.8 Conclusion and Perspectives 25

1.9 References 26

Chapter 2 Introduction Section B: Mass Spectrometry in Integrative Structural Biology 37

2.1 Introduction 38

2.2 Mass Spectrometry and Ion Mobility 42

2.3 Mass Spectrometry and Chemical Crosslinking 46

2.4 Mass Spectrometry and Hydrogen Deuterium Exchange 50

2.5 Integrative Modeling 52

2.6 Conclusion and Perspectives 53

2.7 References 54

Chapter 3 Structural Insights into HipA-HipB Toxin-Antitoxin System in *Shewanella oneidensis* MR-1 63

3.1 Prologue: Structure and function of the HipA-HipB toxin antitoxin system in *E.coli* 64

3.2 Abstract 67
3.3 Introduction 68
3.4 Material and Methods 71
3.4.1 Expression and purification of recombinant proteins and complexes thereof 71
3.4.2 Crystallization and X-ray data collection of $HipAB_{so}$:DNA complex, $HipA_{so}$-AMPPNP-Mg and $HipB_{so}$ 72
3.4.3 Crystal structure determination of $HipA_{so}$-AMPPNP-Mg, $HipB_{so}$ and $HipAB_{so}$:DNA 73
3.4.4 Nano Liquid Chromatography Mass Spectrometry 75
3.4.5 Isothermal Titration Calorimetry 75
3.4.6 Small Angle X-ray Scattering 76
3.4.7 Bacterial growth experiments 76
3.4.8 Radioactive kinase asay 77
3.5 Results 78
3.5.1 $HipA_{so}$ and $HipB_{so}$ constitue a toxin-antitoxin module targeting an a typical operator DNA 78
3.5.2 $HipA_{so}$ and $HipB_{so}$ and operator DNA assemble into a distinct ternary complex 81
3.5.3 The flexible C-terminal tail of $HipB_{so}$ is essential for targeting the $HipA_{so}$ hydrophobic pocket to prevent proteolysis 85
3.5.4 $HipA_{so}$ is conformationally regulated by autophosphorylation 87
3.5.5 Functional investigation of S.oneidensis MR-1 $HipA_{so}$ (SO0706) 90
3.6 Discussion 90
3.7 References 92
Chapter 3 SUPPLEMENTARY MATERIALS 99
Chapter 4 Characterization of the *Shewanella oneidensis* MR-1 HipA-HipB Toxin Antitoxin system Assembly using Mass Spectrometry 107
4.1 Introduction 108
4.2 Results 109
4.2.1 HipA has three different conformational states: apoHipA, ATP bound HipA, phosphorylated HipA 109

4.2.2 Characterization of HipA, HipB and its operator DNA interactions by Ion Mobility Mass Spectrometry reveals a distinct assembly with respect to *S.oneidensis* MR-1 and *E.coli* HipAB complexes 113
4.2.3 Probing the HipA kinase conformational changes upon phosphorylation using chemical crosslinking 116
4.3 Discussion 119
4.4 Experimental procedures 120
4.4.1 Preparation of Protein and Protein-DNA complex Samples 120
4.4.2 ESI Ion mobility MS Analysis 120
4.4.3 Chemical Crosslink and LC-MS analysis of crosslinked peptides 121
4.5 References 123
Chapter 5 Modeling *Pseudomonas aeruginosa* Type II Secretion System XcpP and XcpQ Periplasmic Domain Assembly via Mass Spectrometry based Integrative Structural Biology 125
5.1 Introduction 126
5.2 Results 132
5.2.1 Mass Spectrometry Characterization of *P.aeruginosa* $XcpP_C$ and $XcpQ_D$ 132
5.2.2 Study of the *P.aeruginosa* XcpP and XcpQ Assembly using Chemical Crosslinking and Mass Spectrometry 136
5.3 Discussion 143
5.4 Experimental procedures 146
5.4.1 Protein overexpression and purification 146
5.4.2 Chemical Crosslink, Proteolysis and FTICR MS 147
5.4.3 Ion mobility Mass Spectrometry Analysis 148
5.5 References 149
Chapter 6 General Discussion: Conclusion and Perspectives 153
6.1 What needs still to be done in the study of Shewanella oneidensis MR-1 HipA-HipB toxin antitoxin system molecular mechanism? 154
6.2 Can the HipA-HipB TA system be an antimicrobial drug target 156
6.3 How far can we go with a mass spectrometry based integrative structural biology appraoch in the study of type 2 secretion system 157

6.4 References 158
Summary 161
Samenvatting 164
Acknowledgements 167
Curriculum Vitae .....

Preface

Objectives and Achievements

This PhD thesis outcome of a hybrid project between the Unit for Mass Spectrometry and Proteomics and the Unit for Structural Biology and Biophysics in the Laboratory for Protein Biochemistry and Biomolecular Engineering (L-ProBE) (http://www.lprobe.ugent.be). The Unit for Mass spectrometry and Proteomics is mainly working on use of mass spectrometry and proteomics for the analysis of proteins involved in microbial physiological processes. When the PhD project started off, a novel Waters SYNAPT HDMS mass spectrometry has been introduced in the lab. This instrument combines the triple quadrupole and Time of Flight configuration with a traveling wave ion mobility analyzer which can be used both for LC-MS analysis of complex peptide samples and for the investigation of intact proteins and the assembly of protein complexes. The Unit for Structural biology and Biophysics focuses on the elucidation of the structure function relationships for soluble and membrane associated protein using a in-house state-of-the-art platform. Multiple structural biology techniques such as X-ray crystallography, small angle X-ray scattering and biophysical techniques such as Isothermal Titration Calorimetry, Surface Plasmon Resonance are available to carry out interdisciplinary researches. The PhD project aimed to integrate all the available techniques as an integrative structural biology approach and to apply it on a set of protein that play an important role in the antibiotic resistance and virulence of bacteria.

Bacteria are reported to be responsible for most of infectious diseases which are worldwide responsible for more than 17 million deaths each year, despite the introduction of antibiotics in the mid 1990's. The lack of introduction of novel antibiotics and the growing resistance of bacteria against a wide range of antibiotics poses a huge challenge for the next decades. Multidrug resistance, tolerance and persistence are major causes of failure to respond to the conventional antibiotic treatment and further results in prolonged illness, increased mortality and higher costs for social security.

The PhD project focuses on two aspects of bacterial infections and resistance against antibiotics. We used structural biology approaches to understand the mechanisms of bacterial persistence by studying a toxin antitoxin system module. Furthermore, we investigated the structural assembly of a bacterial type 2 secretion system which is crucial for the transportation of virulence factors The objectives and achievements of this thesis are listed in the contents below:

Chapter one (Introduction section A). We reviewed the function and classification of toxin antitoxin systems, the latest development of the role of toxin antitoxin systems in biological phenomena such as persistence, biofilm formation, and virulence. We conclude with perspectives of toxin antitoxin systems as potential targets in the antimicrobial drug discovery.

Chapter two (Introduction section B). We highlight the advantages and principles of an integrative structural biology approach. We focused on the role of biological mass spectrometry in this field and discuss different mass spectrometry based approaches including Ion Mobility Mass Spectrometry, Chemical crosslinking Mass Spectrometry and Hydrogen Deuterium exchange Mass Spectrometry.

Chapter three. Here we report the results of the integrative structural biology approach involving mainly X-ray crystallography, Small angle X-ray scattering and other biochemical and biophysical techniques in the study of *Shewanella oneidensis* HipA-HipB toxin antitoxin system We provide an in depth understanding of the HipA-HipB module molecular mechanism by analysis of the atomic resolution three dimensional structure.

Chapter four. We further investigated the *Shewanella oneidensis* HipA-HipB toxin antitoxin system assembly by mass spectrometry including Ion Mobility Mass Spectrometry and chemical crosslinking Mass Spectrometry. Complementary insights were observed including

the transition conformation change of HipA and overall assembly of HipA-HipB and its operator DNA complex.

Chapter five. The integrative structural biology approach was applied in the study of the type 2 secretion system of the opportunistic human pathogen *Pseudomonas aeruginosa*. We focused on two components of this system, i.e. the inner membrane XcpP and outer membrane secretin XcpQ. In a collaborative effort with Ruben Van der Meeren from the unit for Structural Biology; we obtained multiple restraint data from chemical crosslinking Mass Spectrometry, affinity studies and structural modeling. We proposed a transition interaction motion of the XcpP and XcpQ periplasmic domain which could complement to understand the transportation of the exoprotein from periplasmic space into the type 2 secretion barrel.

Chapter six. General discussion: conclusion and perspectives of the PhD thesis.

Chapter 1

Introduction Section A

Toxin-Antitoxin systems: their role in persistence, biofilm formation and pathogenicity

The contents of this chapter are adapted from a minireview, accepted for publication in Pathogens and Disease (2014):

Yurong Wen, Ester Behiels, Bart Devreese. Toxin-Antitoxin systems: their role in persistence, biofilm formation and pathogenicity.

1.1 Abstract

One of the most pertinent recent outcomes of molecular microbiology efforts to understand bacterial behavior is the discovery of a wide range of Toxin-Antitoxin (TA) systems that are tightly controlling bacterial persistence. While TA systems were originally linked to control over the genetic material, for example plasmid maintenance, it is now clear that they are involved in essential cellular processes like replication, gene expression and cell wall synthesis. Toxin activity is induced stochastically or after environmental stimuli, resulting in silencing of the above mentioned biological processes and entry in a dormant state. In this minireview, we highlight the recent developments in research on these intriguing systems with a focus on their role in biofilms and in bacterial virulence. We discuss their potential as targets in antimicrobial drug discovery.

1.2 Introduction

Almost all bacteria possess genes whose expression can inhibit cell growth and even result in cell death when overexpressed, reflecting similarities to eukaryotic programmed cell death or apoptosis. Among those, Toxin-Antitoxin (TA) systems contribute largely to the bacterial epigenetic regulatory machinery controlling bacterial survival (Engelberg-Kulka et al., 2006; Yamaguchi et al., 2011). TA modules usually consist of a stable toxin and a degradation-prone antitoxin, mostly encoded in an operon, resulting in tight co-transcription and co-translation of the toxin and antitoxin (Gerdes et al., 1985; Ogura and Hiraga, 1983). The cognate antitoxin is either a protein or a small RNA molecule that counteracts the toxin activity by acting as a direct inhibitor or by controlling toxin production. They were first reported as addiction systems leading to plasmid maintenance. In plasmid-free daughter cells, the toxin becomes active by degradation of the instable antitoxin that can no longer be replenished, leading to post-segregational killing (Gerdes et al. 1986). Later on, they were also discovered in chromosomal genomes where they typically play a role in the stabilization of the bacterial chromosome and integrative conjugative elements (Wozniak and Waldor, 2009). Today, the definition of TA modules is no longer restricted to entities that provide control over genetic material, but expanded to a wide range of biological functions including growth control, defense against phages, biofilm formation, persistence, programmed cell death (PCD), general stress response etc…(Wang & Wood 2011, Engelberg-Kulka H et al. 2006). Basically, toxins within TA modules are proteins whose activity usually leads to inhibition of cell growth by interfering with cellular processes such as DNA replication, translation, cytoskeletal polymerization, phage infections abortive and cell division (Figure 1.1).

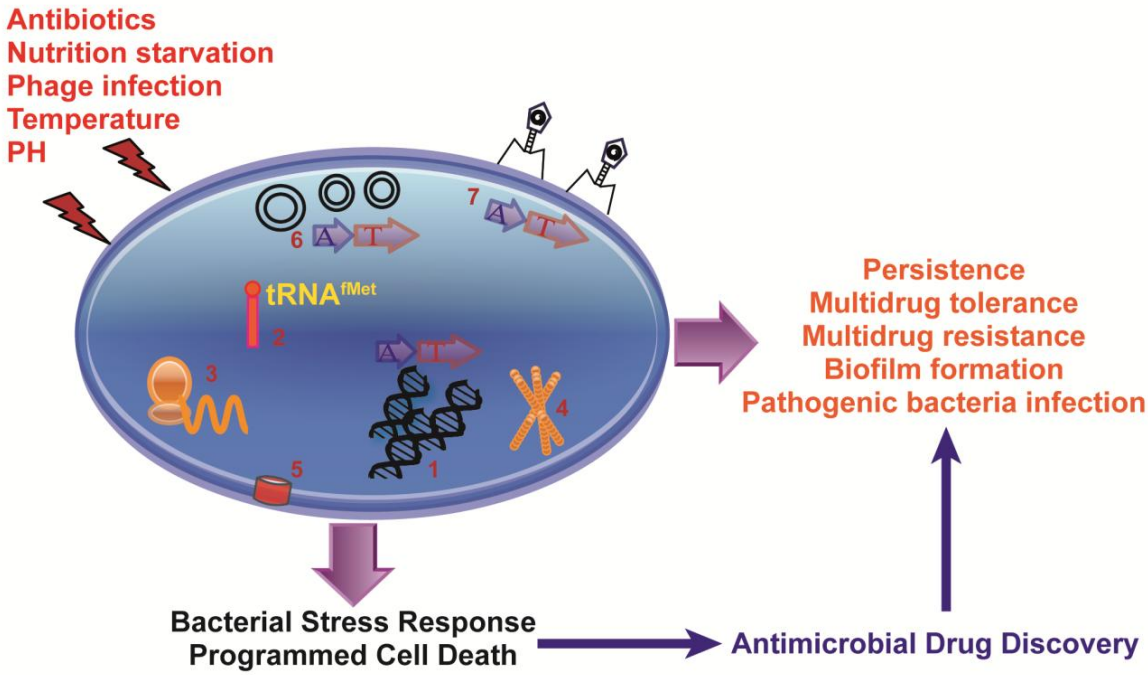


Figure 1.1 Toxin-Antitoxin systems are involved in a broad range of cellular processes. The numbers refer to: 1. DNA replication, 2. tRNA related translation, 3. Macromolecular synthesis, 4. Cytoskeletal polymerization, 5. Cell wall disruption, 6. Plasmid maintenance and 7. Phage infections abortive. The investigation of TA system molecular mechanism could lead to the antimicrobial drug discovery dealing with a number of issues inadequately tackled and covered by current therapies.

1.3 Function and classification of Toxin-Antitoxin systems

Given the diversity in structure and function of TA systems, Hayes proposed a classification into three major types based on the nature and action of the antitoxin (Hayes, 2003). However, recent discoveries resulted in the addition of new classes of TA systems (Masuda et al., 2012; Wang et al., 2013).

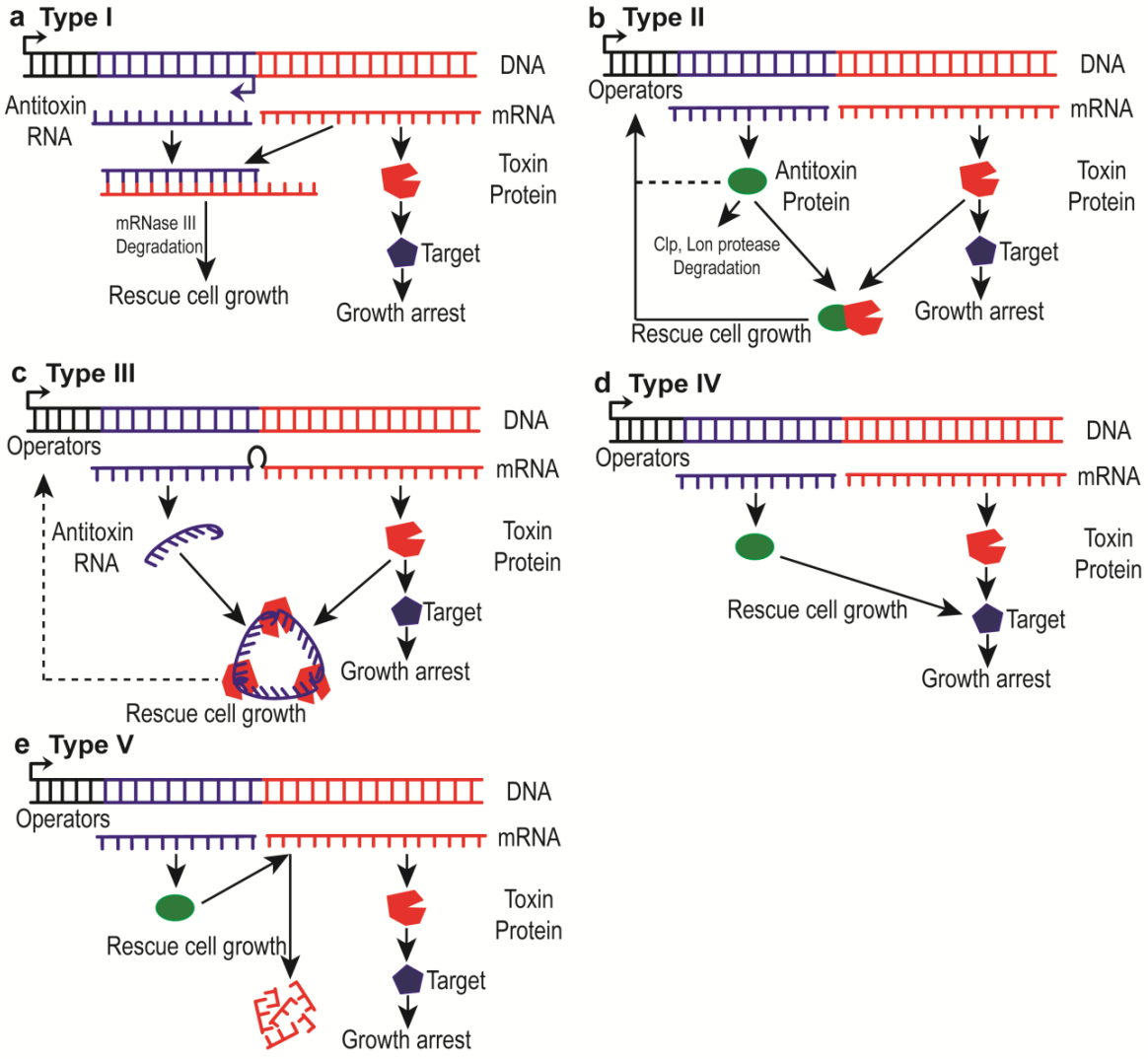


Figure 1.2 Typical mode of action of the 5 classes of Toxin-Antitoxin systems

1.3.1 Type I TA systems (Figure 1.2a)

The type I TA system antitoxins are small regulatory RNA's (sRNA), typically of 50 to 200 nucleotides. Their genes are usually adjacent but antisense to the toxin. Almost all type I TA system toxins are small hydrophobic proteins that are assumed to generate pores in the inner membrane, disrupting the membrane potential to inhibit ATP synthesis and consequently blocking energy-demanding processes such as protein synthesis.

The antitoxin sRNA acts by base pairing to the co-transcribed toxin mRNA, blocking ribosome binding, followed by degradation of the toxin-antitoxin RNA duplex by RNase to suppress the translation of the toxin (Fozo et al., 2008; Van Melderen and De Bast, 2009). The type I TA system was first found in plasmids and attributes to plasmid maintenance through post-segregational killing. Additional type I TA systems evolved in bacterial chromosomes, although no sequence homology to the plasmid-borne systems was reported (Fozo et al., 2010). It was indicated that the type I TA systems can also stabilize chromosomal segments and maintain integrative conjugative elements to prevent the spread of phage infection (Wozniak and Waldor, 2009).

The best studied type I TA system, *hok*/*sok* was found originally in plasmid R1, but homologues were also discovered in other plasmids, e.g. R100, F (Gerdes et al., 1990) and later on the *Escherichia coli* chromosome (Pedersen and Gerdes, 1999). Ectopic *hok* expression leads to cell content leakage and ghost cell formation (Gerdes et al., 1986b). Sok RNA is very unstable as it is rapidly degraded by RNase E with a half-life of only 30s. Formation of a *hok*-sok mRNA duplex likewise results in rapid removal by RNase III mediated degradation to neutralize the Hok toxicity (Gerdes et al., 1992). Synthetic anti-Sok peptide nucleic acid were designed to block the *hok* mRNA-sok RNA interaction, which successfully resulted in the accumulation of Hok protein and bacterial cell death (Faridani et al., 2006).

Well-characterized type I TA system families further include LdrD-RdlD, TisB-IstR, ShoB-OhsC, SymE-SymR (Faridani et al., 2006; Gerdes and Wagner, 2007). The latter has an atypical toxin function; SymE is an RNase. The system is SOS-responsive, and is potentially involved in clearance of damaged RNAs (Kawano et al., 2007). Bioinformatics resulted in the discovery that loci containing type I toxins are widely spread in bacterial chromosomes. The type I TA systems can occur repeatedly within the chromosome, such as LdrD repeats (3-7

times), LbsA (2-7 times), the gene encoding the type I toxin with the broadest distribution, Hok, is repeated 4-15 times in *E.coli* (Fozo et al., 2008). The evolution and distribution of type I TA system originates from lineage specific duplication (Fozo et al., 2010).

1.3.2 Type II TA systems (Figure 1.2b).

In type II TA systems, both the toxin and the antitoxin are proteins. The toxin acts as an inhibitor of essential cellular functions such as replication or protein synthesis. The type II antitoxin is a DNA binding protein that suppresses the transcription of the toxin. When released from the nucleoid, the type II antitoxins are generally labile and can be degraded by Lon or Clp proteases (Aizenman et al. 1996; Van Melderen, 1996). Most studies on type II TA systems involve MazE-MazF, HipA-HipB, RelB-RelE, MqsR-MqsA, Phd-Doc, and VapB-VapC. Typically, the Type II antitoxins form a complex with their toxins sequestering the toxins in an inactive form in the nucleoid, although there are exceptions where the toxin-antitoxin complex does not bind to DNA, e.g. in MqsR-MqsA (Brown et al., 2013).

MazF-MazE was the first identified chromosomal addiction module in *E.coli* (Masuda et al. 1993, Engelberg-Kulka et al. 2005). MazF is an endoribonuclease which can specifically cleave mRNA at the ACA site resulting in protein synthesis inhibition (Zhang et al., 2003). MazE can be degraded by the ATP dependent ClpAP serine protease. The MazF-MazE mediated cell death requires the involvement of the quorum-sensing extracellular death factor EDF (Kolodkin-Gal et al., 2007). Similar activities are found in the RelE family, but the latter differs from the others as it cleaves only in the presence of the ribosome (reviewed in (Cook et al., 2013; Park et al., 2013)). VapC, classified as a PIN domain ribonuclease, was originally reported to inhibit translation by cleaving the initiator $tRNA^{fMet}$ (Winther and Gerdes, 2011), while *Mycobacterium tuberculosis* VapC20 is recently demonstrated to cleave

the 23SrRNA in the region were eukaryotic toxins like α-sarcin and ricin act (Winther et al., 2013).

Among the type II TA system family, HipA-HipB is specifically recognized for its role in persister formation. The toxin HipA shows sequence similarity to eukaryotic serine/threonine kinases. It was originally proposed to phosphorylate elongation factor Tu (EF-Tu) resulting in inhibition of protein biosynthesis (Schumacher et al., 2009, 2012). However, this was not confirmed, and more recently, glutamyl-tRNA synthetase (GltX) was discovered to be the actual substrate (Germain et al., 2013). The HipB antitoxin is a helix-turn-helix protein that neutralizes HipA by forming a complex that sequesters HipA in the nucleoid after HipB dependent binding to the operators upstream of the HipAB operon, suppressing the transcription of HipA (Correia et al., 2006; Schumacher et al., 2009, 2012).

The Doc toxin from the Doc-PhD system was first reported to acts as an inhibitor of translational elongation through binding to the 30S ribosomal subunit. However the protein shows similarity to AMPylating enzymes and recently it was discovered that it actually evolved to be a phosphorylating enzyme with EF-Tu as the substrate for its kinase activity (Castro-Roa et al., 2013).

The wide distribution and diversity of type II TA systems was recognized in TADB, a type II TA system database derived from integrated experimental and bioinformatics approaches (Shao et al., 2011). Type II TA systems were both characterized on plasmid and on chromosomal genomes and it was proposed that the evolution of type II TA systems from plasmid maintenance genes towards the more diverse function of chromosomal distributed systems originates from horizontal gene transfer (Leplae et al., 2011).

1.3.3 Type III TA systems (Figure 1.2c).

Here, the toxin protein function is inhibited by directly interacting with pseudoknots antitoxin RNA, forming an RNA protein complex. The pseudoknots antitoxin RNA is coded by a short tandem repeat upstream of the toxin gene (Blower et al., 2012). ToxI-ToxN, originally discovered in the phytopathogen *Pectobacterium atrosepticum*, was the first type III TA system for which a detailed molecular mechanism was revealed (Fineran et al., 2009; Short et al., 2013). ToxN is an endoribonuclease sharing structural similarity to type II TA systems such as MazF. ToxN is able to cleave ToxI, its own mRNA and other cellular mRNAs (Blower et al., 2011). The sRNA antitoxin acts as a specific inhibitor for the cognate toxin. The type III TA systems fall in three evolutionary related subfamilies with respectively ToxI-ToxN, TenpI-TenpN and CptI-CptN being the prototypic representatives (Blower et al., 2012). So far, all seem to share potential function to abort phage infection.

1.3.4 Type IV and V TA system (Figure 1.2d and 1.2e).

These are recently two new discovered TA systems modules. In the type IV TA system, the toxin CbtA inhibits the polymerization of the cytoskeletal proteins MreB and FtsZ which are essential in cell division. The antitoxin protein CbeA (YeeU) interferes by directly binding with the toxin targets MreB and FtsZ rather than forming a toxin-antitoxin complex as in other systems (Masuda et al., 2012). In the type V TA system, the antitoxin GhoS was identified as an endoribonuclease and cleaves specifically the toxin GhoT mRNA, which encodes a membrane lytic peptide (Wang et al., 2013).

1.3.5 Other TA systems

With the increasing availability and analysis of prokaryotic genomes, we expect novel TA modules to be discovered and the current classification system might require revision. A

recent large scale cloning experiment pointed already in this direction (Sberro et al., 2013). During finishing the draft of this review, this was illustrated by the description of the *Caulobacter crescentus* SocAB system that has a mode of action not included in the current classification criteria of TA systems. SocB toxin blocks replication elongation and the SocA antitoxin acts as a proteolytic adapter for protease ClpXP dependent degradation of SocB (Aakre et al., 2013).

1.4 Toxin-Antitoxin systems and persistence

TA systems are widely distributed in bacteria. At least 36 TA systems are discovered in the *E.coli* K-12 genome and as many as 65 TA systems in *Mycobacterium tuberculosis* H37Rv (Shao et al., 2011; Yamaguchi et al., 2011). The observation of large pools of TA systems in pathogens like *Mycobacterium tuberculosis* was reported to increase the capacity of entering the persistence state, further resulting in multidrug tolerance (Ramage et al., 2009).

Persistence was already described in the early 1940's during studies of the mechanism of penicillin when it was observed that a small population of persister cells survived as dormant cells, different from traditional antibiotic resistance mechanisms (Bigger, 1944; Hobby et al., 1942) . The characteristics used to define the persister phenotype are still constantly updated and include dormancy, multidrug tolerance and a stochastic nature. Persistence can be described as the ability of a few cells within a genetically homogenous population to survive different types of stress (Scherrer and Moyed, 1988). These persisters are not mutants but phenotypic variants of the wild type. They show the same minimal inhibitory concentration (MIC) for antibiotic treatment as the rest of the population and upon regrowth they are as sensitive as is the original population. Persistence should thus be regarded as a population phenomenon and not as an individual trait, and is characterized by a biphasic killing pattern in a time- or dose-response experiment: at first, fast killing is observed, which is followed by

survival of a small fraction that is virtually insensitive to increasing antibiotic concentrations (Moyed and Bertrand, 1983; Moyed and Broderick, 1986). A small number of persistent cells seems to present in any bacterial culture, very small (less than 0.1%) in log phase cells up to 1% of the population in stationary phase cells, reflecting a stochastic induction of the persistence mechanism (Balaban et al., 2004). However, .persistence is dramatically increased upon arrest of protein synthesis upon environmental stimuli (Kwan et al., 2013). The environmental conditions to stimulate formation of persister cells are diverse and include antibiotic pressure, starvation, oxidative stress, extreme pH and temperature and the host immune system (Wang and Wood, 2011).

It took until 1983 to identify the first gene that is involved in persistence. In *E.coli* K-12, it was found that the *hipA7* allele showed a ratio of 10^{-2} of persisters upon ampicillin treatment, whereas for the wild type strain this ratio is as low as 10^{-5} to 10^{-6} (Moyed and Bertrand, 1983). Moyed and his coworkers systematically described the persister phenomenon (Scherrer and Moyed, 1988) and explored the *hip* gene structure and mechanism of *hipA* autoregulation (Black et al., 1991, 1994). HipA7 is the result of a double mutation (G22S, D291A) and has similar effects as overexpression of HipA (Korch et al., 2003).

Despite the evidence for the involvement of TA systems in the development of bacterial persistence, the effects of deletion of a single TA system on persistence are variable (Keren et al. 2004, Kim & Wood 2010). Deletion of *tisB*, *mqsR* or *hipAB* has a dramatic effect on the level of persistence, whereas deletion of *mazEF* does not. The effect is more pronounced in stationary-phase or biofilm cells where there is a higher frequency of persistence formation. Redundancy of TA systems, in particular of the type II class, on most bacterial genomes explains this phenomenon. For example, progressive decrease of survival upon antibiotic treatment was only observed when stepwise up to 10 endoribonuclease encoding TA systems

were disrupted in *E.coli*, ultimately resulting in a 100-200 fold reduction compared to the wide type *E.coli* of persistence frequency (Maisonneuve et al., 2011).

Induction of TA system activity is modulated by the signaling nucleotide (p)ppGpp acting via a cascade involving Lon protease, inorganic phosphate and other stress induced signaling pathways. Stochastic variation leads to a higher level of (p)ppGpp in a subset of exponentially growing cells that switch to a slow growth mode and activate the TA systems (Amato et al., 2013; Maisonneuve et al., 2013). Furthermore, Kwan et al. proved that chemical inhibition of protein synthesis also increases the persister-like cell formation. This work included an important warning, as many antibiotics like fluoroquinolones act on the protein synthesis machinery and could provoke the formation of cells that develop persistence to other drugs, for example in combination therapies (Kwan et al., 2013).

Recent work revealed that the autoregulation of TA systems, which involves conditional cooperativity, results in a threshold phenomenon that could explain the spontaneous (Cataudella et al., 2012) and stochastic persister cell formation (Gelens et al., 2013; Rotem et al., 2010). Further, Fasani and Savageau provided a model explaining that the redundancy of type II TA systems in prokaryotic genomes increases the frequency of persister cell formation and how the TA systems give rise to a bimodal population of persisters and normal cells (Fasani and Savageau, 2013).

1.5 Toxin-Antitoxin systems and biofilm formation

Most bacteria in the environment live in biofilms, multispecies consortia embedded in an extracellular matrix adhering to both biotic and abiotic surfaces (Stewart and Franklin, 2008). Biofilm formation is a well-controlled developmental process with distinctive steps including initial attachment to the surface, maturation of the biofilm and detachment of cells and dispersal (Hall-Stoodley and Stoodley, 2002). Many chronic infections are associated with

the biofilm formation capability of the pathogenic bacteria (von Rosenvinge et al., 2013). Examples of biofilm-associated infections include cystic fibrosis (CF) by *Pseudomonas aeruginosa* (Mulcahy et al., 2013) and tuberculosis (TB) by *Mycobacterium tuberculosis* (Bjarnsholt, 2013). The capability to form biofilm is demonstrated to be the main strategy that the pathogenic bacteria like *Pseudomonas* use to survive against the host immune defense (Rybtke et al. 2011).

Besides the involvement of pathogenicity development and persistence, biofilm also plays an important role in engineering applications. The ability that *Shewanella oneidensis* MR-1 can respire on the insoluble metal substances is dependent on its biofilm formation capacity. Researchers express high interest in *Shewanella oneidensis* MR-1 for its role in biogeochemistry, its potential for bioremediation of heavy metals and aromatic compounds, nanowire synthesis application as well as for its use in microbial fuel cells (Beveridge et al., 2009; El-Naggar et al., 2010; Heidelberg et al., 2002). All of these applications rely on the potential of the organism to form biofilms at inorganic surfaces.

It is widely accepted, mainly through the work of Kim Lewis and co-workers, that biofilms provide a shelter for the survival of persistence cells. Upon antibiotic treatment, the majority of cells are killed, leaving only persister cells that survive. The immune system is supposed to clear these cells, but fails to reach biofilm embedded cells that then may repopulate the biofilm. This is followed by the release of planktonic cells that can spread and cause symptoms (Lewis, 2007, 2010).

However, a direct role for TA systems in controlling the transformation of planktonic to biofilm cells has long been debated. The first TA system that was reported to play a direct role in biofilm formation involved the *E.coli* b3022 gene product, later renamed as MqsR which forms a type II TA system with MqsA (Kasari et al., 2010; Ren et al., 2004). This system cross-talks between the AI-2 dependent quorum-sensing system and the regulation of

motility (Barrios et al., 2006). MqsA represses also the expression of the general stress response regulators RpoS and CsgD, leading to reduced formation of the signaling nucleotide c-di-GMP, curli production and overall reduction of biofilm formation (Soo and Wood, 2013). Upon degradation of MqsA by Lon protease, triggered by oxidative stress, a switch to the biofilm mode of growth is made (Wang et al., 2013).

It is suggested that programmed cell death could be an altruistic mechanism to allow a small subpopulation of biofilm cells to survive by releasing essential nutrients in the biofilm community where diffusion is limited. As already mentioned, the TA module mostly linked to PCD *in E.coli* is MazEF. Indeed, KO mutation of *mazEF*, as well as of *yafQ*/*dinJ*, was reported in different studies to result in a prominent decrease in biofilm formation (Kim et al. 2009, Kolodkin-Gal et al. 2009). It is suggested that expression of *mazEF* is a stress response to cellular damage and results in a reversible bacteriostasis (Bayles, 2013). However, if the cellular damage, e.g. DNA damage, enters a certain, yet poorly understood, threshold, the cell enters into PCD. This is followed by apoptotic-like processes, for example the activation of holins, leading ultimately to cell lysis (Rice et al., 2007)

Also other TA systems are often found differentially regulated when performing transcriptomic or proteomic comparisons between planktonic and biofilm cells. For example, several TA systems are upregulated in biofilm cells in *Treponema denticola*, an oral spirochaete strongly associated with chronic periodontitis (Mitchell et al., 2010). In the HipA-HipB TA system, interruption of the HipA gene by transposon mutagenesis could decrease biofilm formation in *Shewanella oneidensis* MR-1 (Theunissen et al., 2010). This was confirmed in *E.coli* where it was related to the release of extracellular DNA (eDNA) which is essential as an adhesion molecule for the biofilm formation (Zhao et al., 2013).

1.6 Toxin-Antitoxin systems and virulence

TA systems are not only involved in general bacterial physiological processes such as the general stress response, persistence and biofilm formation but also have a direct impact on the pathogenicity of bacteria. It was recently discovered that a surprising correlation exists between the number of TA modules in the genome and the virulence capacity of bacteria (Georgiades and Raoult, 2011). TA systems are omnipresent on mobile genetic elements that have been found in many pathogenic bacteria and that are involved in antibiotic resistance and virulence as pathogenicity islands (Ma et al., 2013). One of the main functions of TA systems is to stabilize and maintain the plasmid population, such as the mvpA-mvpt (VapBC) on a *Shigella flexneri* virulence plasmid Pmysh6000 (Sayeed et al., 2000) and HigBA in *Proteus vulgaris* (Hurley and Woychik, 2009). Deletion of VapBC homologues in non-typeable *Haemophilus influenzae* leads to strong reduction of virulence in tissue and animal models for otitis media (Ren et al., 2012). A remarkable example of the ambiguous effect of antibiotic treatment involves VapC in *Ricketsia* : VapC is released by the bacterium upon chloramphenicol treatment and causes host cell apoptosis, presumably via its RNase activity (Audoly et al., 2011).

Type II TA systems with toxin endoribonuclease activity are involved in the regulation of virulence factors expression. The *S. aureus* MazF recognizes specific pentad sequences which are overrepresented in proteins involved in virulence. More specifically, it is involved in the degradation of the mRNA for SraP, a protein directly involved in host-pathogen interaction (Zhu et al., 2009). Expression of this protein is growth phase dependent, presumably under control of the MazE antitoxin. A similar role for tuning the production of virulence factors was found in *Clostridium difficile* (Rothenbacher et al., 2012). Remarkably,

a MazEF TA system has been demonstrated on plasmids obtained from all vancomycin resistant enterococci (VRE) (Moritz and Hergenrother, 2007).

The 3 component TA system ω-ε-ζ system was discovered from a clinically isolated gram positive pathogen *Streptococcus pyogenes* plasmid pSM19035, later it was also identified on other pathogenic gram positive bacteria's multiple resistance plasmids such as found in *Enterococcus faecalis* and *E. faecium* (Moritz and Hergenrother, 2007). The ω-ε-ζ was reported to contribute to their virulence and proved to stabilize the plasmid involved in resistance to erythromycin and lincomycins (Mutschler and Meinhart, 2011).

Some type I toxins are described to be directly involved in host cell lysis. The *S. aureus* Pepa1 toxin is a pore-forming peptide that can cause bacterial cell death. It is typically released from its SprA antitoxin upon oxidative stress for example after internalization of the *Staphylococcus* in the phagolysosomes. This is suggested to be an example of altruistic behavior, as the peptide can also drive erythrocyte lysis, resulting in release of slowly dividing cells that thus can escape the immune system (Sayed et al., 2012).

1.7 Toxin-Antitoxin systems and antimicrobial drug development

Due to the limited effectiveness of current drugs and the increasing multidrug tolerance and resistance, there is a growing need for novel antimicrobial agents. The targets of the TA toxins are mostly involved in biological processes such as DNA replication, mRNA synthesis, macromolecular synthesis, cell wall synthesis, phage infection and cytoskeletal polymerization. There is a remarkable overlap between these and antibiotics targets (Table 1.1, Figure 1.3). This, together with involvement of TA systems in programmed cell death indicates that investigation of TA systems can provide alternative antimicrobial strategies

(Park et al., 2013). Structural biology efforts to study TA systems can direct antibiotics discovery both by rational drug design (Mandal et al., 2009) or prodrug screening. It should be realized that the redundancy of TA systems and other pathways to persistence formation form a serious bottleneck to directly use them as drug targets (Lewis, 2012, 2013). However, understanding their mode of action could lead to the discovery of better suited downstream processes as was recently shown by the use of acyldepsipeptides to continuously activate ClpP, even in the absence of ATP generating processes in dormant cells (Conlon et al., 2013).

Breaking the balance of toxin-antitoxin levels offers the possibility to induce suicide in bacteria. Derived from CcdB, which targets DNA gyrase resulting in DNA fragmentation and cell death in *E.coli*, a peptide was synthesized that inhibits the DNA gyrase and other topoisomerases killing the bacteria (Trovatti et al., 2008).

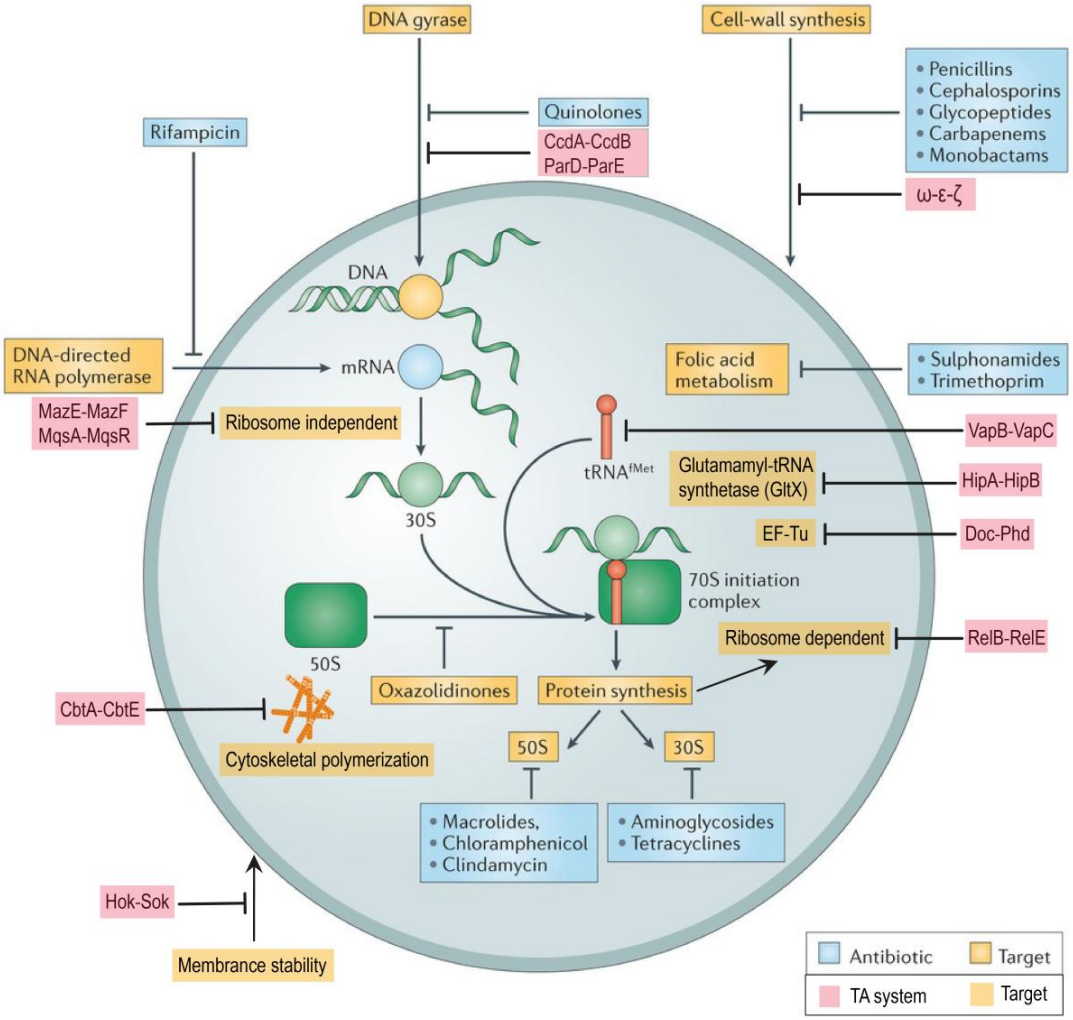


Figure 1.3 Targets of antibiotics and Targets of Toxins. The procedures involved in antibiotic killing and the growth arrest strategy of the toxins are highly similar. Thus the investigation of Toxin-Antitoxin systems could lead to the antimicrobial drug discovery. Figure modified and adapted from Kim Lewis Nature Reviews Drug Discovery (Lewis, 2013).

Table 1.1 List of molecular functions in which Toxin-Antitoxin systems are involved. The remarkable synergy between targets of toxins and antibiotics is illustrated by some key examples.

Targets of toxins			**Targets of antibiotics**		
Involved cellular process	**Target**	**Toxin-Antitoxin systems**	**Involved cellular process**	**Target**	**Antibiotic class**
DNA replication	DNA gyrase	CcdA-CcdB ParD-ParE	DNA replication	DNA gyrase	Quinolones
tRNA synthesis	Glutamyl tRNA synthetase (GltX)	HipA-HipB	mRNA synthesis	RNA polymerase	Rifampicin
Macromolecular synthesis	Ribosome dependent	RelB-RelE	Macromolecular synthesis	50S ribosome	Macrolides, Chloramphenicol, Clindamycin
	EF-Tu	Doc-Phd			
	$tRNA^{fMet}$	VapB-VapC			
	Ribosome independent	MazE-MazF MqsA-MqsR		30S ribosome	Aminoglycosides, Tetracyclines
Cell-wall synthesis	UDP-N-acetyl-glucosamine (UNAG)	ω-ε-ζ	Cell-wall synthesis	PBP Peptidoglycan Cell membrane	Penicillins, Cephalosporins, Glycopeptides, Carbapenems, Monobactames
Membrane disruption	Inner membrane	Sok-Hok TisB-IstR	Folic-acid metabolism		Sulphonamides, Trimethoprim
Phage infections		ToxN-ToxI			
Cytoskeletal polymerization	MreB and FtsZ	CbtA-CbtE			

In the ω-ε-ζ TA system, the toxin ζ was shown to phosphorylate the peptidoglycan precursor UNAG, further resulting in the competitive inhibition of MurA which plays an important role in the peptidoglycan synthesis (Mutschler et al., 2011). The enzymatic product of toxin ζ, UNAG-3P was described to be a potential lead agent for a new broad spectrum antibiotic (Mutschler and Meinhart, 2011).

More TA system applications deal with the endoribonucleases type of toxins other than involved in the antibacterial drug discovery include antiviral gene therapy and recombinant protein technology (Chono et al., 2011; Shapira et al., 2012; Suzuki et al., 2005). Antiviral therapy is based on the fact that some endoribonucleases, e.g. MazF also induce eukaryotic cell death. Bringing *mazF* under an HIV specific regulator provokes cell death of infected cells (Chono et al., 2011). Alternatively, MazF was introduced in a so-called zymoxin, in a fusion of rationally designed antitoxin peptides, a viral protease cleavage site. Upon infection, the viral protease will release the MazF protein (Shapira et al., 2011, 2012).

1.8 Conclusion and Perspectives

After more than two decades of investigation in TA systems, their involvement in a broad range of microbial phenomena is obvious. Understanding the role of the bacterial TA systems gave dramatic insight to understand the molecular mechanisms of bacterial stress response, and how persistence and biofilm formation are induced. TA systems control therefore mechanisms that contribute largely to the failure of the current antibiotics. The investigation of TA systems starts to raise examples of how the study of fundamental mechanisms evolves to biotechnological and medical applications such as antimicrobial drug discovery. Nevertheless, in the pace with more and more characterized and uncharacterized TA systems were discovered and their molecular mechanisms will be explored, the detailed function and evolution of each TA system will be revealed. Definitely, the investigation of TA systems will lead and play a crucial role in the microbiology field.

1.9 References

Aakre, C.D., Phung, T.N., Huang, D., and Laub, M.T. (2013). A Bacterial Toxin Inhibits DNA Replication Elongation through a Direct Interaction with the β Sliding Clamp. Mol. Cell 52, 617-628.

Aizenman, E., Engelberg-Kulka, H., and Glaser, G. (1996). An *Escherichia coli* chromosomal "addiction module" regulated by 3',5'-bispyrophosphate: a model for programmed bacterial cell death. Proc. Natl. Acad. Sci. U. S. A. *93*, 6059–6063.

Amato, S.M., Orman, M. a, and Brynildsen, M.P. (2013). Metabolic control of persister formation in *Escherichia coli*. Mol. Cell *50*, 475–487.

Audoly, G., Vincentelli, R., Edouard, S., Georgiades, K., Mediannikov, O., Gimenez, G., Socolovschi, C., Mège, J.-L., Cambillau, C., and Raoult, D. (2011). Effect of rickettsial toxin VapC on its eukaryotic host. PLoS One *6*, e26528.

Balaban, N.Q., Merrin, J., Chait, R., Kowalik, L., and Leibler, S. (2004). Bacterial persistence as a phenotypic switch. Science *305*, 1622–1625.

Barrios, A.F.G., Zuo, R., Yang, L., Bentley, W.E., Thomas, K., Gonza, F., Hashimoto, Y., and Wood, T.K. (2006). Autoinducer 2 Controls Biofilm Formation in *Escherichia coli* through a Novel Motility Quorum-Sensing Regulator (MqsR , B3022). J. Bacteriol. *188*, 305–316.

Bayles, K.W. (2013). Bacterial programmed cell death: making sense of a paradox. Nat. Rev. Microbiol. *12*, 63–69.

Beveridge, T.J., Chang, I.S., Kim, B.H., Kim, S., Culley, D.E., Reed, S.B., Margaret, F., Saffarini, D.A., Hill, E.A., Shi, L., et al. (2009). Stochastic hybrid modeling of DNA replication across a complete genome. Proc. Natl. Acad. Sci. *106*, 9535–9535.

Bigger, J.W. (1944). Treatment of staphylococcal infections with penicillin. Lancet *244*, 497–500.

Bjarnsholt, T. (2013). The role of bacterial biofilms in chronic infections. APMIS. Suppl. 1–51.

Black, D.S., Kelly, A.J., Mardis, M.J., and Moyed, H.S. (1991). Structure and organization of hip, an operon that affects lethality due to inhibition of peptidoglycan or DNA synthesis. J. Bacteriol. *173*, 5732–5739.

Black, D.S., Irwin, B., and Moyed, H.S. (1994). Autoregulation of hip, an operon that affects lethality due to inhibition of peptidoglycan or DNA synthesis. J. Bacteriol. *176*, 4081–4091.

Blower, T.R., Pei, X.Y., Short, F.L., Fineran, P.C., Humphreys, D.P., Luisi, B.F., and Salmond, G.P.C. (2011). A processed noncoding RNA regulates an altruistic bacterial antiviral system. Nat. Struct. Mol. Biol. *18*, 185–190.

Blower, T.R., Short, F.L., Rao, F., Mizuguchi, K., Pei, X.Y., Fineran, P.C., Luisi, B.F., and Salmond, G.P.C. (2012). Identification and classification of bacterial Type III toxin-antitoxin systems encoded in chromosomal and plasmid genomes. Nucleic Acids Res. *40*, 6158–6173.

Brown, B.L., Lord, D.M., Grigoriu, S., Peti, W., and Page, R. (2013). The *Escherichia coli* toxin MqsR destabilizes the transcriptional repression complex formed between the antitoxin MqsA and the mqsRA operon promoter. J. Biol. Chem. *288*, 1286–1294.

Castro-Roa, D., Garcia-Pino, A., De Gieter, S., van Nuland, N. a J., Loris, R., and Zenkin, N. (2013). The Fic protein Doc uses an inverted substrate to phosphorylate and inactivate EF-Tu. Nat. Chem. Biol. doi:10.1038/nchembio.1364.

Cataudella, I., Trusina, A., Sneppen, K., Gerdes, K., and Mitarai, N. (2012). Conditional cooperativity in toxin-antitoxin regulation prevents random toxin activation and promotes fast translational recovery. Nucleic Acids Res. *40*, 6424–6434.

Chono, H., Matsumoto, K., Tsuda, H., Saito, N., Lee, K., Kim, S., Shibata, H., Ageyama, N., Terao, K., Yasutomi, Y., et al. (2011). Acquisition of HIV-1 resistance in T lymphocytes using an ACA-specific E. coli mRNA interferase. Hum. Gene Ther. *22*, 35–43.

Conlon, B.P., Nakayasu, E.S., Fleck, L.E., LaFleur, M.D., Isabella, V.M., Coleman, K., Leonard, S.N., Smith, R.D., Adkins, J.N., and Lewis, K. (2013). Activated ClpP kills persisters and eradicates a chronic biofilm infection. Nature *503*, 365–370.

Cook, G.M., Robson, J.R., Frampton, R. a, McKenzie, J., Przybilski, R., Fineran, P.C., and Arcus, V.L. (2013). Ribonucleases in bacterial toxin-antitoxin systems. Biochim. Biophys. Acta *1829*, 523–531.

Correia, F.F., D'Onofrio, A., Rejtar, T., Li, L., Karger, B.L., Makarova, K., Koonin, E. V, and Lewis, K. (2006). Kinase activity of overexpressed HipA is required for growth arrest and multidrug tolerance in *Escherichia coli*. J. Bacteriol. *188*, 8360–8367.

El-Naggar, M.Y., Wanger, G., Leung, K.M., Yuzvinsky, T.D., Southam, G., Yang, J., Lau, W.M., Nealson, K.H., and Gorby, Y. a (2010). Electrical transport along bacterial nanowires from *Shewanella oneidensis* MR-1. Proc. Natl. Acad. Sci. U. S. A. *107*, 18127–18131.

Engelberg-Kulka, H., Amitai, S., Kolodkin-Gal, I., and Hazan, R. (2006). Bacterial programmed cell death and multicellular behavior in bacteria. PLoS Genet. *2*, e135.

Faridani, O.R., Nikravesh, A., Pandey, D.P., Gerdes, K., and Good, L. (2006). Competitive inhibition of natural antisense Sok-RNA interactions activates Hok-mediated cell killing in *Escherichia coli*. Nucleic Acids Res. *34*, 5915–5922.

Fasani, R. a, and Savageau, M. a (2013). Molecular mechanisms of multiple toxin-antitoxin systems are coordinated to govern the persister phenotype. Proc. Natl. Acad. Sci. U. S. A. *110*, E2528–37.

Fineran, P.C., Blower, T.R., Foulds, I.J., Humphreys, D.P., Lilley, K.S., and Salmond, G.P.C. (2009). The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. Proc. Natl. Acad. Sci. U. S. A. *106*, 894–899.

Fozo, E.M., Hemm, M.R., and Storz, G. (2008). Small Toxic Proteins and the Antisense RNAs That Repress Them. Microbiol. Mol. Biol. Rev. *72*, 579–589.

Fozo, E.M., Makarova, K.S., Shabalina, S. a, Yutin, N., Koonin, E. V, and Storz, G. (2010). Abundance of type I toxin-antitoxin systems in bacteria: searches for new candidates and discovery of novel families. Nucleic Acids Res. *38*, 3743–3759.

Gelens, L., Hill, L., Vandervelde, A., Danckaert, J., and Loris, R. (2013). A General Model for Toxin-Antitoxin Module Dynamics Can Explain Persister Cell Formation in E. coli. PLoS Comput. Biol. *9*, e1003190.

Georgiades, K., and Raoult, D. (2011). Genomes of the most dangerous epidemic bacteria have a virulence repertoire characterized by fewer genes but more toxin-antitoxin modules. PLoS One *6*, e17962.

Gerdes, K., and Wagner, E.G.H. (2007). RNA antitoxins. Curr. Opin. Microbiol. *10*, 117–124.

Gerdes, K., Larsen, J.E., and Molin, S. (1985). Stable inheritance of plasmid R1 requires two different loci. J. Bacteriol. *161*, 292–298.

Gerdes, K., Rasmussen, P.B., and Molin, S. (1986a). Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. Proc. Natl. Acad. Sci. U. S. A. *83*, 3116–3120.

Gerdes, K., Bech, F.W., Jorgensen, S.T., Lobner-olesen, A., Rasmussen, P.B., Atlung, T., Boe, L., Karlstrom, O., Molin, S., and Von Meyenburg, K. (1986b). Mechanism of postsegregational killing by the hok gene product of the parB system of plasmid R1 and its homology with the relF gene product of the E. coli relB operon. EMBO J. *5*, 2023–2029.

Gerdes, K., Poulsen, L.K., Thisted, T., Nielsen, A.K., Martinussen, J., and Andreasen, P.H. (1990). The hok killer gene family in gram-negative bacteria. New Biol. *2*, 946–956.

Gerdes, K., Nielsen, a, Thorsted, P., and Wagner, E.G. (1992). Mechanism of killer gene activation. Antisense RNA-dependent RNase III cleavage ensures rapid turn-over of the stable hok, srnB and pndA effector messenger RNAs. J. Mol. Biol. *226*, 637–649.

Germain, E., Castro-Roa, D., Zenkin, N., and Gerdes, K. (2013). Molecular Mechanism of Bacterial Persistence by HipA. Mol. Cell http://dx.doi.org/10.1016/j.molcel.2013.08.045.

Hall-Stoodley, L., and Stoodley, P. (2002). Developmental regulation of microbial biofilms. Curr. Opin. Biotechnol. *13*, 228–233.

Hayes, F. (2003). Toxins-antitoxins: plasmid maintenance, programmed cell death, and cell cycle arrest. Science *301*, 1496–1499.

Heidelberg, J.F., Paulsen, I.T., Nelson, K.E., Gaidos, E.J., Nelson, W.C., Read, T.D., Eisen, J. a, Seshadri, R., Ward, N., Methe, B., et al. (2002). Genome sequence of the dissimilatory metal ion-reducing bacterium Shewanella oneidensis. Nat. Biotechnol. *20*, 1118–1123.

Hobby, G.L., Meyer, K., and Chaffee, E. (1942). Observations on the Mechanism of Action of Penicillin. Exp. Biol. Med. *50* , 281–285.

Hurley, J.M., and Woychik, N. a (2009). Bacterial toxin HigB associates with ribosomes and mediates translation-dependent mRNA cleavage at A-rich sites. J. Biol. Chem. *284*, 18605–18613.

Kasari, V., Kurg, K., Margus, T., Tenson, T., and Kaldalu, N. (2010). The *Escherichia coli* mqsR and ygiT genes encode a new toxin-antitoxin pair. J. Bacteriol. *192*, 2908–2919.

Kawano, M., Aravind, L., and Storz, G. (2007). An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin. Mol. Microbiol. *64*, 738–754.

Keren, I., Shah, D., Spoering, A., Kaldalu, N., and Lewis, K. (2004). Specialized Persister Cells and the Mechanism of Multidrug Tolerance in *Escherichia coli*. J. Bacteriol. *186*, 8172–8180.

Kolodkin-Gal, I., Hazan, R., Gaathon, A., Carmeli, S., and Engelberg-Kulka, H. (2007). A linear pentapeptide is a quorum-sensing factor required for mazEF-mediated cell death in *Escherichia coli*. Science *318*, 652–655.

Korch, S.B., Henderson, T. a., and Hill, T.M. (2003). Characterization of the hipA7 allele of Escherichia coli and evidence that high persistence is governed by (p)ppGpp synthesis. Mol. Microbiol. *50*, 1199–1213.

Kwan, B.W., Valenta, J. a, Benedik, M.J., and Wood, T.K. (2013). Arrested protein synthesis increases persister-like cell formation. Antimicrob. Agents Chemother. *57*, 1468–1473.

Leplae, R., Geeraerts, D., Hallez, R., Guglielmini, J., Drèze, P., and Van Melderen, L. (2011). Diversity of bacterial type II toxin-antitoxin systems: a comprehensive search and functional analysis of novel families. Nucleic Acids Res. *39*, 5513–5525.

Lewis, K. (2007). Persister cells, dormancy and infectious disease. Nat. Rev. Microbiol. *5*, 48–56.

Lewis, K. (2010). Persister cells. Annu. Rev. Microbiol. *64*, 357–372.

Lewis, K. (2012). Recover the lost art of drug discovery. Nature *485*, 439–440.

Lewis, K. (2013). Platforms for antibiotic discovery. Nat. Rev. Drug Discov. *12*, 371–387.

Ma, Z., Geng, J., Yi, L., Xu, B., Jia, R., Li, Y., Meng, Q., Fan, H., and Hu, S. (2013). Insight into the specific virulence related genes and toxin-antitoxin virulent pathogenicity islands in swine streptococcosis pathogen *Streptococcus equi ssp*. *zooepidemicus* strain ATCC35246. BMC Genomics *14*, 377.

Maisonneuve, E., Shakespeare, L.J., Girke, M., and Gerdes, K. (2011). Bacterial persistence by RNA endonucleases. PNAS *108*, 13206–13211.

Maisonneuve, E., Castro-Camargo, M., and Gerdes, K. (2013). (p)ppGpp Controls Bacterial Persistence by Stochastic Induction of Toxin-Antitoxin Activity. Cell *154*, 1140–1150.

Mandal, S., Moudgil, M., and Mandal, S.K. (2009). Rational drug design. Eur. J. Pharmacol. *625*, 90–100.

Masuda, H., Tan, Q., Awano, N., Wu, K.-P., and Inouye, M. (2012). YeeU enhances the bundling of cytoskeletal polymers of MreB and FtsZ, antagonizing the CbtA (YeeV) toxicity in *Escherichia coli*. Mol. Microbiol. *84*, 979–989.

Masuda, Y., Miyakawa, K., Nishimura, Y., and Ohtsubo, E. (1993). chpA and chpB, *Escherichia coli* chromosomal homologs of the pem locus responsible for stable maintenance of plasmid R100. J. Bacteriol. *175*, 6850–6856.

Mitchell, H.L., Dashper, S.G., Catmull, D. V, Paolini, R. a, Cleal, S.M., Slakeski, N., Tan, K.H., and Reynolds, E.C. (2010). *Treponema denticola* biofilm-induced expression of a bacteriophage, toxin-antitoxin systems and transposases. Microbiology *156*, 774–788.

Moritz, E.M., and Hergenrother, P.J. (2007). Toxin-antitoxin systems are ubiquitous and plasmid-encoded in vancomycin-resistant enterococci. Proc. Natl. Acad. Sci. U. S. A. *104*, 311–316.

Moyed, H.S., and Bertrand, K.P. (1983). HipA , a Newly Recognized Gene of *Escherichia coli* K-12 That Affects Frequency of Persistence After Inhibition of Murein Synthesis. J. Bacteriol. *155*, 768–775.

Moyed, H.S., and Broderick, S.H. (1986). Molecular cloning and expression of hipA, a gene of *Escherichia coli* K-12 that affects frequency of persistence after inhibition of murein synthesis. J. Bacteriol. *166*, 399–403.

Mulcahy, L.R., Isabella, V.M., and Lewis, K. (2013). *Pseudomonas aeruginosa* Biofilms in Disease. Microb. Ecol. DOI 10.1007/s00248–013–0297–x.

Mutschler, H., and Meinhart, A. (2011). E/Z Systems: Their Role in Resistance, Virulence, and Their Potential for Antibiotic Development. J. Mol. Med. (Berl). *89*, 1183–1194.

Mutschler, H., Gebhardt, M., Shoeman, R.L., and Meinhart, A. (2011). A novel mechanism of programmed cell death in bacteria by toxin-antitoxin systems corrupts peptidoglycan synthesis. PLoS Biol. *9*, e1001033.

Ogura, T., and Hiraga, S. (1983). Mini-F plasmid genes that couple host cell division to plasmid proliferation. Proc. Natl. Acad. Sci. U. S. A. *80*, 4784–4788.

Park, S.J., Son, W.S., and Lee, B.-J. (2013). Structural overview of toxin-antitoxin systems in infectious bacteria: a target for developing antimicrobial agents. Biochim. Biophys. Acta *1834*, 1155–1167.

Pedersen, K., and Gerdes, K. (1999). Multiple hok genes on the chromosome of *Escherichia coli*. Mol. Microbiol. *32*, 1090–1102.

Ramage, H.R., Connolly, L.E., and Cox, J.S. (2009). Comprehensive functional analysis of *Mycobacterium tuberculosis* toxin-antitoxin systems: implications for pathogenesis, stress responses, and evolution. PLoS Genet. *5*, e1000767.

Ren, D., Bedzyk, L. a, Thomas, S.M., Ye, R.W., and Wood, T.K. (2004). Gene expression in *Escherichia coli* biofilms. Appl. Microbiol. Biotechnol. *64*, 515–524.

Ren, D., Walker, A.N., and Daines, D. a (2012). Toxin-antitoxin loci vapBC-1 and vapXD contribute to survival and virulence in nontypeable *Haemophilus influenzae*. BMC Microbiol. *12*, 263.

Rice, K.C., Mann, E.E., Endres, J.L., Weiss, E.C., Cassat, J.E., Smeltzer, M.S., and Bayles, K.W. (2007). The cidA murein hydrolase regulator contributes to DNA release and biofilm development in *Staphylococcus aureus*. Proc. Natl. Acad. Sci. U. S. A. *104*, 8113–8118.

Von Rosenvinge, E.C., O'May, G. a, Macfarlane, S., Macfarlane, G.T., and Shirtliff, M.E. (2013). Microbial biofilms and gastrointestinal diseases. Pathog. Dis. *67*, 25–38.

Rotem, E., Loinger, A., Ronin, I., Levin-Reisman, I., Gabay, C., Shoresh, N., Biham, O., and Balaban, N.Q. (2010). Regulation of phenotypic variability by a threshold-based mechanism underlies bacterial persistence. Proc. Natl. Acad. Sci. U. S. A. *107*, 12541–12546.

Rothenbacher, F.P., Suzuki, M., Hurley, J.M., Montville, T.J., Kirn, T.J., Ouyang, M., and Woychik, N. a (2012). *Clostridium difficile* MazF toxin exhibits selective, not global, mRNA cleavage. J. Bacteriol. *194*, 3464–3474.

Rybtke T. Morten, Jensen O. Peter , Hoiby Niels , Givskov Michael, T.T.-N. and B.T. (2011). The Implication of *Pseudomonas aeruginosa* Biofilms in Infections. In Inflammation & Allergy-Drug Targets, 10,. 141–157.

Sayed, N., Nonin-Lecomte, S., Réty, S., and Felden, B. (2012). Functional and structural insights of a *Staphylococcus aureus* apoptotic-like membrane peptide from a toxin-antitoxin module. J. Biol. Chem. *287*, 43454–43463.

Sayeed, S., Reaves, L., Radnedge, L., and Austin, S. (2000). The stability region of the large virulence plasmid of *Shigella flexneri* encodes an efficient postsegregational killing system. J. Bacteriol. *182*, 2416–2421.

Sberro, H., Leavitt, A., Kiro, R., Koh, E., Peleg, Y., Qimron, U., and Sorek, R. (2013). Discovery of Functional Toxin / Antitoxin Systems in Bacteria by Shotgun Cloning. Mol. Cell *50*, 136–148.

Scherrer, R., and Moyed, H.S. (1988). Conditional impairment of cell division and altered lethality in hipA mutants of *Escherichia coli* K-12. J. Bacteriol. *170*, 3321–3326.

Schumacher, M. a, Piro, K.M., Xu, W., Hansen, S., Lewis, K., and Brennan, R.G. (2009). Molecular mechanisms of HipA-mediated multidrug tolerance and its neutralization by HipB. Science *323*, 396–401.

Schumacher, M.A., Min, J., Link, T.M., Guan, Z., Xu, W., Ahn, Y.-H., Soderblom, E.J., Kurie, J.M., Evdokimov, A., Moseley, M.A., et al. (2012). Role of Unusual P Loop Ejection and Autophosphorylation in HipA-Mediated Persistence and Multidrug Tolerance. Cell Rep. *2*, 518–525.

Shao, Y., Harrison, E.M., Bi, D., Tai, C., He, X., Ou, H.-Y., Rajakumar, K., and Deng, Z. (2011). TADB: a web-based resource for Type 2 toxin-antitoxin loci in bacteria and archaea. Nucleic Acids Res. *39*, D606–11.

Shapira, A., Gal-Tanamy, M., Nahary, L., Litvak-Greenfeld, D., Zemel, R., Tur-Kaspa, R., and Benhar, I. (2011). Engineered toxins “zymoxins” are activated by the HCV NS3 protease by removal of an inhibitory protein domain. PLoS One *6*, e15916.

Shapira, A., Shapira, S., Gal-Tanamy, M., Zemel, R., Tur-Kaspa, R., and Benhar, I. (2012). Removal of hepatitis C virus-infected cells by a zymogenized bacterial toxin. PLoS One *7*, e32320.

Short, F.L., Pei, X.Y., Blower, T.R., Ong, S.-L., Fineran, P.C., Luisi, B.F., and Salmond, G.P.C. (2013). Selectivity and self-assembly in the control of a bacterial toxin by an antitoxic noncoding RNA pseudoknot. Proc. Natl. Acad. Sci. U. S. A. *110*, E241–9.

Soo, V.W.C., and Wood, T.K. (2013). Antitoxin MqsA Represses Curli Formation Through the Master Biofilm Regulator CsgD. Sci. Rep. *3*, 3186.

Stewart, P.S., and Franklin, M.J. (2008). Physiological heterogeneity in biofilms. Nat. Rev. Microbiol. *6*, 199–210.

Suzuki, M., Zhang, J., Liu, M., Woychik, N. a, and Inouye, M. (2005). Single protein production in living cells facilitated by an mRNA interferase. Mol. Cell *18*, 253–261.

Theunissen, S., De Smet, L., Dansercoer, A., Motte, B., Coenye, T., Van Beeumen, J.J., Devreese, B., Savvides, S.N., and Vergauwen, B. (2010). The 285 kDa Bap/RTX hybrid cell surface protein (SO4317) of *Shewanella oneidensis* MR-1 is a key mediator of biofilm formation. Res. Microbiol. *161*, 144–152.

Trovatti, E., Cotrim, C. a, Garrido, S.S., Barros, R.S., and Marchetto, R. (2008). Peptides based on CcdB protein as novel inhibitors of bacterial topoisomerases. Bioorg. Med. Chem. Lett. *18*, 6161–6164.

Van Melderen, L. (1996). ATP-dependent Degradation of CcdA by Lon Protease. EFFECTS OF SECONDARY STRUCTURE AND HETEROLOGOUS SUBUNIT INTERACTIONS. J. Biol. Chem. *271*, 27730–27738.

Van Melderen, L., and Bast, M.S. De (2009). Bacterial Toxin – Antitoxin Systems : More Than Selfish Entities ? PLoS Genet. *5*, e1000437.

Wang, X., and Wood, T.K. (2011). Toxin-antitoxin systems influence biofilm and persister cell formation and the general stress response. Appl. Environ. Microbiol. *77*, 5577–5583.

Wang, X., Lord, D.M., Cheng, H., Osbourne, D.O., Hoon, S., Sanchez-torres, V., Quiroga, C., Zheng, K., Herrmann, T., Peti, W., et al. (2013). A Novel Type V TA System Where mRNA for Toxin GhoT is Cleaved by Antitoxin GhoS. Nat Chem Biol *8*, 855–861.

Winther, K.S., and Gerdes, K. (2011). Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. Proc. Natl. Acad. Sci. U. S. A. *108*, 7403–7407.

Winther, K.S., Brodersen, D.E., Brown, A.K., and Gerdes, K. (2013). VapC20 of *Mycobacterium tuberculosis* cleaves the Sarcin-Ricin loop of 23S rRNA. Nat. Commun. *4*, 2796.

Wozniak, R. a F., and Waldor, M.K. (2009). A toxin-antitoxin system promotes the maintenance of an integrative conjugative element. PLoS Genet. *5*, e1000439.

Yamaguchi, Y., Park, J.-H., and Inouye, M. (2011). Toxin-antitoxin systems in bacteria and archaea. Annu. Rev. Genet. *45*, 61–79.

Zhang, Y., Zhang, J., Hoeflich, K.P., Ikura, M., Qing, G., and Inouye, M. (2003). MazF cleaves cellular mRNAs specifically at ACA to block protein synthesis in *Escherichia coli*. Mol. Cell *12*, 913–923.

Zhao, J., Wang, Q., Li, M., Heijstra, B.D., Wang, S., Liang, Q., and Qi, Q. (2013). *Escherichia coli* toxin gene hipA affects biofilm formation and DNA release. Microbiology *159*, 633–640.

Zhu, L., Inoue, K., Yoshizumi, S., Kobayashi, H., Zhang, Y., Ouyang, M., Kato, F., Sugai, M., and Inouye, M. (2009). *Staphylococcus aureus* MazF specifically cleaves a pentad sequence, UACAU, which is unusually abundant in the mRNA for pathogenic adhesive factor SraP. J. Bacteriol. *191*, 3248–3255.

Chapter 2

Introduction Section B

Mass Spectrometry in Integrative Structural Biology

2.1 Introduction

Since the discovery of the double helix structure of DNA in 1953, the whole genome of more than 10000 organisms was sequenced; including bacteria, archaea, eukaryote and also the DNA sequence of many viruses, phages, viroids, plasmids and organelles was completely determined (http://www.ncbi.nlm.nih.gov/genome). In 2003, as one of the major milestones, the human genome sequence was completed by the Human Genome Project (HGP), with efforts from laboratories all over the world (http://www.genome.gov/). The number of human protein coding genes is estimated to be between 20000 to 25000 (International Human Genome Sequencing Consortium, 2004; Lander et al., 2001; Venter et al., 2001). Compared to the number of genes of other organisms, the number of human genes surprisingly does not corroborate the relative complexity of human life. It has been recently revealed that the biological complexity is correlated with the Protein-Protein Interaction (PPI) network, rather than simply with the total number of proteins. The number of interactions in human cells is estimated to be around 650000, which is tenfold higher than for the fruit fly *Drosophila melanogaster* which has a similar number of genes as human (Duan, 2002; Hart et al., 2006; Stumpf et al., 2008). The Protein-protein Interaction Network (PIN) turns out to be relevant to understand the functional and the evolutionary properties of biological systems (Jin et al., 2013).

A living cell has a wide range of PPIs of structural and functional relevance. To understand the function of macromolecular assemblies, it is essential to explore the characteristics of the individual macromolecules, and then to study their stoichiometry, shapes, structures and finally their assemblies (Sharon, 2010). To achieve this, it is necessary to study subsequently from each component of the assembly the amino acid sequence, the secondary structure, the stoichiometry, the affinity and ultimately the high resolution 3-dimentional structure. Thanks to developments in the atomic resolution technologies such as X-ray crystallography and

Nuclear Magnetic Resonance (NMR), the number of proteins of which the three dimensional structure has been determined will reach a total number of 100000 in the year 2014 as can be found in the Protein Data Bank (PDB) (http://www.rcsb.org/pdb/home/home.do). However, compared to the number of genes in the GenBank database (http://www.ncbi.nlm.nih.gov/genbank/), there is still a difference of two orders of magnitude. Furthermore, over 85% of the structures deposited in PDB are monomers or homomers, whereas in a living cell proteins mostly function as oligomers and are involved in heteromeric complexes. By way of example, there are approximately 6 unique interactions per protein in *Saccharomyces cerevisiae* (Reguly et al., 2006), although the interactions partners are largely unknown. Unfortunately, PINs occur at relative low concentration within physiologically relevant samples. Especially the approaches applied to determine 3-dimentional structures such as X-ray crystallography, NMR spectroscopy or other techniques require large amounts of sample or long acquisition time and poorly address certain groups of proteins, e.g. membrane associated proteins. All those features of the traditional approaches make the determination of the macromolecular assemblies to be a major challenge and there are many hurdles on the way to get the job done in one step.

At present, a macromolecular study cannot be achieved by a single technique and thus requires the development of so-called "Integrative Structural Biology" (ISB), an approach which combines multi-source data information to achieve the determination of biologically relevant macromolecular assembly structures. Such multiple biophysical techniques mainly include X-ray crystallography, nuclear magnetic resonance (NMR), Small Angle X-ray Scattering (SAXS), Electron Microscopy (EM), Mass Spectrometry (MS), Atomic Force Microscopy (AFM), Förster Resonance Energy Transfer (FRET), and computational modeling (Figure 2.1) (Karaca and Bonvin, 2013; Ward et al., 2013).

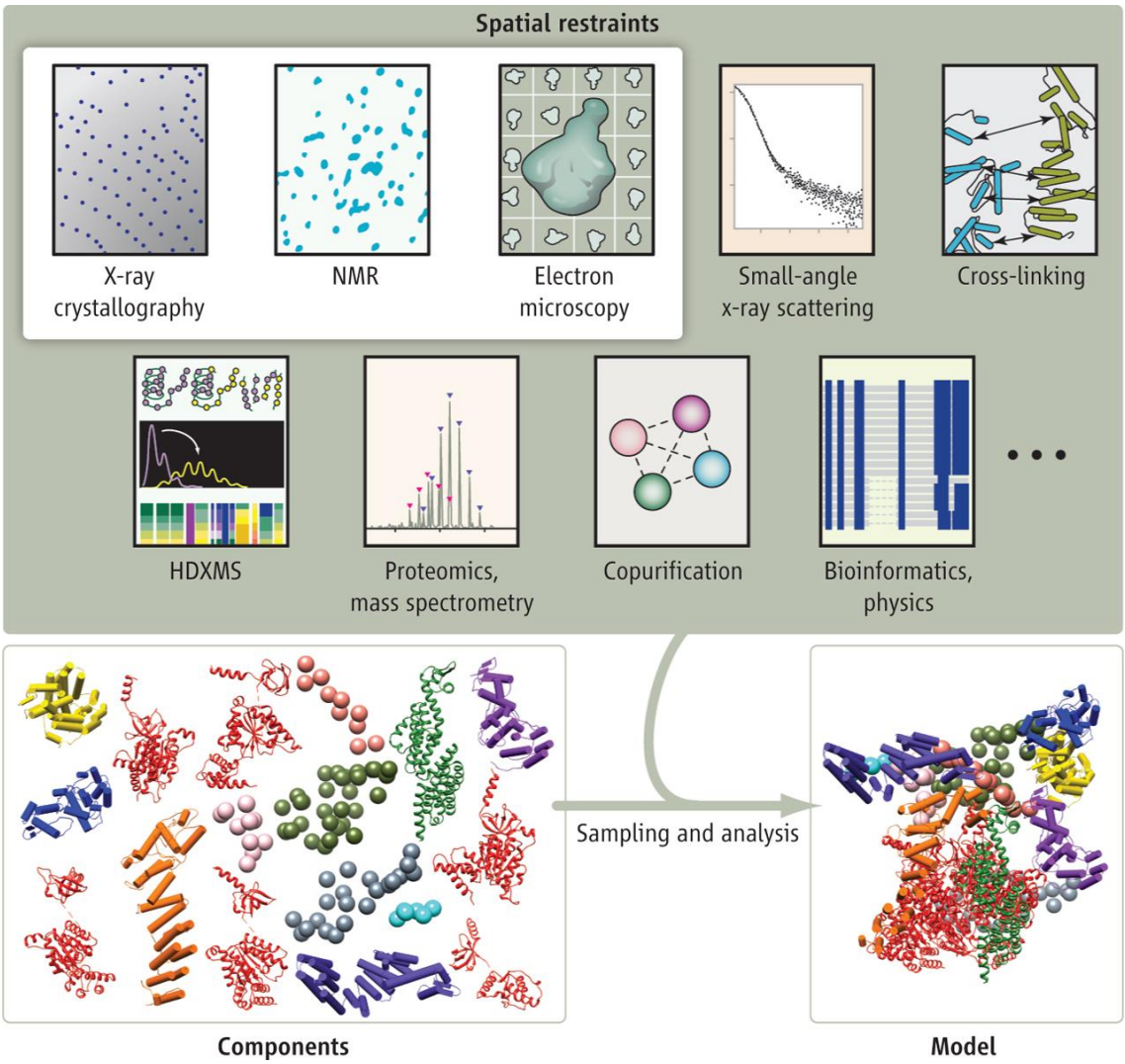


Figure 2.1 Principle of Integrative Structural Biology approaches. Models of macromolecules and their complexes can be constructed by combining different types of information generated by various experimental and theoretical techniques (gray box). The data are converted into spatial restraints, which are combined into a scoring function that guides sampling algorithms to obtain a detailed structural model. Adapted from (Ward et al., 2013)

The individual data source can generate information on the different (spatial) restraints of the biological macromolecular assembly. Hence the combination of all the restraints will lead to the final model through an integrative modeling approach.

In the last decade, the development of mass spectrometry applications in the field of Integrative Structural Biology allowed to establish the bridge between the partial high resolution structure of subunits of a complex and the macromolecular assembly of the multiple subunits. The discovery of electrospray ionization added the wings to the "molecular elephant" which makes the analysis possible of intact proteins and of their assembly with other proteins, DNA, RNA or small chemicals (Fenn, 2003). Together with MALDI, this soft ionization technique has contributed to the widespread use of mass spectrometry both in the academic and industrial fields.

Mass spectrometry, a technology that allows to determine the mass-to-charge (m/z) ratio of ionized molecules, has mainly played an essential role in Proteomics for its capacity to measure the masses of tryptic fragments of proteins with great accuracy and to subsequently identify the peptides by database searches (Aebersold and Mann, 2003). The quantitative nature of mass spectrometry allowed using this method to determine gene expression levels at the protein level. However, since its early days, electrospray ionization has also been used to establish information on protein-protein and protein-ligand interactions. Indeed, it was shown that non-covalent interactions can be maintained during ionization, allowing to determine the mass of the complex (Loo, 1997). Most recently, the development of high resolution mass spectrometry and its combination with ion mobility measurements, chemical crosslinking methods, hydrogen/deuterium exchange, affinity purification and computational processing have shed light on the structure of macromolecular assemblies. Here we summarize the principles of these technologies and their applications in the field of Integrative Structural Biology.

2.2 Mass Spectrometry and Ion Mobility

Ion Mobility spectrometry (IMS) is a technique that allows to separate ionized molecules based on their different mobility in a neutral carrier gas under a weak electric field (Kanu et al., 2008). IMS was first developed to be applied on small chemical molecules such as drugs and explosives. However, it can also be applied to large molecules such as intact proteins and macromolecular assemblies for use in the field of structural biology (Jurneczko and Barran, 2011; Konijnenberg et al., 2013). The MS instrument that is available for the IM measurement contains an ion mobility separation unit apart from the normal ion generation, ion selection and ion mass analysis MS components (Figure 2.2a). Different ion mobility results in a different drift time in the drift tube; compact ions will move faster than extended ones which will result in a shorter Arrival Time Distribution (ATD) (Figure 2.2b). From this drift time, the collision cross section (CCS) can be generated corroborating the shape of the ions (Equation 1) (Bohrer et al., 2008).

$$\Omega = \frac{(18\pi)^{1/2}}{16}\frac{ze}{(k_bT)^{1/2}}\left[\frac{1}{m_I}+\frac{1}{m_N}\right]^{1/2}\times\frac{760}{P}\frac{T}{273.2}\frac{1}{N}\frac{t_DE_D}{L}$$

Equation 1: z is the charge on the ion, k_b is the Boltzmann constant, and P and T are the pressure and the temperature. The mass of the ion and the buffer gas are also considered, where m_I represents the mass of the ion of interest and m_N represents the mass of the buffer gas. The final variables are the neutral number density of the buffer gas N, L is the length of the drift tube, t_D is the time the ion requires reaching the detector, and the cross-section of the ion itself Ω.

In combination with a Time-of-Flight analyzer, the IM-MS data generates a three dimensional spectrum, including the m/z ratio, the intensity and the drift time of ions (Figure 2.2c).

The combination of ESI and IM-MS can thus be used to analyze intact macromolecular assemblies, gaining insight into protein interactions networks and to architectural models (Robinson, 2012; Sharon and Robinson, 2007; Zhou and Robinson, 2010). Initially, the debate on protein stability in vacuum instead of in solution involving the contribution of hydrophobic forces was projected. Later, it was proven that van der Waals attractive forces can provide sufficient stability to retain complex assemblies in the gas phase, and that the loss of water molecules is compensated through the formation of hydrogen bonds between protein residues as suggested by molecular dynamics simulations (van der Spoel et al., 2011). The many successes in the ESI-IM-MS approach have shown that the native-like assembly of complexes is retained without disrupting the ternary macromolecular complex structure in the milli-second timescale when the complex is transferred into the gas phase (Marcoux and Robinson, 2013). Therefore, ESI-IM-MS is a very powerful approach to investigate macromolecular assemblies.

The ESI-IM-MS technique has been applied to solve the assembly not only of soluble protein complexes but also of membrane protein complexes which need to be stabilized by detergents. Initially, the technique was indeed used to determine the compositing units and the stoichiometry of soluble proteins and their ligand-binding properties (van Duijn et al., 2009; Gold et al., 2011; Zhou et al., 2008). Recently, ESI-IM-MS was successfully applied to investigate membrane proteins as huge and complicated as V-type ATPase with its lipid and nucleotide binding characteristics (Barrera and Robinson, 2011; Zhou et al., 2011), and an ABC transporter revealing its drug-binding properties and information on its conformational changes (Marcoux et al., 2013).

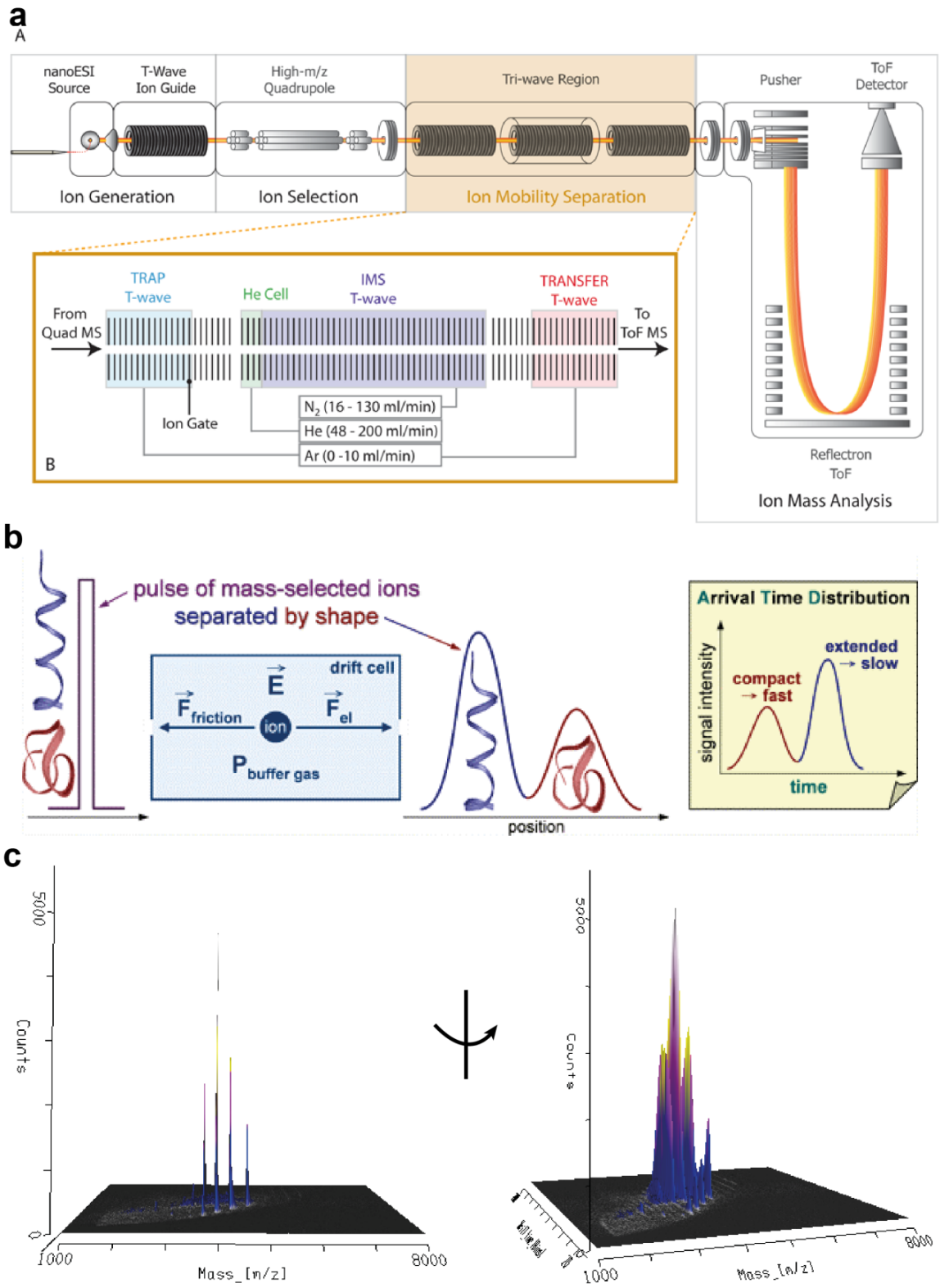

a
A
nanoESI Source
T-Wave Ion Guide
High-m/z Quadrupole
Tri-wave Region
Pusher
ToF Detector
Ion Generation
Ion Selection
Ion Mobility Separation
TRAP T-wave
He Cell
IMS T-wave
TRANSFER T-wave
From Quad MS
To ToF MS
Ion Gate
N2 (16 - 130 ml/min)
He (48 - 200 ml/min)
Ar (0 -10 ml/min)
B
Reflectron ToF
Ion Mass Analysis
b
pulse of mass-selected ions
separated by shape
drift cell
E
F friction
F el
ion
P buffer gas
position
Arrival Time Distribution
signal intensity
compact → fast
extended → slow
time
c
Counts
5000
0
1000
8000
Mass_[m/z]
Counts
5000
1000
8000
Mass_[m/z]

Figure 2.2 Ion Mobility Mass Spectrometry Theory. a. A schematic view of Quadrupole-Ion Mobility Time of Flight Mass Spectrometry instrument (Synapt G2, Waters) B enlargement view of the Ion Mobility Separation part. Adapted from (Zhong et al., 2011). b. Besides the separation through different m/z ratio, the ions can also be distinguished due to distinct arrival time distribution caused by their different shape in the buffer gas, thus generating a third dimension ion separation capacity. Adapted from Bowers Group Website. c. Example of 3-dimensional view of IM-MS separation generated from DriftScope (Waters).

Most recently, the CCS restraints generated from ESI-IM-MS were found to benefit modeling of the overall architecture of multiprotein complex assemblies at low resolution, and has also been shown to moderately correlate with high resolution X-ray analyses (Hall et al., 2012; Politis et al., 2010). Therefore, the ESI-IM-MS approach has shown the potential to generate low resolution models of macromolecular assemblies. To improve resolution and the false discovery rate (FDR) of structural models generated from the CCS restraints, one of the main challenges which remains to be solved is to develop more accurate and standard gas phase CCS calibration methods (Bush et al., 2010). Besides being useful in the study of protein complex assemblies, a nice confirmation of the potential of the ESI-IM-MS approach dealing with macromolecular assembly study is the successful investigation of the intact viruses and virus capsids assemblies which are as large as 18MDa using only minor modifications of commercially available MS instrument (Snijder et al., 2013; Uetrecht et al., 2008). With the breakthrough of high resolution and sensitive commercial MS instrumentations, the ESI-IM-MS technology will in future be more widely used in Integrative Structural Biology.

2.3 Mass Spectrometry and Chemical Crosslinking

Chemical crosslinkers are compounds that are used to covalently connect proteins that form otherwise complexes via weak non-covalent interactions that cannot be sustained during sample preparation and purification for biophysical analysis. It are mostly small chemical molecules having two reactive functional (bifunctional) groups (Sinz, 2003, 2006). These groups can either be homobifunctional or heterobifunctional (Figure 2.3).

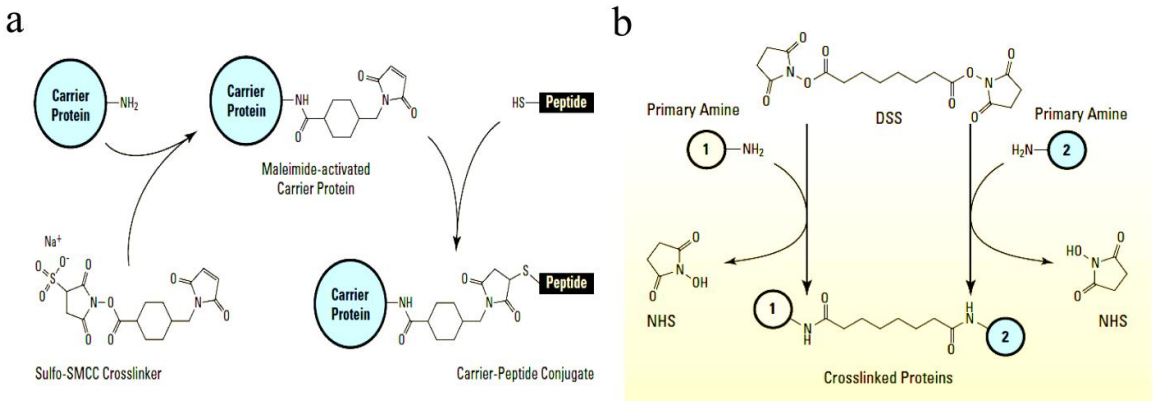


Figure 2.3 Examples of heterobifunctional and homobifuntional crosslinkers. a. The heterobifunctional crosslinker Sulfo-SMCC first labels primary amines with a cysteine reactive functional group, maleimide. Maleimide targets reduced sulfhydryls such as those found in the side chain of a cysteine residue. **b**. the homobiofunctional crosslinker DSS reacts with primary amine groups from both sides resulting in amine to amine conjugation. Adapted from Thermo Scientific Pierce.

There is a spacer between those two reactive groups which should correspond with the distance between the two residues to be crosslinked. The spacer can also be used as a reporter or label for further enrichment and identification of crosslinked products. Mostly, bifunctional crosslinkers can generate three types of crosslinked products: single reactive group crosslinked peptides (type 0); intra crosslinked peptides (type 1) and inter crosslinked peptides (type 2). The type 1 and type 2 crosslinked peptides are the most valuable ones since

they provide significant information on the distance between two different residues, which lead to restraints useful in structure determination and macromolecular architecture modeling (Paramelle et al., 2013; Zybailov et al., 2013). The reactivity and selectivity of a crosslinker are essential for the experimental success. At present, there are more than 100 crosslinkers reported, with most of them being commercially available. Those crosslinkers differ in reactive groups, in length of spacer and in having or not having a specific reporter, such that a widespread range of crosslink experiments can be carried out (Sinz, 2003, 2006).

Recent improvements in mass spectrometric technology have generated a renewed interest in chemical crosslink approaches. Indeed, after selection of suitable crosslinkers, a Top-down or Bottom-up approach can be applied to analyze crosslinked sample followed by high resolution MS analysis, a method referred to as XLMS The principle of the XLMS approach is described in (Figure 2.4). After crosslinking the proteins, mostly a proteolytic digestion is performed to cleave the protein in peptides, followed by LC-MS/MS to identify the cross-linked peptides. Enrichment of the crosslinked peptides is crucial to improve MS detection, since the crosslinked peptides mostly have very low abundance in the digestion mixtures. Strategies implying size exclusion chromatography (SEC), chemical resin pull down enrichment, or reporter affinity purification significantly increase the sensitivity of crosslinked peptides detection (Buncherd et al., 2012; Leitner et al., 2012a; Weisbrod et al., 2013).

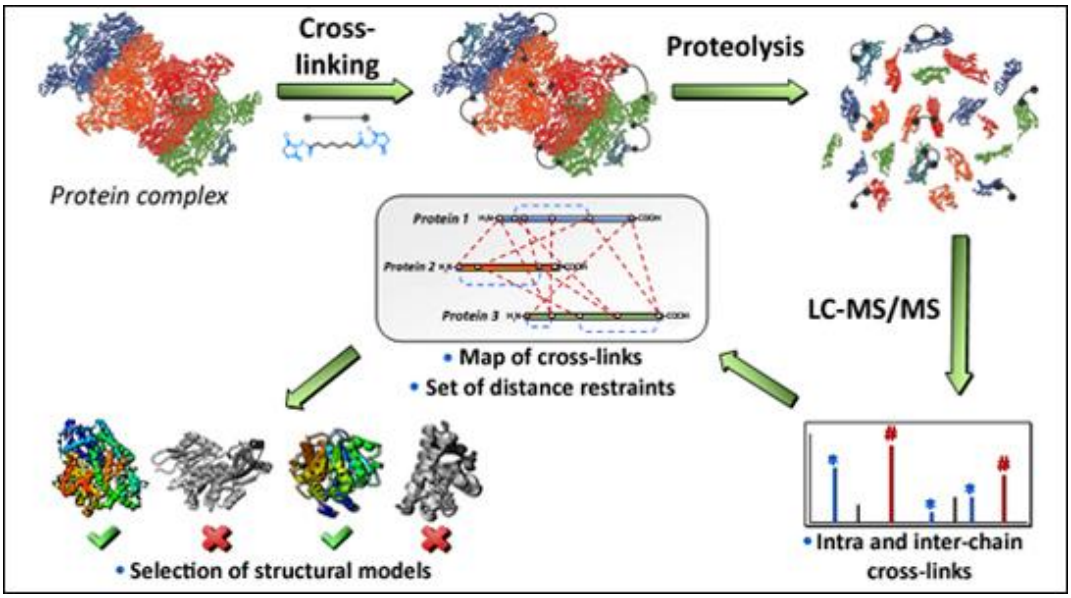


Figure 2.4 Chemical Crosslink Mass Spectrometry approach. A protein complex is crosslinked with selective crosslinker, followed by proteolysis prior to high resolution LCMS/MS analysis. MS data is interpreted and crosslinked peptides with distance restraints are generated for structural models selection. Adapted from Dalton Mass Spectrometry Lab.

Crosslinked peptide identification has always been a big challenge in the XL-MS investigations because a crosslink is essentially different from a classical post-translational modification (PTM). Initially, the crosslink identification algorithms could only be applied in small scale analyses, examples of such algorithms are: xComb (Panchaud et al., 2010), X!Link (Lee et al., 2007), MassMatrix (Xu et al., 2010). Besides the challenge to increase the amount of a MS-detectable crosslinked peptide, the accurate and efficient identification of crosslinks in large scale proteomes and the improvement of false discovery rate of this identification have become an essential problem (Tabb, 2012). Recently, the introduction of novel algorithms presented in software tools like Xquest, Xprophet (http://prottools.ethz.ch/orinner/public/htdocs/xquest/) (Walzthoeni et al., 2012) and Plink, Pfind (http://pfind.ict.ac.cn/software/pLink/index.html) (Yang et al., 2012) has dramatically

made the XLMS approach more accessible for routine applications, effectively improving the false discovery rate, using appropriate isotope labeled crosslinkers.

The XLMS approach is mostly been use to unravel the binary interactions in protein complexes, being combined with other biophysical techniques, such as X-ray crystallography, Electron Microscopy (EM) and Small Angle X-ray Scattering (SAXS) thus contributing to three dimensional model building (Rappsilber, 2011; Stengel et al., 2012). Recently, the distance restraints generated from XLMS were applied to improve the refinement of existing X-ray diffraction data of the eukaryotic chaperonin protein complex TRiC/CCT; significantly improving the atomic resolution model fitting, with 6% Rfree optimization (Leitner et al., 2012b). The combination of XLMS with affinity purification (AP) and Electron Microscopy (EM) was successful to determine the structural interaction network of human protein phosphatase 2A (PP2A) and construct the three dimensional model of the TCP1 ring complex (TRiC) chaperonin interacting with the PP2A regulatory subunit 2ABG (Herzog et al., 2012).

With both technological and informatics improvements described above, the XLMS approach extended from *in vitro* analyses to levels comparable to physiological relevant *in vivo* conditions, allowing to address the whole cell proteome interactome (Zheng et al., 2011). Most recently, Real-time Analysis for Crosslinked peptide Technology (ReACT) was developed which successfully identified 708 unique crosslinked peptide pairs with less than 5% false discovery rate in *E.coli* (Weisbrod et al., 2013). By further improvement of the XLMS approach, sample preparation techniques, design of new crosslinker molecules and advanced database search strategies, the ultimate goal of detecting proteome wide protein interaction networks and topological analyses are achievable. All the above mentioned breakthroughs in XLMS, including the enrichment of crosslinked peptides, higher mass accuracy and improved sensitivity of MS instrumentation, more accurate and efficient crosslink

identification algorithms for large scale proteomes, and the decrease in false discovery rate have leveled up the technique of XLMS in biological applications, especially in the field of integrative structural biology of macromolecular assemblies.

2.4 Mass Spectrometry and Hydrogen Deuterium Exchange

Since its introduction in the 1990s, Hydrogen Deuterium Exchange Mass Spectrometry (HDXMS) has been found to be a powerful approach to probe protein conformational dynamics and protein interactions (Katta et al., 1991; Katta and Chait, 1993). The method uses stable isotope labeling with deuterium to trace the rate at which the hydrogen atom of the protein backbone amide bond undergoes exchange. The hydrogen/deuterium exchange rate is highly dependent on the degree of hydrogen bonding, and therefore reflects the protein conformational mobility, hydrogen bond strength and solvent accessibility. HDXMS can thus provide information on protein conformational changes, and efficiently catch information on the thermodynamics and kinetics of protein unfolding at the moment of equilibrium (Jaswal, 2013).

Experimentally, the approach of HDXMS includes protein labeling, denaturation of the samples, protein fragmentation, LC-MS, and interpretation of the data in context of the structure (Figure 2.5) (Brock, 2012). Although there are quite some discussions about the method concerning aspects of reproducibility, digestion efficiency, protein size and resolution, HDXMS has been used and successfully applied in biopharmaceuticals studies, the development of formulation, and in protein interaction studies (Iacob and Engen, 2012). HDXMS application areas also include the study of excited states, redox states, catalytic transition states, substrate recognition, metastable states, assembly of macromolecular complexes, aggregation, enzyme activation and regulation, and in the interaction studies of protein with small molecules (Brock, 2012). Most recently, it was also involved in the

structure-function investigation of membrane proteins of the G protein coupled receptor (GPCR) family, and drug discovery studies (Katritch et al., 2013; Pacholarz et al., 2012). After two decades of development as routine analytical technique, HDXMS has been applied in a large number of applications, providing complementary information. However, although still a lot of features need to be explored such as the origin of some HDX signals, attained by other methods, the HDXMS will provide very promising and meaningful insights in protein assembly research.

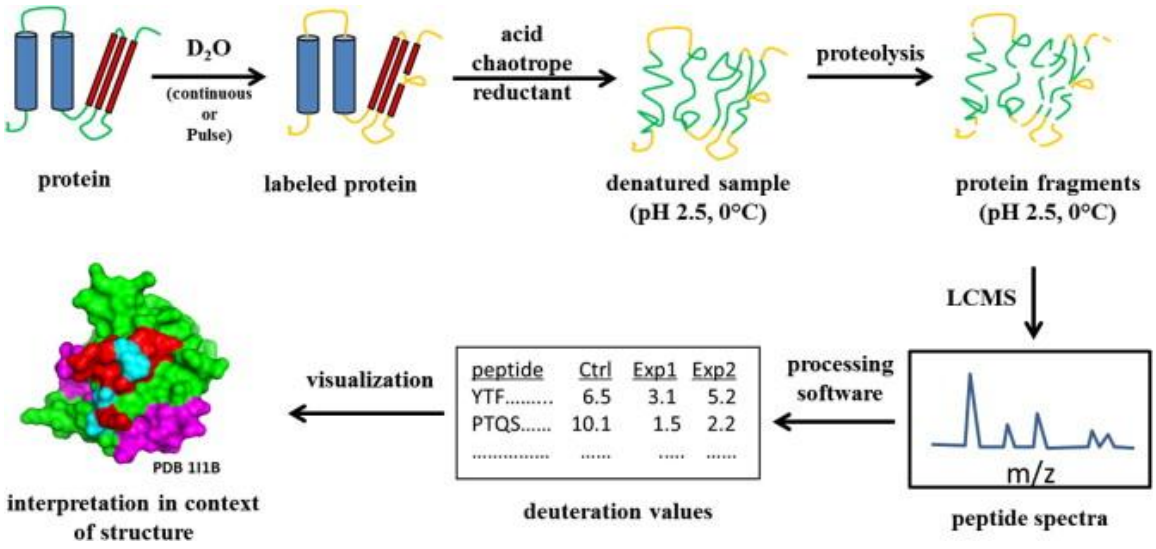


Figure 2.5 Hydrogen Deuterium Exchange Mass Spectrometry approach. Typical fragmentation HDXMS, data analysis, and interpretative workflow. Isotope exchange is initiated by dilution into D_2O. After a set period, the sample is quenched by dilution with denaturant and lowering of the pH and temperature. This is followed by proteolysis and LCMS. Deuteration levels are extracted from the mass spectra and interpretation of the data is performed in the context of available structures and other biophysical data. Adapted from (Brock, 2012)

Besides the techniques of IM, XL and HDX mass spectrometry discussed above, some other techniques have also been applied to study the integrative proteomics and to generate structural restraints. Such techniques include, for example, oxidative foot printing (Guan and

Chance, 2005) and Tandem Affinity Purification (TAP) (Gavin et al., 2002; Hyung and Ruotolo, 2012; Kaiser et al., 2008). Recently, the combination of a different fragmentation technique, Electron Transfer Dissociation (ETD) and a new generation commercial Q-IM-TOF mass spectrometry was applied for protein surface mapping (Lermyte et al., 2014). Those techniques definitely contribute to the understanding the composition, structure and dynamics of macromolecular assemblies.

2.5 Integrative Modeling

The development of the Integrative Structural Biology techniques goes hand in hand with the development of computational methods. Recently, several remarkable platforms were reported, such as the Integrative Modeling Platform package (IMP) (http://www.integrativemodeling.org/) which can deal with restraints from a variety of data including X-ray crystallography, EM, XLMS, IMMS, SAXS and proteomics to build up relevant models (Russel et al., 2012; Webb et al., 2011). The platform, for example, sheds light on the molecular architecture of 26S proteasome which provided further insights on the evolution and function of this enormous complex (Lasker et al., 2012). Another powerful macromolecular structure modeling software, Rosetta (https://www.rosettacommons.org/) was applied on the structure determination of bacterial type III secretion needle (Leaver-Fay et al., 2011; Loquet et al., 2012). The High Ambiguity Data-Driven Docking program (HADDOCK) (http://www.nmr.chem.uu.nl/haddock/) allows to directly apply docking onto generic multicomponent complexes (Karaca and Bonvin, 2013; de Vries et al., 2010). In short, the combination of integrative structural biology experimentation and integrative computational modeling can result in detailed understanding of the phenomenon of macromolecular assemblies; further can lead to a mechanistic understanding of cell function and to cellular tomograms.

2.6 Conclusion and Perspectives

In order to study the structure and function of individual proteins as well as their involvement in large multicomponent assemblies, Integrative Structural Biology is the most appropriate approach to generate data at different levels of resolution. The recent applications of IMMS enlighten the super assembly involved with β barrel membrane protein and intrinsically disordered protein interactions and of XLMS successful apply on the in vivo whole cell samples (Housden et al., 2013; Weisbrod et al., 2013). The development of instrumentation as well as of computational modeling tools in this field allows us to overcome challenges that exist in terms of sensitivity (analysis of protein complexes at physiologically relevant concentrations), coverage (e.g. membrane associated proteins) and speed, (transient interactions). Many challenges are still lying on the way to understand the detailed function, evolution and modulation of cellular machineries. Definitely the Integrative Structural Biology will play an essential role in those studies, and MS as an independent or complementary approach will play a crucial role as a bridge to integrate and model different kinds of data on protein interactions.

2.7 References

Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. Nature *422*, 198–207.

Barrera, N.P., and Robinson, C. V (2011). Advances in the mass spectrometry of membrane proteins: from individual proteins to intact complexes. Annu. Rev. Biochem. *80*, 247–271.

Bohrer, B.C., Merenbloom, S.I., Koeniger, S.L., Hilderbrand, A.E., and Clemmer, D.E. (2008). Biomolecule analysis by ion mobility spectrometry. Annu. Rev. Anal. Chem. (Palo Alto. Calif). *1*, 293–327.

Brock, A. (2012). Fragmentation hydrogen exchange mass spectrometry: a review of methodology and applications. Protein Expr. Purif. *84*, 19–37.

Buncherd, H., Nessen, M. a, Nouse, N., Stelder, S.K., Roseboom, W., Dekker, H.L., Arents, J.C., Smeenk, L.E., Wanner, M.J., van Maarseveen, J.H., et al. (2012). Selective enrichment and identification of cross-linked peptides to study 3-D structures of protein complexes by mass spectrometry. J. Proteomics *75*, 2205–2215.

Bush, M.F., Hall, Z., Giles, K., Hoyes, J., Robinson, C. V, and Ruotolo, B.T. (2010). Collision cross sections of proteins and their complexes: a calibration framework and database for gas-phase structural biology. Anal. Chem. *82*, 9557–9565.

De Vries, S.J., van Dijk, M., and Bonvin, A.M.J.J. (2010). The HADDOCK web server for data-driven biomolecular docking. Nat. Protoc. *5*, 883–897.

Duan, X.J. (2002). Describing Biological Protein Interactions in Terms of Protein States and State Transitions : THE LiveDIP DATABASE. Mol. Cell. Proteomics *1*, 104–116.

Fenn, J.B. (2003). Electrospray wings for molecular elephants (Nobel lecture). Angew. Chem. Int. Ed. Engl. *42*, 3871–3894.

Gavin, A.-C., Bösche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.-M., Cruciat, C.-M., et al. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature *415*, 141–147.

Gold, M.G., Stengel, F., Nygren, P.J., Weisbrod, C.R., Bruce, J.E., Robinson, C. V, Barford, D., and Scott, J.D. (2011). Architecture and dynamics of an A-kinase anchoring protein 79 (AKAP79) signaling complex. Proc. Natl. Acad. Sci. U. S. A. 1–9.

Guan, J.-Q., and Chance, M.R. (2005). Structural proteomics of macromolecular assemblies using oxidative footprinting and mass spectrometry. Trends Biochem. Sci. *30*, 583–592.

Hall, Z., Politis, A., and Robinson, C. V (2012). Structural modeling of heteromeric protein complexes from disassembly pathways and ion mobility-mass spectrometry. Structure *20*, 1596–1609.

Hart, G.T., Ramani, A.K., and Marcotte, E.M. (2006). How complete are current yeast and human protein-interaction networks? Genome Biol. *7*, 120.

Herzog, F., Kahraman, A., Boehringer, D., Mak, R., Bracher, A., Walzthoeni, T., Leitner, A., Beck, M., Hartl, F.-U., Ban, N., et al. (2012). Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. Science *337*, 1348–1352.

Housden, N.G., Hopper, J.T.S., Lukoyanova, N., Rodriguez-Larrea, D., Wojdyla, J. a, Klein, A., Kaminska, R., Bayley, H., Saibil, H.R., Robinson, C. V, et al. (2013). Intrinsically disordered protein threads through the bacterial outer-membrane porin OmpF. Science *340*, 1570–1574.

Hyung, S.-J., and Ruotolo, B.T. (2012). Integrating mass spectrometry of intact protein complexes into structural proteomics. Proteomics *12*, 1547–1564.

Iacob, R.E., and Engen, J.R. (2012). Hydrogen exchange mass spectrometry: are we out of the quicksand? J. Am. Soc. Mass Spectrom. *23*, 1003–1010.

International Human Genome Sequencing Consortium (2004). Finishing the euchromatic sequence of the human genome. Nature *431*, 931–945.

Jaswal, S.S. (2013). Biological insights from hydrogen exchange mass spectrometry. Biochim. Biophys. Acta *1834*, 1188–1201.

Jin, Y., Turaev, D., Weinmaier, T., Rattei, T., and Makse, H. a (2013). The evolutionary dynamics of protein-protein interaction networks inferred from the reconstruction of ancient networks. PLoS One *8*, e58134.

Jurneczko, E., and Barran, P.E. (2011). How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. Analyst *136*, 20–28.

Kaiser, P., Meierhofer, D., Wang, X., and Huang, L. (2008). Tandem Affinity Purification Combined with Mass Spectrometry to Identify Components of Protein Complexe. Methods Mol Biol *439*, 1–16.

Kanu, A.B., Dwivedi, P., Tam, M., Matz, L., and Jr, H.H.H. (2008). Ion mobility – mass spectrometry. J. Mass Spectrom. *43*, 1–22.

Karaca, E., and Bonvin, A.M.J.J. (2013). Advances in integrative modeling of biomolecular complexes. Methods *59*, 372–381.

Katritch, V., Cherezov, V., and Stevens, R.C. (2013). Structure-function of the G protein-coupled receptor superfamily. Annu. Rev. Pharmacol. Toxicol. *53*, 531–556.

Katta, V., and Chait, B.T. (1993). Hydrogen / Deuterium Exchange Electrospray Ionization Mass Spectrometry : A Method for Probing Protein Conformational Changes in Solution. J. Am. Chem. Soc. 4–8.

Katta, V., Chait, B.T., and Carr, S. (1991). Conformational changes in proteins probed by hydrogen-exchange electrospray-ionization mass spectrometry. Rapid Commun. Mass Spectrom. *5*, 214–217.

Konijnenberg, a, Butterer, a, and Sobott, F. (2013). Native ion mobility-mass spectrometry and related methods in structural biology. Biochim. Biophys. Acta *1834*, 1239–1256.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

Lasker, K., Förster, F., Bohn, S., Walzthoeni, T., Villa, E., Unverdorben, P., Beck, F., Aebersold, R., Sali, A., and Baumeister, W. (2012). Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. Proc. Natl. Acad. Sci. U. S. A. *109*, 1380–1387.

Leaver-Fay, A., Tyka, M., Lewis, S.M., Lange, O.F., Thompson, J., Jacak, R., Kaufman, K.W., Renfrew, P.D., Smith, C.A., Sheffler, W., et al. (2011). Chapter nineteen - Rosetta3: An Object-Oriented Software Suite for the Simulation and Design of Macromolecules. In Computer Methods, Part C, M.L.J. and L.B.B.T.-M. in Enzymology, ed. (Academic Press), pp. 545–574.

Lee, Y.J., Lackner, L.L., Nunnari, J.M., and Phinney, B.S. (2007). Shotgun Cross-Linking Analysis for Studying Quaternary and Tertiary Protein Structures research articles. J. Proteome Res. *6*, 3908–3917.

Leitner, A., Reischl, R., Walzthoeni, T., Herzog, F., Bohn, S., Förster, F., and Aebersold, R. (2012a). Expanding the chemical cross-linking toolbox by the use of multiple proteases and enrichment by size exclusion chromatography. Mol. Cell. Proteomics *11*, M111.014126.

Leitner, A., Joachimiak, L. a, Bracher, A., Mönkemeyer, L., Walzthoeni, T., Chen, B., Pechmann, S., Holmes, S., Cong, Y., Ma, B., et al. (2012b). The molecular architecture of the eukaryotic chaperonin TRiC/CCT. Structure *20*, 814–825.

Lermyte, F., Konijnenberg, A., Williams, J.P., Brown, J.M., Valkenborg, D., and Sobott, F. (2014). ETD Allows for Native Surface Mapping of a 150 kDa Noncovalent Complex on a Commercial Q-TWIMS-TOF Instrument. J. Am. Soc. Mass Spectrom. 10.1007/s13361–013–0798–3.

Loo, J. (1997). Studying noncovalent protein complexes by electrospray ionization mass spectrometry. Mass Spectrom. Rev. *16*, 1–23.

Loquet, A., Sgourakis, N.G., Gupta, R., Giller, K., Riedel, D., Goosmann, C., Griesinger, C., Kolbe, M., Baker, D., Becker, S., et al. (2012). Atomic model of the type III secretion system needle. Nature *486*, 276–279.

Marcoux, J., and Robinson, C. (2013). Twenty Years of Gas Phase Structural Biology. Structure *21*, 1541–1550.

Marcoux, J., Wang, S.C., Politis, A., Reading, E., Ma, J., Biggin, P.C., Zhou, M., Tao, H., Zhang, Q., Chang, G., et al. (2013). Mass spectrometry reveals synergistic effects of nucleotides, lipids, and drugs binding to a multidrug resistance efflux pump. Proc. Natl. Acad. Sci. U. S. A. *110*, 9704–9709.

Pacholarz, K.J., Garlish, R.A., Taylor, R.J., and Barran, P.E. (2012). Mass spectrometry based tools to investigate protein-ligand interactions for drug discovery. Chem. Soc. Rev. *41*, 4335–4355.

Panchaud, A., Singh, P., Shaffer, S.A., and Goodlett, D.R. (2010). xComb : A Cross-Linked Peptide Database Approach to Protein - Protein Interaction Analysis research articles. J. Proteome Res. *9*, 2508–2515.

Paramelle, D., Miralles, G., Subra, G., and Martinez, J. (2013). Chemical cross-linkers for protein structure studies by mass spectrometry. Proteomics *13*, 438–456.

Politis, A., Park, A.Y., Hyung, S.-J., Barsky, D., Ruotolo, B.T., and Robinson, C. V (2010). Integrating ion mobility mass spectrometry with molecular modelling to determine the architecture of multiprotein complexes. PLoS One *5*, e12080.

Rappsilber, J. (2011). The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. J. Struct. Biol. *173*, 530–540.

Reguly, T., Breitkreutz, A., Boucher, L., Breitkreutz, B., Hon, G.C., Myers, C.L., Parsons, A., Friesen, H., Oughtred, R., Tong, A., et al. (2006). Comprehensive curation and analysis of global interaction networks in Saccharomyces cerevisiae.

Robinson, C. V (2012). Finding the right balance: a personal journey from individual proteins to membrane-embedded motors: based on a lecture delivered at the 36th FEBS Congress in Torino, Italy, June 2011. FEBS J. *279*, 663–677.

Russel, D., Lasker, K., Webb, B., Velázquez-Muriel, J., Tjioe, E., Schneidman-Duhovny, D., Peterson, B., and Sali, A. (2012). Putting the pieces together: integrative modeling platform software for structure determination of macromolecular assemblies. PLoS Biol. *10*, e1001244.

Sharon, M. (2010). How far can we go with structural mass spectrometry of protein complexes? J. Am. Soc. Mass Spectrom. *21*, 487–500.

Sharon, M., and Robinson, C. V (2007). The role of mass spectrometry in structure elucidation of dynamic protein complexes. Annu. Rev. Biochem. *76*, 167–193.

Sinz, A. (2003). Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes. J. Mass Spectrom. *38*, 1225–1237.

Sinz, A. (2006). Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. Mass Spectrom. Rev. *25*, 663–682.

Snijder, J., Rose, R.J., Veesler, D., Johnson, J.E., and Heck, A.J.R. (2013). Studying 18 MDa virus assemblies with native mass spectrometry. Angew. Chem. Int. Ed. Engl. *52*, 4020–4023.

Stengel, F., Aebersold, R., and Robinson, C. V (2012). Joining forces: integrating proteomics and cross-linking with the mass spectrometry of intact complexes. Mol. Cell. Proteomics *11*, R111.014027.

Stumpf, M.P.H., Thorne, T., Silva, E. De, Stewart, R., An, H.J., and Lappe, M. (2008). Estimating the size of the human interactome MATHEMATICS. Proc. Natl. Acad. Sci. *105*, 6959–6964.

Tabb, D.L. (2012). Evaluating protein interactions through cross-linking mass spectrometry. Nat. Methods *9*, 879–881.

Uetrecht, C., Versluis, C., Watts, N.R., Roos, W.H., Wuite, G.J.L., Wingfield, P.T., Steven, A.C., and Heck, A.J.R. (2008). High-resolution mass spectrometry of viral assemblies: molecular composition and stability of dimorphic hepatitis B virus capsids. Proc. Natl. Acad. Sci. U. S. A. *105*, 9216–9220.

Van der Spoel, D., Marklund, E.G., Larsson, D.S.D., and Caleman, C. (2011). Proteins, lipids, and water in the gas phase. Macromol. Biosci. *11*, 50–59.

Van Duijn, E., Barendregt, A., Synowsky, S., Versluis, C., and Heck, A.J.R. (2009). Chaperonin complexes monitored by ion mobility mass spectrometry. J. Am. Chem. Soc. *131*, 1452–1459.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C. a, Holt, R. a, et al. (2001). The sequence of the human genome. Science *291*, 1304–1351.

Walzthoeni, T., Claassen, M., Leitner, A., Herzog, F., Bohn, S., Förster, F., Beck, M., and Aebersold, R. (2012). False discovery rate estimation for cross-linked peptides identified by mass spectrometry. Nat. Methods *9*, 901–903.

Ward, A.B., Sali, A., and Wilson, I. a (2013). Biochemistry. Integrative structural biology. Science *339*, 913–915.

Webb, B., Lasker, K., Schneidman-Duhovny, D., Tjioe, E., Phillips, J., Kim, S., Velázquez-Muriel, J., Russel, D., and Sali, A. (2011). Modeling of Proteins and Their Assemblies with the Integrative Modeling Platform. In Network Biology SE - 19, G. Cagney, and A. Emili, eds. (Humana Press), pp. 377–397.

Weisbrod, C.R., Chavez, J.D., Eng, J.K., Yang, L., Zheng, C., and Bruce, J.E. (2013). In Vivo Protein Interaction Network Identi fi ed with a Novel Real-Time Cross-Linked Peptide Identi fi cation Strategy. J. Proteome Res. *12*, 1569–1579.

Xu, H., Hsu, P., Zhang, O.L., Tsai, M., and Freitas, M.A. (2010). Database Search Algorithm for Identification of Intact Cross-Links in Proteins and Peptides Using Tandem Mass Spectrometry research articles. J. Proteome Res. *9*, 3384–3393.

Yang, B., Wu, Y.-J., Zhu, M., Fan, S.-B., Lin, J., Zhang, K., Li, S., Chi, H., Li, Y.-X., Chen, H.-F., et al. (2012). Identification of cross-linked peptides from complex samples. Nat. Methods *9*, 904–906.

Zheng, C., Yang, L., Hoopmann, M.R., Eng, J.K., Tang, X., Weisbrod, C.R., and Bruce, J.E. (2011). Cross-linking measurements of in vivo protein complex topologies. Mol. Cell. Proteomics *10*, M110.006841.

Zhong, Y., Hyung, S.-J., and Ruotolo, B.T. (2011). Characterizing the resolution and accuracy of a second-generation traveling-wave ion mobility separator for biomolecular ions. Analyst *136*, 3534–3541.

Zhou, M., and Robinson, C. V (2010). When proteomics meets structural biology. Trends Biochem. Sci. *35*, 522–529.

Zhou, M., Sandercock, A.M., Fraser, C.S., Ridlova, G., Stephens, E., Schenauer, M.R., Yokoi-Fong, T., Barsky, D., Leary, J. a, Hershey, J.W., et al. (2008). Mass spectrometry reveals modularity and a complete subunit interaction map of the eukaryotic translation factor eIF3. Proc. Natl. Acad. Sci. U. S. A. *105*, 18139–18144.

Zhou, M., Morgner, N., Barrera, N.P., Politis, A., Isaacson, S.C., Matak-Vinković, D., Murata, T., Bernal, R. a, Stock, D., and Robinson, C. V (2011). Mass spectrometry of intact V-type ATPases reveals bound lipids and the effects of nucleotide binding. Science *334*, 380–385.

Zybailov, B.L., Glazko, G. V, Jaiswal, M., and Raney, K.D. (2013). Large Scale Chemical Cross-linking Mass Spectrometry Perspectives. Proteomics Bioinforma. *S2*, 1–11.

Chapter 3

Structural Insights into HipA-HipB Toxin-Antitoxin System in *Shewanella oneidensis* MR-1

The contents of this chapter adapted from a manuscript:

Yurong Wen, Ester Behiels, Jan Felix, Jonathan Elegheert, Bjorn Vergauwen, Bart Devreese*, Savvas Savvides*. Structural Insights into Toxin-Antitoxin System HipA-HipB in *Shewanella oneidensis* MR-1.Under submission.

3.1 Prologue: Structure and function of the HipA-HipB toxin-antitoxin system in *E.coli*

HipA is recognized for its crucial role in persistence and multidrug tolerance. It was firstly discovered by Harris Moyed during a target screening in the search of genes linked to a high frequency of persister cells in 1983 (Moyed and Bertrand, 1983). Compared to the wild type *E.coli* K-12, A 1000 times higher frequency of persister formation upon treatment of ampicillin was observed after mutation of the *hip* locus (*hipA7* allele). This mutation reduces the lethality of selective inhibition of peptidoglycan synthesis and results in a severe growth defect at low temperature, hence persistence was linked to cold sensitivity (Scherrer and Moyed, 1988). Only until 2003, the *hipA7* was sequenced and a double mutation, G22S and D291A, was revealed compared to the wild type *hipA* locus (Korch et al., 2003). Subsequent experiments indicated that the single mutation D291A was sufficient for the high persistence phenotype in *E.coli* and the ectopic induction of *hipA* could inhibit DNA replication, transcription and translation (Korch and Hill, 2006). Black et al. initially described that the *hipA* and the upstream neighbor gene *hipB* form an operon and Falla et al. firstly classified the *hipA*-*hipB* operon as a toxin antitoxin module. Indeed, *hipA* expression is toxic to the cell in the absence of HipB. HipB was further demonstrated to be a Cro-like DNA binding protein using a classical helix-turn-helix DNA binding motif and (Black et al., 1991, 1994; Falla and Chopra, 1998). HipA was further identified to have a eukaryotic serine/threonine kinase-like folding with closest similarity to the PI 3/4 kinase superfamily through sequence comparative analysis. HipA can be autophosphorylated at the residue Ser150 in the presence of ATP and Mg^{2+} in vitro. Mutation of the potential ATP and Mg^{2+}-binding site residues D309 or D332 revealed the importance of this phosphorylation in the HipA function since it resulted in the defect of cellular growth inhibition upon HipA overexpression (Correia et al., 2006).

A ground break understanding of the HipA-HipB molecular mechanism came from resolving the atomic crystal structure of *E.coli* ternary complex of HipA-HipB with its operator DNA sequence (Schumacher et al., 2009). HipB does not directly counteract the HipA toxin function by itself. It rather acts by the formation of a heterotetramer HipA-$HipB_2$-HipA sandwich complex where the HipA molecules use their N- and C-terminus to bind the 2 HipB molecules from the dimer respectively. This complex was thought to further bind with the palindromic operator DNA sequence of the *hipA-hipB* operon thus neutralizing the HipA toxicity by suppressing its expression (Schumacher et al., 2009). Recently, crosstalks between multiple toxin-antitoxin systems and also in the HipA-HipB module were denoted; the transcription repressor HipB was revealed not only to negatively regulate its own operators from the HipA-HipB module but it also binds to as many as 39 other putative palindromic sequences. For example, this results in the negative regulation of RelA expression (Kasari et al., 2013; Lin et al., 2013).

The structural analysis of HipA further revealed its kinase activity. HipA binds with ATP and Mg^{2+} and autophosphorylates at Ser150. The structural data revealed that this autophosphorylation was through interaction of two HipA molecules and that it uses an unusual phosphorylation loop (pLoop) ejection mechanism (Schumacher et al., 2009, 2012). The phosphorylation of HipA could serve for the autoregulation mechanism and deactivates the HipA kinase activity (Schumacher et al., 2012). In the work of Schumacher et al, the essential elongation factor Tu (EFTu) was first reported as the target of HipA kinase; HipA would then inhibit macromolecular synthesis by interrupting the elongation process via phosphorylating EFTu (Schumacher et al., 2009). However, this could not be confirmed by other groups and most recently, EFTu was confirmed not to be the actual substrate of HipA. A genetic high throughput screening by overexpressing approximately 4120 genes in *E.coli* indicated that glutamyl tRNA synthetase (GltX) was the only gene of which overexpression

neutralizes the toxicity of HipA (Germain et al., 2013a). The cold sensitivity phenotype of *hipA7* was also reported to be suppressed by the overexpression of GltX (Kaspy et al., 2013a) and further experiments revealed that HipA can phosphorylate the GltX at the Ser239 which is near the active center and can inhibit aminoacylation (Germain et al., 2013a). The phosphorylation of GltX could result in the accumulation of uncharged $tRNA^{Glu}$ of which binding at the A site of the ribosome may lead to the activation of the stringent response. Indeed, synthesis of (p)ppGpp through the RelA mediated pathway was proposed to be the mechanism to generate the type I persistence (Bokinsky et al., 2013a; Germain et al., 2013b; Kaspy et al., 2013a).

The threshold phenomenon of Toxin-antitoxin systems like HipA-HipB is believed to play a role in the co-existence of dormant (persisters) cells in a bacterial populations. It was indicated that in other toxin-antitoxin systems, mainly type II such as Doc-Phd or RelA-RelB, conditional cooperativity is essential to keep a storage pool of the toxin to generate the type I persistence (Cataudella et al., 2013; Gelens et al., 2013). Conditional cooperativity means that suppression of toxin production by binding of the toxin antitoxin complex to their operator will only occur when the toxins are in the proper stoichiometry compared to the antitoxins level. Conditional cooperativity was reported to stabilize the antitoxin level in rapid growing cells, preventing the random activation of toxin and promotes fast translational recovery (Cataudella et al., 2012). However, for the HipA-HipB module in the persister cell formation, conditional cooperativity was not demonstrated and the simple threshold models were still far from satisfactory because it ignores several aspects concerning the key HipA-HipB autoregulation pathway such as the autophosphorylation of HipA and the reversion growth arrest phenomenon of phosphorylated HipA (Feng et al., 2014; Rotem et al., 2010). Hence, regarding the molecular mechanism of HipA-HipB as a regulator of persister formation, there are still far more questions to answer, such as: what is role of HipA

autophosphorylation involved in the complex formation between HipA-HipB and its operator DNA, how the cell manages to keep a storage toxin HipA pool generate persister cells without conditional cooperativity like the other type II toxin antitoxin systems, or is there another mechanism of possible conditional cooperativity in the HipA-HipB module.

3.2 Abstract

Nearly all bacteria exhibit a type of phenotypic growth described as persistence that is thought to underlie antibiotic resistance, infection relapse, and recalcitrant chronic conditions. The chromosomally encoded high-persistence (Hip) antitoxin/toxin proteins HipB and HipA from *Shewanella oneidensis*, an opportunistic pathogen with unusual respiratory capacities, constitute a type II toxin-antitoxin protein couple. Here we show that phosphorylated HipA can engage in an unexpected ternary complex with HipB and double stranded operator DNA that is distinct from the prototypical counterpart complex from E.coli. The structure of HipB in complex with operator DNA reveals a flexible C-terminus that is sequestered by HipA in the ternary complex, indicative of its role in binding HipA to abolish its function in persistence. The structure of a HipA mutant devoid of phosphorylation capacity in complex with a non-hydrolysable ATP analogue uncovers that HipA autophosphorylation is coupled to an unusual conformational collapse of its pLoop. However, HipA is unable to phosphorylate the translation factor Elongation factor Tu (EFTu), contrary to previous reports but in agreement with more recent findings. Our studies suggest that the phosphorylation state of HipA is an important factor in persistence and that the structural and mechanistic diversity of HipAB modules as regulatory factors in bacterial persistence is broader than previously thought.

3.3 Introduction

To cope with unpredictable environmental fluctuations bacteria resorted to population diversity as a risk-spreading strategy. By creating a phenotypic or genetically heterogeneous population bacteria essentially increase the probability that at least a subpopulation is well adapted to the actual environment (Acar et al., 2008; Avery, 2006; Hersh et al., 2004). A striking example of such phenotypic diversity is the emergence of persister cells as a small subpopulation of cells able to survive against an antibiotic challenge by entering a dormant state with increased tolerance against a wide variety of stresses (Lewis, 2010). Recently, a tight correlation between multidrug tolerant (MDT) persister cells and biofilm formation has been shown, further the crucial role in investigation of persistence and biofilm formation can lead to the understanding of recalcitrant infection disease and antimicrobial drug discovery (Balaban et al., 2013; Lewis, 2013). While bacterial persistence was discovered almost 70 years ago, its understanding at the molecular level is a relatively recent development and has coincided with the discovery of the *hipA* gene in *Escherichia coli* (Bigger, 1944; Moyed and Bertrand, 1983; Moyed and Broderick, 1986). The product of *hipA* and its upstream linked gene *hipB* form a toxin antitoxin module, which together are recognized as a major factor involved in persistence (Black et al., 1994). The toxin HipA is a kinase and its activity is essential for growth arrest and multidrug tolerance in *E.coli* K12 (Correia et al., 2006). A first functional annotation of HipA by Schumacher et al. had linked HipA to phosphorylation of elongation factor Tu (EFTu) resulting in inhibition of protein biosynthesis. However, more recent studies have demonstrated that EFTu is not the phosphorylation target of HipA but glutamyl-tRNA synthetase (GltX) instead (Germain et al., 2013a). Furthermore, HipA has been shown to trigger growth arrest by controlling RelA-mediated biosynthesis of the alarmone (p)ppGpp (Bokinsky et al., 2013b; Korch et al., 2003). The function of HipB has been much clearer in that it utilizes its DNA-binding capacity to bind to operators upstream

of the *hipAB* operon to neutralize HipA in the nucleolus, thereby serving as an antitoxin. Formation of a HipAB complex suppresses the transcription and translation of HipA (Schumacher et al., 2009, 2012). However, the detailed mechanism of growth arrest and growth arrest rescue has remained unclear to go along with lack of knowledge of how bacteria emerge from the dormancy and how they repopulate the growth phase.

We previously demonstrated that interruption of the SO0706 gene involved in *Shewanella oneidensis* MR-1 resulted in a 50% decrease in biofilm formation (Theunissen et al., 2010) This gene encodes a protein sharing 28% sequence identity with *E.coli* HipA. *S.oneidensis* MR-1 is a respiratory diverse model bacterium that attracted interest for its role in biogeochemistry, its potential for bioremediation of heavy metals and aromatic compounds as well as for its use in microbial fuel cells (Gorby et al., 2006; Heidelberg et al., 2002). Many of these applications rely on the potential of the organism to form biofilms on inorganic surfaces. Molecular studies on the mechanisms of biofilm formation by this organism established a role for a type IV pilus system, the presence of extracellular DNA and the mxdABCD gene cluster, encoding a polysaccharide biogenesis system (Thormann et al., 2004). A microarray analysis of the early stage of biofilm formation revealed the involvement of different metabolic processes and cyclic di-GMP dependent pathways which was recently confirmed by molecular characterization of PdeB, a c-di-GMP phosphodiesterase (Chao et al., 2013; Thormann et al., 2006) . The sequencing of the *S.oneidensis* MR-1 genome (Heidelberg et al., 2002) inspired efforts to gain insight into the regulatory and metabolic network and to get systematic level understanding of its capacities. In general, research has mainly focused on the respiratory and redox potential of *S.oneidensis* and applications thereof in bioremediation and electromicrobiology. Although biofilm formation is an important aspect of the natural occurrence of *S.oneidensis* MR-1 and crucial for its biotechnological application, rather few efforts have been undertaken to understand the fundamentals of this

phenomenon in this organism. Given the strong correlation between biofilm formation and persistence, as well as the potential involvement of SO0706 in this process we further investigated the structure and function of this protein.

Here, we report the results of a broad range of biochemical, molecular and structural biology experiments on the SO0705:SO0706 operon encoding a HipA-HipB related toxin-antitoxin module. We provide the atomic structure of the ternary $HipAB_{so}$ protein complex with its double stranded operator DNA and of $HipA_{so}$ bound to an ATP analogue. Interestingly, in the $HipAB_{so}$:DNA crystal only the phosphorylated $HipA_{so}$ co-crystallized with $HipB_{so}$ and DNA form the ternary complex. The phosphorylation of $HipA_{so}$, which uses an unusual pLoop conformation collapse from buried to exposed, is thought to be the trigger for growth arrest. However our direct observation of phosphorylated $HipA_{so}$ in a ternary complex indicates that $HipB_{so}$ could rescue growth arrest by directly targeting the active phosphorylated $HipA_{so}$ population. Furthermore, C-terminus of $HipB_{so}$ undergoes dramatic conformational changes between the unbound form and when in complex with $HipA_{so}$, indicating a role of the proteolysis regulation mechanism. Furthermore, we find that *S.oneidensis* MR-1 $HipA_{so}$ is neither able to bind to nor can phosphorylate $EFTu_{so}$ in agreement with recent findings for the *E.coli* counterpart.

3.4 Material and Methods

3.4.1 Expression and purification of recombinant proteins and complexes thereof

Expression constructs for *S.oneidensis* MR-1 $HipA_{so}$ (SO0706) and $HipB_{so}$ (SO0705) were established based on primers MR01, MR02 for $HipA_{so}$ and MR03, MR04 for HipB (**Table S1**), and were cloned into the pET15b expression vector (Merck Millipore). The D306Q mutant of HipAso was generated via overlapping PCR using the flanking primer MR01 and internal primer MR06 in one reaction and the flanking primer MR02 and internal primer MR05 in another reaction. The PCR products were gel purified and used in a final overlap PCR round using the flanking primers MR01 and MR02. An expression construct for *S.oneidensis* MR-1 $EFTu_{so}$ (SO0229) was PCR amplified with primer MR07, MR08 from *S.oneidensis* MR-1 genomic DNA (Table S1) and was cloned into a TOPO pCR2.1 vector (Invitrogen). *S.oneidensis* MR-1 $HipA_{so}$ and $HipB_{so}$ were expressed in *E.coli* BL21 DE3 and $EFTu_{so}$ in *E.coli* C43. Bacterial cultures were inducing for 1 h ($HipA_{so}$) or 4h ($HipB_{so}$ and $EFTu_{so}$) with 1mM isopropyl β-D-1-thiogalactopyranoside (IPTG) (Duchefa Biochemie) and were grown to an OD600 of 0.6-0.7 in LB medium supplied with Carbenicillin (100µg/ml) at 37ºC. The cells were harvested by centrifugation at 7000g for 10 min at 4°C. After resuspending the cell pellets in 25 mM Na_2HPO4 pH 7.5, 500 mM NaCl buffer with EDTA free protease inhibitor, the cellular suspension was sonicated and the soluble fraction was clarified by centrifugation at 75000g for 30 min. The clarified lysate was loaded on Ni-NTA column (Qiagen) and pre-equilibrated with buffer A (25 mM Na2HPO4 pH 7.5, 500 mM NaCl, 10mM imidazole), washed with buffer B (25 mM Na2HPO4 pH 7.5, 500 mM NaCl, 50mM imidazole) and eluted with gradient buffer C (25 mM Na2HPO4 pH 7.5, 500 mM NaCl, 500mM imidazole). The recombinant proteins were concentrated and subsequently further purified by size exclusion chromatography on a Prep-Grade Hiload 16/60

Superdex200 column (GE Healthcare) equilibrated with buffer (10mM HEPES, 300mM NaCl and 5% glycerol pH7.5). All procedures for the overexpression and purification of $EFTu_{so}$ protein are the same as for except that we added 0.2mM GDP and 5mM $MgCl_2$ in the buffers used for purifications. The N-terminal 6xHis-tag in HipB was cleaved with biotinylated thrombin (Novagen) during overnight incubation and repurified with size exclusion chromatography. Pure fractions were pooled and proteins were concentrated for further experiments.

To assemble the $HipAB_{so}$ in complex with its operator DNA, purified $HipA_{so}$ and $HipB_{so}$ were mixed with double-stranded operator DNA obtained by annealing two single primers 5'-ATTAGGTGTACTTATCTACACTTTTT-3' and 5'-AAAAAGTGTAGATAAGTACACCTAAT-3' (IDT). The $HipAB_{so}$-DNA complex was purified by size exclusion chromatography.

3.4.2 Crystallization and X-ray data collection of $HipAB_{so}$: DNA complex, $HipA_{so}$-AMPPNP-Mg and $HipB_{so}$

All crystallization trial experiments were performed via the vapor diffusion method using a Mosquito robot (TTP LabTech). Crystals of $HipA_{so}$ in complex with $Mg2^{+}$-AMPPNP grew in two days 20 °C by mixing 200 nL purified recombinant $HipA_{so}$ with 2mM AMPPNP (Sigma) and 10mM $MgCl_2$ with an equal volume of reservoir solution (0.2M sodium chloride, 0.1M Na HEPES pH7.0, 25% Polyethylene glycol 4000) were equilibrated against 200 nL reservoir solution. For X-ray data collection under cryogenic conditions the crystals were harvested from the drop and frozen in a cryostream using crystallization condition supplemented with 20% glycerol as cryoprotection solvent. Crystals of the $HipA_{so}$-AMPPNP complex belonged to spacegroup $P2_12_12_1$ with a=60.70 Å, b=75.83 Å, c=110.32 Å and α=β=γ=90° and diffracted to 1.83 Å resolution. Crystals of recombinant $HipB_{so}$ were grown overnight from

purified $HipB_{so}$ after cleavage of the N-terminal 6xHis-tag. From 0.1M Na HEPES pH7.5, 10% v/v 2-propanol, 20% PEG 4000 and diffracted to 1.85 Å. The crystals belonged to space group $C222_1$ with a=53.46Å, b=72.95Å, c=86.91Å and α=β=γ=90°.

Crystals of recombinant $HipAB_{so}$ in complex with its operator DNA complex were grown in the crystallization geometries described above from reservoirs containing 0.2M calcium acetate hydrate pH7.5, 20% Polyethylene glycol 3350. For cryoprotection, the $HipAB_{so}$: DNA crystals were first incubated for several minutes in the crystallization condition supplemented with 20% glycerol. We identified two crystal forms that diffracted to 3.39 Å and 3.79 Å respectively. The first crystal form diffracted to 3.39Å belonged to spacegroup $P2_12_12_1$ with a=57.82 Å, b=122.45 Å, c=189.85 Å and α=β=γ= 90°. The second crystal form in spacegroup $P2_1$ with a= 72.24 Å, b= 57.33 Å, c= 171.38 Å and α=γ= 90°, β= 95.8° diffracted to 3.79 Å resolution. All X-ray diffraction data were processed with XDS and the XIA2 routine (Kabsch, 2010; Winter, 2010).

3.4.3 Crystal structure determination of $HipA_{so}$-AMPPNP-Mg^{2+}, $HipB_{so}$ and $HipAB_{so}$: DNA

The structure of $HipA_{so}$-AMPPNP-Mg^{2+} was determined via the molecular replacement platform BALBES (Long et al., 2008), using the *S.oneidensis* MR-1 $HipA_{so}$ sequence and the crystal structure of *E.coli* HipA (PDB: 3DNU) structure as input. The structure was refined extended to completion using Phenix.AutoMR, Phenix. Refine and Phenix.AutoBuild (Adams et al., 2010). The final HipA-AMPPNP structure contains one molecule of HipA per asymmetric unit cell including the all 433 residues from HipA and 5 residues from the affinity purification tag, 1 AMPPNP molecular, 2 Mg2+ ions, 1 Na+ ion, and 486 water molecules. The structure has excellent geometry which was indicated by the Ramachandran analysis in the X-ray data collection and refinement statistics table.

To determine the structure of $HipB_{so}$, the *S.oneidensis* MR-1 $HipB_{so}$ sequence and a subunit from the crystal structure of dimeric *E.coli* HipB (PDB: 3DNV) were used to create a truncated search model lacking loop regions and non-conserved side-chains using CCP4 Chainsaw. The structure was determined using molecular replacement protocols implemented in MOLREP and Phenix.Phaser and was refined using Phenix.Refine (Adams et al., 2010; Vagin and Teplyakov, 1997). The final $HipB_{so}$ structure, which contains 1 dimer $HipB_{so}$ in the crystal asymmetric unit including residues 24-87 for one subunit and residues 17-87 for the second subunit, 123 water molecules in the dataset of 1.85 Å resolution and including residues 17-87 for both subunits, 39 water molecules in the dataset of 2.35 Å resolution.

The $HipAB_{so}$:DNA complex structure was determined by molecular replacement protocols implemented in Phenix.AutoMR using the HipA-AMPPNP monomer and the $HipB_{so}$ dimer as search models. The limited diffraction of crystal forms 1 and 2 for the $HipAB_{so}$:DNA complex necessitated the use of weak data at the diffraction limit of the crystals as indicated by half-data set Pearson correlation coefficient (CC1/2) (Karplus and Diederichs, 2012), and the quality of the resultant electron-density difference maps. Both the orthorhombic and monoclinic crystals contain 2 monomers of $HipA_{so}$, 1 dimer of $HipB_{so}$ and double stranded DNA in the asymmetric unit cell. After initial rounds of refinement, clear electron density could be observed for the DNA duplex which allowed modeling of the complete DNA sequence in COOT (Emsley and Cowtan, 2004). The final $HipAB_{so}$: DNA structure contains 2 $HipA_{so}$ which were found to be the phosphorylated form, with remodeling of the phosphorylation loop including residues 7-433, 1 dimer $HipB_{so}$ including residues 20-87, 1 double stranded DNA with 24 and 25 nucleotides in each strand.

3.4.4 Nano Liquid Chromatography Mass Spectrometry

Freshly purified $HipA_{so}$ in 10mM ATP 20mM $MgCl_2$ and $HipA_{so}$ treated with lambda phosphatase overnight were diluted 1:50 with 50mM ammonium bicarbonate buffer. Sequencing-grade trypsin (Sigma) was added in 1:50 ratio to the diluted protein and digested overnight at 37 degrees. 1pmol HipA digestion sample from those 3 groups was applied for LCMS analysis on a nanoACQUITY UPLC coupled to a Synapt mass spectrometer used in the Q-TOF mode (Waters). A nanoACQUITY UPLC@2G-V/M Trap column 108µm*20mm 5µ symmetry C18 (Waters) was used as the trap column at 15 µL/min for 1min. A nanoACQUITY UPLC 1.8µm HSS T3 75µm*250mm column (Waters) was used as the separation column at 0.25 µL/min with a gradient as follows: 3%B to 40%B 0-30min, 40%B to 85%B 30-31min, 85%B 31-35 min, 85%B-3%B 35-35.5min, 3%B until 65min using 0.1 formic acid as buffer A and 100% Acetonitrile as buffer B. All the Nano LCMS data were processed with the PLGS 2.5 software (Waters).

3.4.5 Isothermal Titration Calorimetry

Experiments were carried out on a VP-ITC MicroCalorimeter at 25 degrees and data were analyzed using the origin ITC analysis software package (MicroCal). Purified HipA, HipB and DNA were matched into the buffer containing 25mM Na_2HPO4, 300mM NaCl and 5% glycerol at pH7.5. The concentrations of $HipA_{so}$, $HipB_{so}$ and DNA were measured by a Nanodrop spectrophotometer using an extinction coefficient generated from ProtParam (Uniprot Knowledgebase). All samples were degassed before the experiments. An initial injection of 2ul was followed by 10ul injection with 250s or 350s apart. The raw data were corrected and fitted to the one binding site model, and apparent molar reaction enthalpy (ΔH), apparent entropy (ΔS), dissociation constant (K_D) and stoichiometry of binding (N) were generated.

3.4.6 Small Angle X-ray Scattering

Small angle X-ray Scattering (SAXS) measurements of $HipB_{so}$, $HipB_{so}DNA$, $HipAB_{so}$, and $HipAB_{so}DNA$ were carried out at Beamline SWING in SOLEIL (SOLEIL, Paris, France) with online HPLC separation (Shodex) in 25mM HEPES, 300mM NaCl and 5% glycerol pH7.5 buffer. Buffer subtraction and extrapolation were performed with the program FOXTROT or PRIMUS. The Guinier region was evaluated by PRIMUS using the Guinier approximation. The Porod volume was evaluated using PRIMUS, the radius of gyration Rg, forward scattering I0, maximum particle dimension Dmax and the distance distribution function P(r) were evaluated using GNOM. Theoretical scattering patterns from available structures and generated models, together with their respective discrepancy to the experimental data, were calculated using CRYSOL from ATSAS package(Petoukhov et al., 2012; Svergun et al., 1995). The molecular weight estimation was generated with the protocol described before (Rambo and Tainer, 2013).

3.4.7 Bacterial growth experiments

To construct the HipAB low copy expression vector, $HipAB_{so}$ with native promoter region was PCR purified from genomic DNA of *S.oneidensis* MR-1 with the primers MR13 and MR14 (**Table S1**). The obtained PCR product was gel-purified and subcloned into the TOPO pCR2.1 vector (Invitrogen) prior digestion to ApaLI and XhoI. The digested fragment was cloned into the ApaLI and XhoI digested pACYCDuet expression vector (Merck Millipore), thereby producing the pACYCDuet-$HipAB_{so}$ vector. To obtain the overexpression vectors, The $HipA_{so}$, $HipB_{so}$, $HipAB_{so}$ gene were amplified from the *S.oneidensis* MR-1 with the following primers: MR9 and MR12 for HipAso; MR10 and MR11 for HipBso and MR10, MR12 for $HipAB_{so}$ (**Table S1**). The obtained PCR products were gel purified and cloned into pBAD202/D-TOPO vector (Invitrogen) following the manufacturer's instructions.

Subsequently the thioredoxin sequence was removed by digestion with NcoI, gel purification and relegation, producing respectively the pBADHipAso, pBADHipBso, pBADHipABso vectors. Plasmid DNA was introduced into S.oneidensis MR-1 cells by electroporation using a 2.5kV pulse. After electroporation cells were immediately suspended in 0.5mL LB, incubated at 28 degrees for 1 hour and plated on the LB agar with appropriate antibiotics. Bacteria were routinely grown on Lysogeny Broth (LB) agar plates or shaking (200 rpm) in LB medium at 28℃ for *S.oneidensis* strains. When appropriate, media were supplemented with antibiotics in the following concentrations: carbenicillin (Cb) 100μg/mL, chloramphenicol (Cm) 25μg/mL, kanamycin (Km) 25μg/mL. For the growth curves overnight grown precultures were diluted 1:20 to 1:30 to an optical density at 600 nm (OD600) of ~0.3. Aerobic growth was monitored directly, without the need to transfer culture samples into regular cuvettes, using a Shimadzu 1240 Mini Single-Beam UV-Vis spectrophotometer. Each growth curve was derived from three independent experiments for which the mean (+-SEM) is plotted.

3.4.8 Radioactive kinase assay

A standard radioactive kinase assay was performed in 11μl of reaction buffer (25mM HEPES pH 7.5, 150 mM NaCl, 25 mM β-glycerophosphate, 1 mM Na3VO4, 100μM ATP, and 100 μM MgCl2) in the presence of 10μCi of [γ-32P] ATP and the purified HipA and EF-Tu. The reaction was allowed to incubate for 15 min at room temperature and quenched by the addition of 1/5 volume of SDS-PAGE loading buffer. The ^{32}P labeled proteins were separated by SDS-PAGE, followed by autoradiography of the gels. As a negative control a reaction without ^{32}P and a reaction with BSA were included.

3.5 Results

3.5.1 $HipA_{so}$ and $HipB_{so}$ constitute a toxin-antitoxin module targeting an atypical operator DNA

Transposon mutagenesis of SO0706 led to a strong reduction of biofilm formation in *S.oneidensis* MR-1 (Theunissen et al., 2010). Sequence analysis revealed that the protein product of SO0706 ($HipA_{so}$) shows clear homology to *E.coli* HipA with 28% identity (**Figure S1**). The protein product of the adjacent gene, SO0705 ($HipB_{so}$) is annotated as a toxin-antitoxin antidote transcriptional repressor (Heidelberg et al., 2002). Despite significant similarities between the *S. oneidensis* and *E. coli* sequences, there are also striking differences in the operon structure which may reflect the functional diversity of such toxin-antitoxin systems. For instance, the SO0705 and SO0706 genes are separated by 3 bp instead of overlapping by 1 bp as seen in the *hipAB* operon structure in *E.coli*. Furthermore, a palindromic operator sequence, $GTGTA(N)_6TACAC$ (N indicates random nucleotide), is repeated four times upstream of the $hipAB_{so}$ gene. We note that the random nucleotide sequence within the operator is 2 bp shorter than in the *E.coli* HipAB system. Interestingly, the 4^{th} operator is inside of the $HipB_{so}$ open reading frame and the distance between the neighbor operon is also much larger than in *E.coli* (**Figure S2a**).

To investigate the function of $HipA_{so}$ and $HipB_{so}$ we used *S.oneidensis* MR-1 cDNA to isolate the genes for $HipB_{so}$, $HipA_{so}$, $HipAB_{so}$ and its operators DNA. By employing PCR primer pairs designed to span the entire region of $HipB_{so}$ and $HipA_{so}$ we were able to show that the two genes are transcriptionally coupled (**Figure S2b**). Complementation and overexpression mutants were created to study the $HipAB_{so}$ physiological function and its role in persistence. We constructed pBAD series of vectors which were previously shown to allow a well-controlled expression system to study the toxic effects of $HipA_{so}$ (Korch et al., 2003).

We observed that overexpression of $HipA_{so}$ using a pBAD vector in *S.oneidensis* MR-1 causes a temporary inhibition of growth in both the wild type and the $\Delta HipBA_{so}$ strain (**Figure S2c**). However, when $HipA_{so}$ was overexpressed in the $\Delta HipB_{so}$ strain growth inhibition was acute as previously reported in *E.coli* (Black et al., 1991). This strongly suggests that mild overexpression of $HipA_{so}$ only causes a temporary inhibition but not termination of growth (Korch and Hill, 2006).

To enable studies towards obtaining structural and mechanistic insights into the role of $HipA_{so}$ and $HipB_{so}$ in persistence, we overexpressed and purified recombinant $HipA_{so}$ and $HipB_{so}$. Elution profiles during size exclusion chromatography (SEC), showed that $HipA_{so}$ eluted with as a 50 kDa monomer while $HipB_{so}$ purified as a dimer with an apparent molecular mass of 30kDa. Simple mixing of the proteins resulted in the formation of a complex between $HipA_{so}$ and $HipB_{so}$ with an apparent molecular mass of 130 kDa, consistent with a 2:2 stoichiometry. Furthermore, this complex can bind the single operator DNA sequence forming a homogeneous population (**Figure 1a**). In addition, estimation of the molecular weight estimation from Small Angle X-ray scattering experiments correlated well with the retention volumes in SEC (**Table S2**).

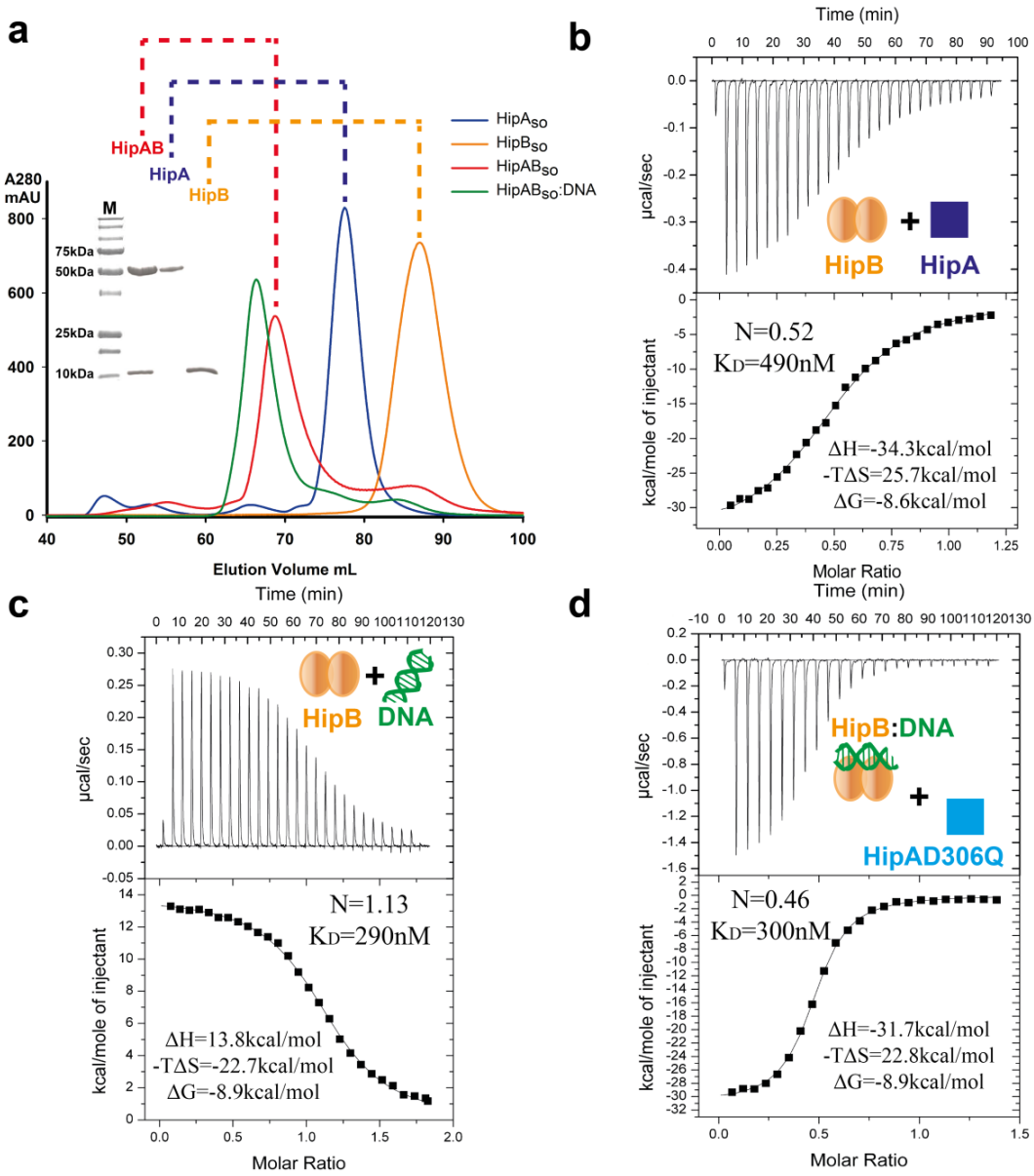


Figure 1: Biochemical reconstitution and characterization of HipB$_{so}$ (SO0705) and HipA$_{so}$ (SO0706) complexes (a) SEC profiles of HipB$_{so}$ (Orange), HipA$_{so}$ (Blue), HipA$_{so}$-HipB$_{so}$ complex (Red) and HipA$_{so}$-HipB$_{so}$: DNA complex (Green). SDS-PAGE analysis of peak fraction is accompanied with the SEC profile. (b) Thermodynamic characterization of the interaction between HipA$_{so}$ and HipB$_{so}$ (c) Thermodynamic characterization of the interaction between HipB$_{so}$ and single DNA operator (d) Thermodynamic characterization between HipB$_{so}$:DNA complex and non-phosphorylated mutant HipA$_{so}$D306Q. The interaction stoichiometry and affinities are shown together with ΔH, -ΔTS and ΔG.

We used Isothermal Titration Calorimetry (ITC) to characterize the interaction between $HipA_{so}$ and $HipB_{so}$, $HipB_{so}$ and its operator DNA, and the recruitment of $HipA_{so}$ to the $HipB_{so}$:DNA complex (**Figure 1b-1d**). These experiments demonstrated that 2 monomers of $HipA_{so}$ bind 1 dimer $HipB_{so}$ with an equilibrium dissociation constant (K_D) of 490nM and via an enthalpically-driven binding event (**Figure 1b**). On the other hand, the interaction of $HipB_{so}$ with operator DNA (26 bp) is endothermic, resulting in an entropically-driven binding event to establish a high-affinity interaction (K_D = 290nM) (**Figure 1c**). Titration of double operator DNA (76bp) with $HipB_{so}$ revealed a similar affinity and thermodynamic fingerprint indicating that there is no cooperativity between the neighbor operators of $HipB_{so}$ binding (**Table S2 and FigureS3c**). Finally, we characterized the recruitment of $HipA_{so}$ to a $HipB_{so}$:DNA complex. For this study we employed preformed $HipB_{so}$:DNA complex and $HipA_{so}$ carrying a D306Q mutation and found that $HipA_{so}$ can bind to the $HipB_{so}$: DNA complex with a K_D of 300 nM and a thermodynamic binding profile that is very similar to the HipA:HipB interaction in the absence of DNA (**Figure 1d**) (**Table S1**).

3.5.2 $HipA_{so}$, $HipB_{so}$ and operator DNA assemble into a distinct ternary complex

To dissect the structural basis of the assembly of the $HipAB_{so}$:DNA ternary complex we crystallized the complex in two different crystal forms and determined crystal structures to 3.39 Å and 3.79 Å resolution (**Table 1**). The *S.oneidensis* MR-1 $HipAB_{so}$:DNA ternary complex assembles as a compact $HipB_{so}$ dimer bound by a 25 bp duplex operator DNA, with each of the $HipB_{so}$ subunits interacting with a different $HipA_{so}$ subunit (**Figure 2a**). The HipABso:DNA ternary assembly bears striking differences compared to the previously characterized *E.coli* sandwich HipABso:DNA ternary complex (PDB:3DNV).

Table 1. X-Ray data collection and refinement statistics

Crystal	HipAB:DNA Crystal form 1	HipAB:DNA Crystal form 2	HipA:AMPPNP:Mg	HipB Dataset1	HipB Dataset2
Data collection					
Space group	$P2_12_12_1$	$P2_1$	$P2_12_12_1$	$C222_1$	$C222_1$
a , b , c (Å)	57.82, 122.45, 189.85	72.24, 57.33, 171.38	60.70, 75.83, 110.32	53.46, 72.95, 86.91	53.57, 71.57, 85.81
α , β , γ (°)	90.0, 90.0, 90.0	90.0, 95.8, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution(Å)	49.39-3.39(3.51-3.39)	47.57-3.79(3.92-3.79)	44.61-1.83(1.9-1.83)	43.12-1.85(1.95-1.85)	38.36-2.35(2.43-2.35)
Rmeas [a]	0.260 (1.112)	0.281 (1.051)	0.115 (0.911)	0.074 (0.761)	0.147 (0.935)
CC(1/2)	0.989 (0.673)	0.983 (0.538)	0.999 (0.904)	0.998 (0.742)	0.995 (0.744)
I/σ(I)	6.78 (1.81)	4.40 (1.21)	19.96 (3.17)	11.66 (1.68)	11.21 (2.15)
Completeness (%)	98.93 (92.72)	97.03 (84.10)	99.43 (95.29)	99.46 (99.20)	99.71 (97.55)
Multiplicity	6.5 (6.4)	3.4 (2.9)	13.1 (12.5)	4.5 (4.4)	6.3 (6.3)
Wilson B-factor	75.47	93.12	21.25	31.00	40.13
Refinement					
Total reflections	124924 (11122)	46748 (3379)	592535 (53603)	67170 (6306)	45147 (4242)
Unique reflections	19211 (1732)	13858 (1164)	45100 (4273)	14778 (1440)	7125 (676)
Rwork/Rfree	0.2560/0.3154	0.2364/0.3064	0.1558/0.1754	0.1861/0.2161	0.2040/0.2238
Number of Atoms					
Protein	8805	8763	3466	1019	1070
Ligand or Ion			34		
Water			486	123	39
Average B-factors	46.44	80.30	27.20	33.90	49.20
Protein	46.44	80.30	26.40	33.20	49.20
Ligand			11.60		
Water			33.80	39.50	47.60
Ramachandran Favored/Allowed	0.978/0.020	0.970/0.027	0.984/0.016	0.992/0.008	0.993/0.007
Root-Mean-Square Deviations					
Bond Lengths (Å)	0.003	0.005	0.005	0.004	0.004
Bond Angles (°)	0.80	0.95	0.95	0.68	0.88

[a] $R_{meas} = \sum_h \sqrt{n_h/(n_h-1)} \sum_h \sum_i |I(h,i)-\langle I(h)\rangle| / \sum_h \sum_i I(h,i)$, where n_h is the multiplicity, I(h,i) is the intensity of the i^{th} measurement of reflection h, and ‹I(h)› is the average value over multiple measurements. Statistics for the highest resolution shell are shown in parentheses.

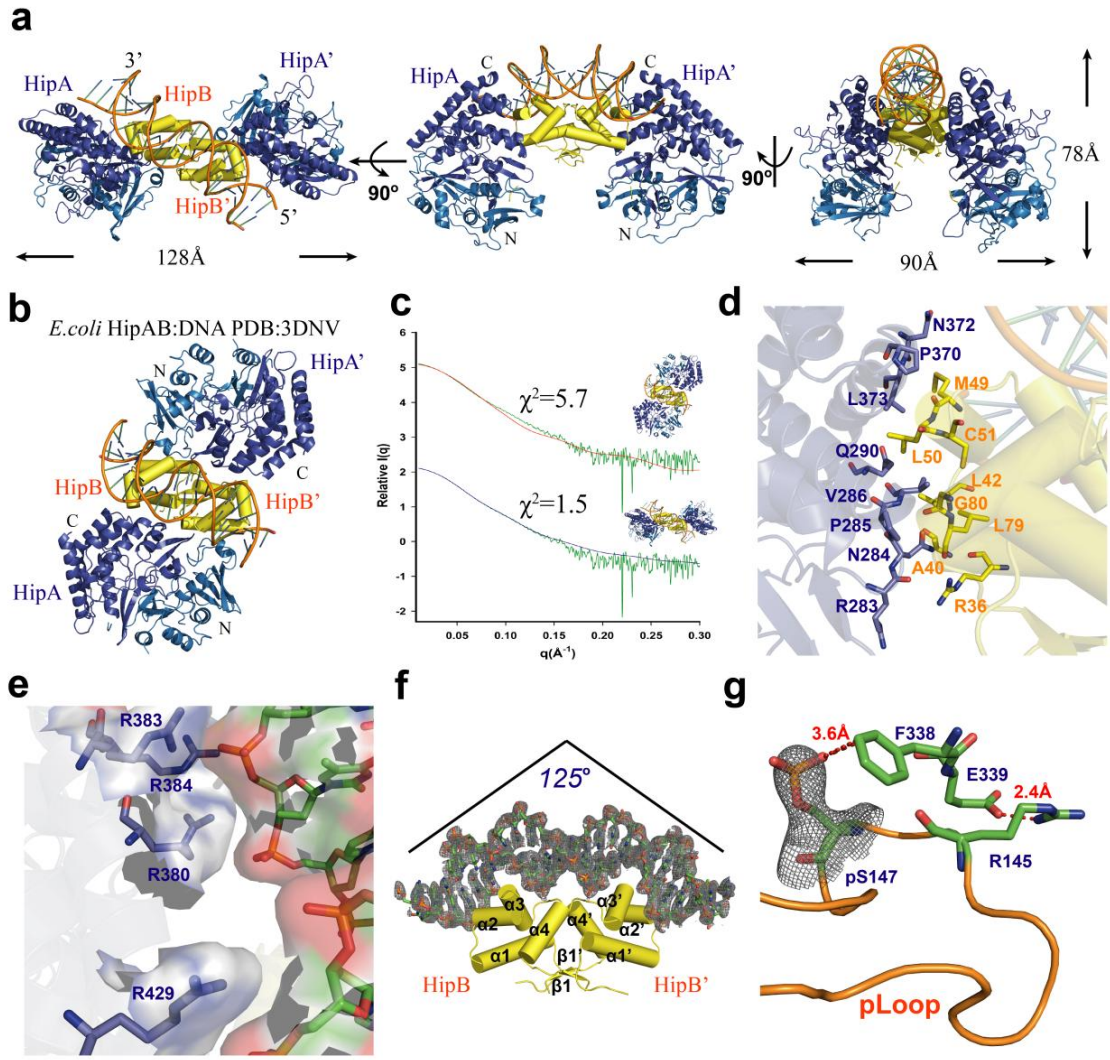


Figure 2: Structural studies of the $HipA_{so}$-$HipB_{so}$: DNA complex (a) X-ray crystal structure of $HipA_{so}$-$HipB_{so}$: DNA complex, $HipB_{so}$ dimer (Yellow) interacts with 2 $HipA_{so}$ C terminus helix bundles (Blue), the $HipA_{so}$ N terminus (Skyblue) has no interaction with $HipB_{so}$. The other 90 degrees rotation views are shown with dimension parameters. (b) *S.oneidensis* MR-1 $HipA_{so}$-$HipB_{so}$: DNA assembles in a distinct way compared to the *E.coli* HipAB: DNA (PDB: 3DNV). The *E.coli* HipAB: DNA structure is shown in the same DNA top view as the left panel in Figure 2a. Instead, in the *E.coli* HipAB: DNA both C and N terminus of HipA interact with both of the HipB dimer molecules while in *S.oneidensis* MR-1 $HipAB_{so}$: DNA assembly only the $HipA_{so}$ C terminus interacts with single $HipB_{so}$ molecules. (c) The SAXS data seal the distinct assembly, showing a fit of *S.oneidensis* MR-1 $HipA_{so}$-$HipB_{so}$: DNA to its solution structure with χ^2=1.5 while the fit of *E.coli* HipA-HipB: DNA is

with χ^2=5.7. (d) The interaction interface between the $HipA_{so}$ and $HipB_{so}$ are shown with involved residues labeled. (e) The interaction interface between $HipB_{so}$ and DNA are shown, the four residues R380, R383, R384 and R429 from $HipA_{so}$ form a positive patch with DNA (f) Interaction between $HipB_{so}$ and DNA resulted in 125 degrees of bending of the normally linear DNA. $HipB_{so}$ mainly uses its α2, α3 and α4 helix turn helix motif to tightly interact with DNA, and detail interactions are plotted with NUCPLOT in Figure S3a. The density of DNA is shown in 2Fo-Fc contoured at 1σ. (g) Both copies of $HipA_{so}$ in the two ternary complexes are phosphorylated at Ser147. This phosphorylated residue is anchored through anion π interaction with $HipA_{so}$ F338 and hydrogen bond is formed between R145 and F338 or E339. The density of pSer147 is shown in 2Fo-Fc=1σ

Instead of binding a single $HipA_{so}$ C terminus helix bundle at each side of the dimer HipBso molecules, in *E.coli* the HipA uses its C and N terminus to bind with both the dimer HipB molecules simultaneous. Viewed from the top of the DNA double helix, there is a 90 degrees rotation of the HipAso out of the sandwich core plate (**Figure 2b**). The *S.oneidensis* MR-1 HipABso:DNA assembly was confirmed to be similar in solution and in crystal by a Small Angle X-ray scattering experiment with χ2=1.5 which sealed the distinct assembly of *E.coli* with a fit χ2=5.7 (**Figure 2c**). The HipA:HipB binding interface is mainly hydrophobic with a buried surface of 470$Å^2$ and features interactions between the C terminal helical-bundle domain of HipA and a hydrophobic patch on each HipB subunit (**Figure 2d**). In addition to the anticipated extensive interactions of HipB with DNA, HipA also makes specific interactions with the phosphodiester flank of the DNA duplex via a positively charged patch defined by R380, R383 and R384 and R429 with an interface around 150 Å2 determined by PDB PISA server (Krissinel and Henrick, 2007) (**Figure 2e**). $HipB_{so}$ interacts tightly with its operator DNA using mainly its α2, α3 and part of α4, forming the expected helix-turn-helix DNA binding motif (**Figure 2f** and **Figure S3a**). Interestingly, binding of duplex DNA to the

$HipAB_{so}$ complex causes the DNA to bend by about 125 degrees compared to linear DNA. $HipB_{so}$ bends the operator DNA more extensively that *E.coli* HipB does (125 degrees vs 110 degrees). The latter can be explained by the fact that the distance between each operators' binding site is shorter.

The $HipA_{so}$-$HipB_{so}$-DNA complex is characterized by two unexpected features. In the first instance, HipA is phosphorylated at Ser147 (**Figure 2g**). The phosphate group of pSer147 is observed at full occupancy in both crystal forms of the $HipA_{so}$-$HipB_{so}$-DNA complex and is anchored by an anion-π interaction with F338. This binding mode is further stabilized by a hydrogen bonds between main-chain atoms of R145 and F338, and a salt-bridge between E339 (**Figure 2g**).

3.5.3 The flexible C-terminal tail of $HipB_{so}$ is essential for targeting the $HipA_{so}$ hydrophobic pocket to prevent proteolysis

An additional striking difference between *E.coli* and *S.oneidensis* HipAB:DNA complexes concerns the structural order of the C-terminus of HipB. We posit that this could have functional implications in terms of the post-translational regulation. In *E.coli*, the C-terminus of HipB is unstructured, even in the HipAB:DNA complex and it has been demonstrated that the C-terminal tryptophan is proteolytically cleaved by the ATP-dependent Lon protease (Hansen et al., 2012). Interestingly, in our structure of the $HipAB_{so}$:DNA to 3.79Å resolution, we can trace the C terminus of $HipB_{so}$ into a hydrophobic pocket of $HipA_{so}$. Moreover, in our second structure of $HipAB_{so}$:DNA to 3.39Å resolution, we see clear density in the same pocket for the last three C terminal residues GWY of $HipB_{so}$ (**Figure 3a** and **Figure S3b**).

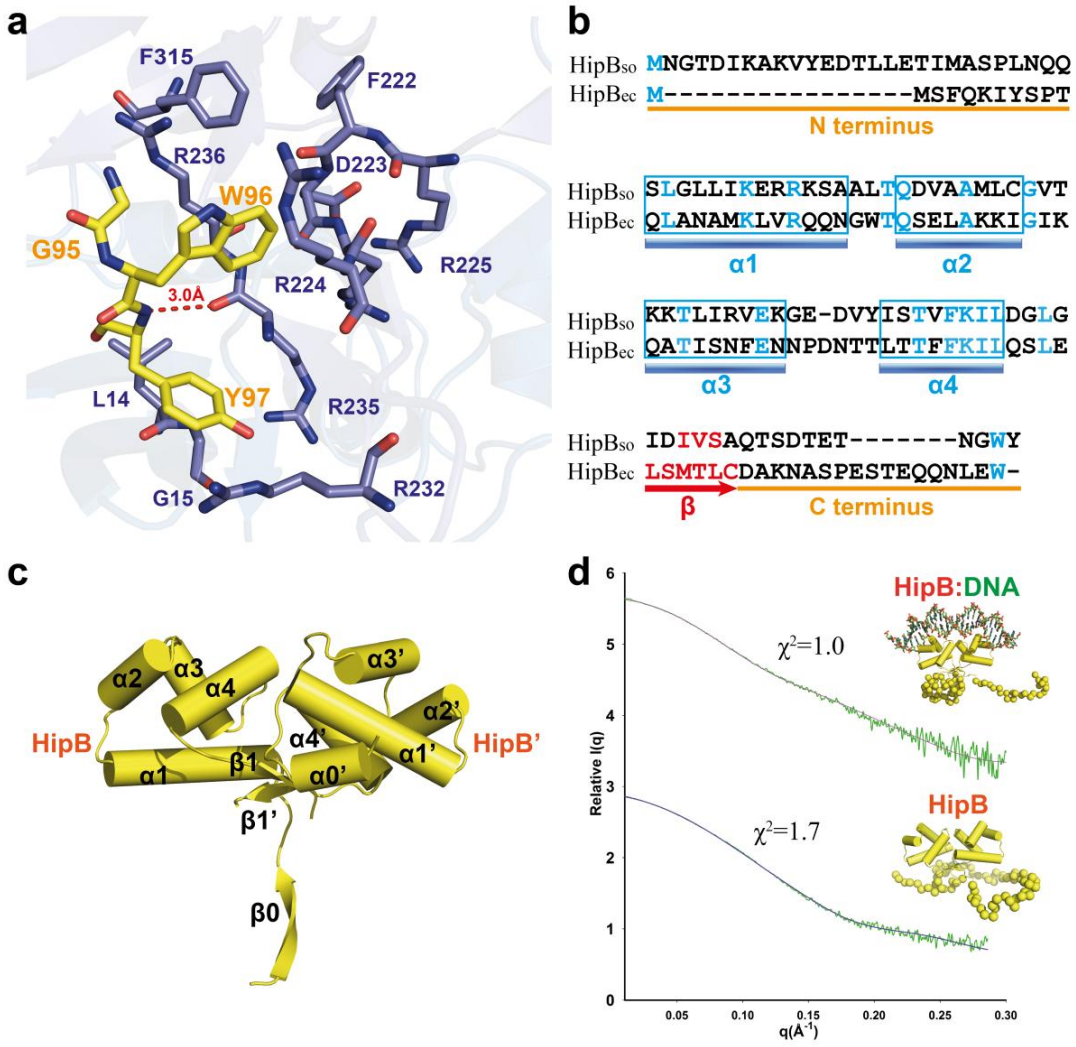


Figure 3: The flexible C-terminal tail of HipB$_{so}$ is accommodated by a hydrophobic pocket in HipA$_{so}$ (a) The HipB$_{so}$ interacts with the hydrophobic core between HipA$_{so}$ N and C terminus by forming an R232$_{HipA}$-W96$_{HipB}$-R236$_{HipA}$-F315$_{HipA}$ cation π interaction and one hydrogen bond with R235. The interaction between HipB$_{so}$ C terminus GWY with HipA$_{so}$ is plotted by Ligplot in Figure S3b. (b) Although with only 18% of identity, the secondary structures of *S.oneidensis* MR-1 and *E.coli* HipB are highly conserved and the disordered C terminus W residue involved in targeting the HipA hydrophobic core is present in both. (c) The difference in structure of HipB$_{so}$ N termini is observed in the 2.35 resolution crystal form. The N termini from the dimer HipB$_{so}$ form a α0' helix and β0 sheet respectively. (d) SAXS modeling indicated that both HipB$_{so}$ and HipB$_{so}$: DNA complex have a disordered N- and C-terminus, which indicates that the binding of HipB$_{so}$ to HipA$_{so}$ is essential to keep the structure of the HipB$_{so}$ C-terminus ordered.

As it is widely accepted that the protease-sensitivity of the antitoxins is a mechanism for the regulation of toxin activity (Brzozowska and Zielenkiewicz, 2013; Donegan et al., 2010; Hansen et al., 2012), we pursued structural studies of HipB outside the context of the ternary complex with HipA and DNA to obtain insights into the plasticity and structural order of the C-terminus in $HipB_{so}$.. We crystallized the HipB alone and solved the structure to 1.85Å and 2.35Å resolution (**Table 1**). The crystal structures revealed that both the N- and C-terminal segments in $HipB_{so}$ are flexible. Moreover, a structure-based sequence alignment reveals a 16 residues disordered extension in $HipB_{so}$ of which the function is yet unknown (Pei et al., 2008) (**Figure 3b**). Interestingly, in our structure of $HipB_{so}$ to 2.35 Å resolution, the N termini of the two $HipB_{so}$ monomers adopt two different secondary structures, namely a β0 strand and α0' helix respectively. We also constructed the truncated form lacking the 16 residues N-terminal extension and proved there was no interference of this disordered N terminus on the $HipA_{so}$ and $HipB_{so}$ binding by ITC experiment (**Figure S3c**). The C-terminus of $HipB_{so}$ alone is also flexible as no density can be observed in both of the high resolution crystal datasets (**Figure 3c**). Furthermore, SAXS experiments indicated that both in $HipB_{so}$ alone and in $HipB_{so}$ bound with operator DNA, the C terminus of $HipB_{so}$ is disordered (**Figure 3d**). Thus, the C terminus of $HipB_{so}$ is flexible and becomes ordered in the $HipAB_{so}$:DNA complex.

3.5.4 $HipA_{so}$ is conformationally regulated by autophosphorylation

In *E.coli*, HipA kinase activity is essential for growth arrest and multidrug tolerance (Correia et al., 2006). To further investigate the mechanism of $HipA_{so}$ in *S. oneidensis* MR-1, we solved the structure of $HipA_{so}$ in complex with the non-hydrolyzable ATP analogue AMPPNP to 1.83Å resolution (**Table 1**).

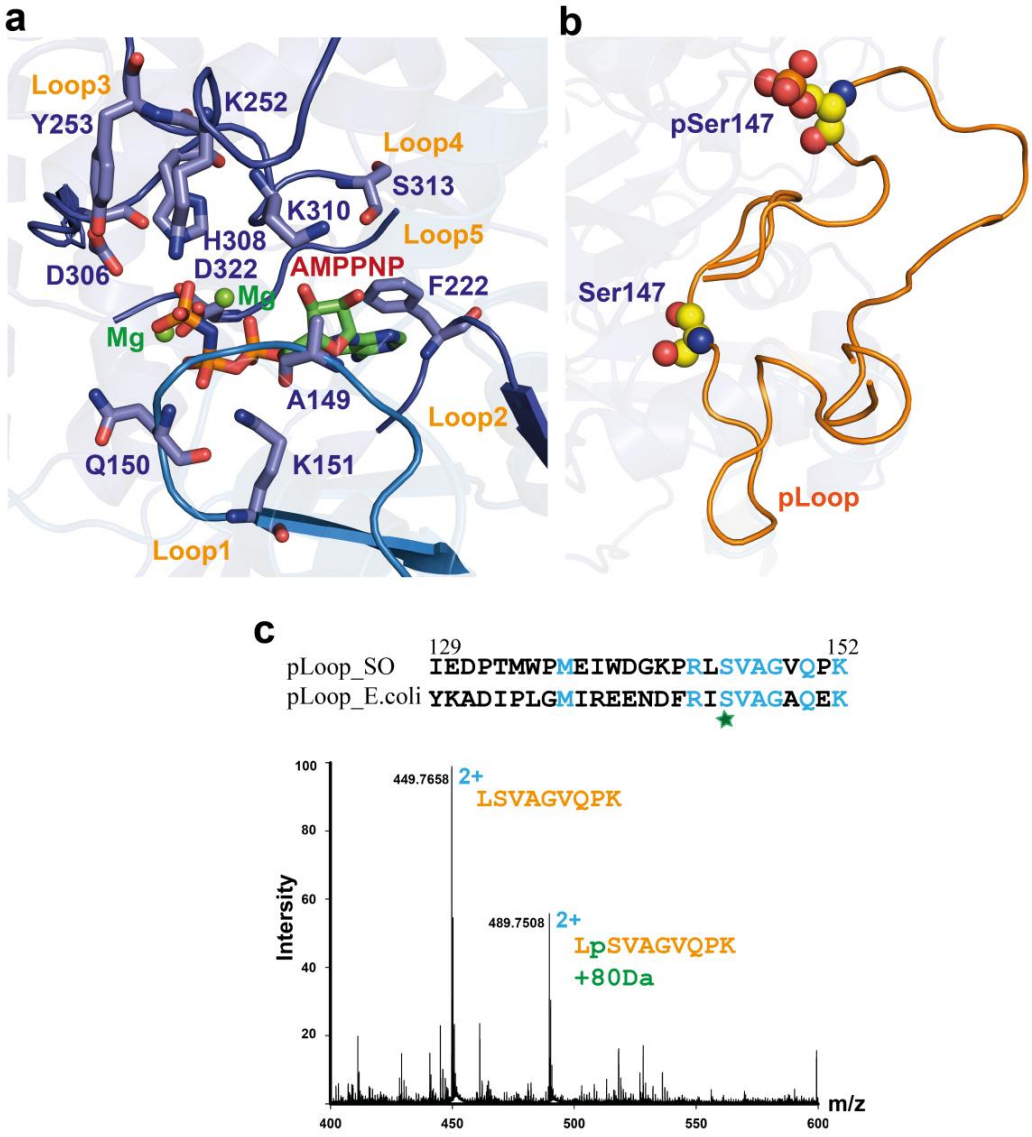


Figure 4: *S.oneidensis* MR-1 $HipA_{so}$ is regulated conformationally upon autophosphorylation (a) 5 loops in the $HipA_{so}$ structure are involved in the binding of AMPPNP and 2 Mg^{2+} ions, the residues involved in the Mg-AMPPMP binding are shown in stick and labeled in blue. The two Mg^{2+} ions are shown in green spheres. Detailed interaction between $HipA_{so}$ and AMPPNP, Mg^{2+} are plotted by Ligplot in Figure S4a. (b) Detailed view of the pLoop ejection after phosphorylation of $HipA_{so}$.Ser147 from AMPPNP binding $HipA_{so}$ and pSer147 from $pHipA_{so}$ are shown in sphere with colored in C yellow, N blue, O red, P Orange. (c) The sequence alignment between *S.oneidensis* MR-1 and *E.coli* is accompanied with nanoLCMS identification of the phosphorylated peptide, the peptide LSVAGVQPK was observed at charge 2+ state in two forms differing 80Da in molecular weight.

This structural analysis at high resolution revealed binding of AMPPNP and Mg^{2+} at high occupancy and delineation of the full interaction landscape (**Figure 4a** and **Figure S4a**)**.** The ATP binding site in $HipA_{so}$ is engulfed by the N- and C-terminal domains and most interacting residues are contributed by five loop regions (**Figure S4b**). All residues involved in Mg^{2+} and ATP binding are highly conserved in *S.oneidensis* MR-1 and *E.coli* (**Figure 4a**) (**Figure S1**). We propose that phosphorylation of $HipA_{so}$ is essential for the formation of the ternary $HipAB_{so}$:DNA complex. This could be a common feature for the HipA kinase family. The $HipA_{so}$ pLoop consists of residues 129-152 and is ejected out of the pocket compared to the non-phosphorylated form (**Figure 4b**). Furthermore, the pLoop in autophosphorylated $HipA_{so}$ as observed in the $HipAB_{so}$:DNA complex, projects to the surface. Such activation mechanism involving ejection of the pLoop upon phosphorylation has also been proposed for in *E.coli* HipA (Schumacher et al., 2012). To investigate the autophosphorylation capacity of $HipA_{SO}$, we incubated freshly purified HipA with ATP and Mg^{2+} overnight and observed a dramatic increase of the corresponding phosphorylated peptide intensity from approximately 8% to almost 50% relative to the signal of the non-phosphorylated peptide in LC-MS (**Fig. 4c**).

In *E.coli*, $HipA_{so}$ is autophosphorylated at Ser150. Similarly, our structural analysis revealed that in *S.oneidensis* MR-1, HipA can autophosphorylate at Ser147 which aligns with Ser150 in *E.coli* HipA. We note that both of our crystal forms of the $HipAB_{so}$:DNA complex contain phosphorylated HipA in sharp contrast to the *E.coli* HipAB complex (Evdokimov et al., 2009; Schumacher et al., 2012). Upon superposition of the N-terminal domains of phosphorylated and unphosphorylated $HipA_{so}$ (derived from the AMPPNP complex) we observed a rigid-body shift of the $HipA_{so}$ C-terminal domain by about 2 Å shifts (**Figure 4Sc**). Given that this part of $HipA_{so}$ interacts with DNA in the ternary complex with $HipB_{so}$,

we propose that the phosphorylated form of $HipA_{so}$ may be important in the complex assembly and stability.

3.5.5 Functional investigation of *S.oneidensis* MR-1 $HipA_{so}$ (SO0706)

HipA was the first toxin protein that was linked to the phenomenon of persistence and, its overexpression in *E.coli* resulted in multiple effects including growth arrest, inhibition of macromolecular synthesis as well as DNA replication and transcription (Correia et al., 2006; Korch and Hill, 2006; Korch et al., 2003; Pei et al., 2008). EFTu was originally identified as the substrate of the kinase activity of HipA in *E.coli,* thereby creating a functional link between HipA and the inhibition of translation (Schumacher et al., 2009). However, repeated attempts to show that $HipA_{so}$ has an analogous activity failed, leading to the conclusion that $EFTu_{so}$ does most likely not constitute a substrate for $HipA_{so}$ (**Figure S5a**). Furthermore, ITC and pull-down experiments revealed the absence of interaction between these proteins thereby corroborating the functional assay (**Figure S5b**).

3.6 Discussion

We here provided structural insights into the toxin antitoxin system HipA-HipB and its operator DNA assembly in *S.oneidensis* MR-1, which can now serve as a second example of such complexes next to the previously characterized *E.coli* HipAB:DNA complex. The *S.oneidensis* MR-1 $HipAB_{so}$:DNA assembly is sharply distinct from the described *E.coli* assembly, providing new insights into the structural and mechanistic diversity of HipAB modules and cautioning against the use of a single model organism for deriving functional generalizations. In *E.coli* the phosphorylation of HipA evokes pLoop ejection and is thought to release HipA from HipB in the HipAB:DNA complex (Schumacher et al., 2012). However, we have now provided direct structural evidence that phosphorylated HipA can participate in

the $HipAB_{so}$:DNA complex form, and further that the phosphorylation site is blocked and anchored through anion-π interaction. This suggests that HipB may play an active role in recruiting activated HipA to avoid growth arrest. We also observed that upon release of HipA, the HipB structure is destabilized, probably making it more prone to proteolytic degradation. This could help to propagate the effect of HipA-regulated persistence. This raises a new question about the possible molecular trigger for the release of HipA from HipB given that p-loop ejection does not appear to do so.

We further provide information that EFTu is unlikely to be the substrate of toxin HipA in the case of *S.oneidensis*, which is in agreement with more recent findings. While it was originally believed that it HipA halted ribosomal activity by phosphorylating EFTu (Schumacher et al., 2009, 2012), it was recently established that its natural substrate is actually glutamyl-tRNA-synthetase. By phosporylating this enzyme at its ATP binding site, uncharged $tRNA^{Glu}$ accumulates and the stringent response is activated (Germain et al., 2013a; Kaspy et al., 2013b). Mining the *S.oneidensis* MR-1 genome database and the Toxin-Antitoxin database (TADB) revealed that multiple homologues of the HipAB toxin antitoxin system exist in *S.oneidensis* MR-1 (Shao et al., 2011). Besides SO0705 ($HipB_{so}$), SO0706 ($HipA_{so}$); SO3169, SO3170 and SO0062, SO0063 are also annotated as members of the HipAB toxin antitoxin system. Interestingly, despite the low identity (15-20%) between the predicted HipA-like proteins SO3170, SO0063 and SO0706, the residues involved in Mg^{2+}-ATP binding and Ser phosphorylation are highly conserved (**Figure S6**). Moreover, upstream of those predicted HipA like proteins, there is also one helix-turn-helix DNA binding HipB-like protein and a potential operator. Thus, it appears that there might be functional redundancy of such proteins in *S.oneidensis* which may complicate future studies in delineating the exact role of HipA activity in this organism.

3.7 References

Acar, M., Mettetal, J.T., and van Oudenaarden, A. (2008). Stochastic switching as a survival strategy in fluctuating environments. Nat. Genet. *40*, 471–475.

Adams, P.D., Afonine, P. V, Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D. Biol. Crystallogr. *66*, 213–221.

Avery, S. V (2006). Microbial cell individuality and the underlying sources of heterogeneity. Nat Rev Micro *4*, 577–587.

Balaban, N.Q., Gerdes, K., Lewis, K., and McKinney, J.D. (2013). A problem of persistence: still more questions than answers? Nat. Rev. Microbiol. *11*, 587–591.

Bigger, J.W. (1944). Treatment of staphylococcal infections with penicillin. Lancet *244*, 497–500.

Black, D.S., Kelly, A.J., Mardis, M.J., and Moyed, H.S. (1991). Structure and organization of hip, an operon that affects lethality due to inhibition of peptidoglycan or DNA synthesis. J. Bacteriol. *173*, 5732–5739.

Black, D.S., Irwin, B., and Moyed, H.S. (1994). Autoregulation of hip, an operon that affects lethality due to inhibition of peptidoglycan or DNA synthesis. J. Bacteriol. *176*, 4081–4091.

Bokinsky, G., Baidoo, E.E.K., Akella, S., Burd, H., Weaver, D., Alonso-Gutierrez, J., García-Martín, H., Lee, T.S., and Keasling, J.D. (2013a). HipA-triggered growth arrest and β-lactam tolerance in *Escherichia coli* are mediated by RelA-dependent ppGpp synthesis. J. Bacteriol. *195*, 3173–3182.

Bokinsky, G., Baidoo, E.E.K., Akella, S., Burd, H., Weaver, D., Alonso-Gutierrez, J., García-Martín, H., Lee, T.S., and Keasling, J.D. (2013b). HipA-triggered growth arrest and β-lactam tolerance in *Escherichia coli* is mediated by RelA-dependent ppGpp synthesis. J. Bacteriol. *195*, 3173–3182.

Brzozowska, I., and Zielenkiewicz, U. (2013). Regulation of toxin-antitoxin systems by proteolysis. Plasmid *70*, 33–41.

Cataudella, I., Trusina, A., Sneppen, K., Gerdes, K., and Mitarai, N. (2012). Conditional cooperativity in toxin-antitoxin regulation prevents random toxin activation and promotes fast translational recovery. Nucleic Acids Res. *40*, 6424–6434.

Cataudella, I., Sneppen, K., Gerdes, K., and Mitarai, N. (2013). Conditional cooperativity of toxin - antitoxin regulation can mediate bistability between growth and dormancy. PLoS Comput. Biol. *9*, e1003174.

Chao, L., Rakshe, S., Leff, M., and Spormann, A.M. (2013). PdeB, a Cyclic Di-GMP-Specific Phosphodiesterase That Regulates Shewanella oneidensis MR-1 Motility and Biofilm Formation. J. Bacteriol. *195*, 3827–3833.

Correia, F.F., D'Onofrio, A., Rejtar, T., Li, L., Karger, B.L., Makarova, K., Koonin, E. V, and Lewis, K. (2006). Kinase activity of overexpressed HipA is required for growth arrest and multidrug tolerance in *Escherichia coli*. J. Bacteriol. *188*, 8360–8367.

Donegan, N.P., Thompson, E.T., Fu, Z., and Cheung, A.L. (2010). Proteolytic regulation of toxin-antitoxin systems by ClpPC in *Staphylococcus aureus*. J. Bacteriol. *192*, 1416–1422.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D. Biol. Crystallogr. *60*, 2126–2132.

Evdokimov, A., Voznesensky, I., Fennell, K., Anderson, M., Smith, J.F., and Fisher, D.A. (2009). New kinase regulation mechanism found in HipBA: a bacterial persistence switch. Acta Crystallogr. D. Biol. Crystallogr. *65*, 875–879.

Falla, T.J., and Chopra, I. (1998). Joint tolerance to beta-lactam and fluoroquinolone antibiotics in *Escherichia coli* results from overexpression of hipA. Antimicrob. Agents Chemother. *42*, 3282–3284.

Feng, J., Kessler, D.A., Ben-Jacob, E., and Levine, H. (2014). Growth feedback as a basis for persister bistability. Proc. Natl. Acad. Sci. U. S. A. *111*, 544–549.

Gelens, L., Hill, L., Vandervelde, A., Danckaert, J., and Loris, R. (2013). A General Model for Toxin-Antitoxin Module Dynamics Can Explain Persister Cell Formation in E. coli. PLoS Comput. Biol. *9*, e1003190.

Germain, E., Castro-Roa, D., Zenkin, N., and Gerdes, K. (2013a). Molecular Mechanism of Bacterial Persistence by HipA. Mol. Cell http://dx.doi.org/10.1016/j.molcel.2013.08.045.

Germain, E., Castro-Roa, D., Zenkin, N., and Gerdes, K. (2013b). Molecular mechanism of bacterial persistence by HipA. Mol. Cell *52*, 248–254.

Gorby, Y.A., Yanina, S., McLean, J.S., Rosso, K.M., Moyles, D., Dohnalkova, A., Beveridge, T.J., Chang, I.S., Kim, B.H., Kim, K.S., et al. (2006). Electrically conductive bacterial nanowires produced by *Shewanella oneidensis* strain MR-1 and other microorganisms. Proc. Natl. Acad. Sci. U. S. A. *103*, 11358–11363.

Hansen, S., Vulić, M., Min, J., Yen, T.-J., Schumacher, M. a, Brennan, R.G., and Lewis, K. (2012). Regulation of the Escherichia coli HipBA toxin-antitoxin system by proteolysis. PLoS One *7*, e39185.

Heidelberg, J.F., Paulsen, I.T., Nelson, K.E., Gaidos, E.J., Nelson, W.C., Read, T.D., Eisen, J. a, Seshadri, R., Ward, N., Methe, B., et al. (2002). Genome sequence of the dissimilatory metal ion-reducing bacterium *Shewanella oneidensis*. Nat. Biotechnol. *20*, 1118–1123.

Hersh, M.N., Ponder, R.G., Hastings, P.J., and Rosenberg, S.M. (2004). Adaptive mutation and amplification in *Escherichia coli*: two pathways of genome adaptation under stress. Res. Microbiol. *155*, 352–359.

Kabsch, W. (2010). XDS. Acta Crystallogr. D. Biol. Crystallogr. *66*, 125–132.

Karplus, P.A., and Diederichs, K. (2012). Linking crystallographic model and data quality. Science *336*, 1030–1033.

Kasari, V., Mets, T., Tenson, T., and Kaldalu, N. (2013). Transcriptional cross-activation between toxin-antitoxin systems of *Escherichia coli*. BMC Microbiol. *13*, 45.

Kaspy, I., Rotem, E., Weiss, N., Ronin, I., Balaban, N.Q., and Glaser, G. (2013a). HipA-mediated antibiotic persistence via phosphorylation of the glutamyl-tRNA-synthetase. Nat. Commun. *4*, 3001.

Kaspy, I., Rotem, E., Weiss, N., Ronin, I., Balaban, N.Q., and Glaser, G. (2013b). HipA-mediated antibiotic persistence via phosphorylation of the glutamyl-tRNA-synthetase. Nat. Commun. *4*, 1–7.

Korch, S.B., and Hill, T.M. (2006). Ectopic overexpression of wild-type and mutant hipA genes in *Escherichia coli*: effects on macromolecular synthesis and persister formation. J. Bacteriol. *188*, 3826–3836.

Korch, S.B., Henderson, T. a., and Hill, T.M. (2003). Characterization of the hipA7 allele of *Escherichia coli* and evidence that high persistence is governed by (p)ppGpp synthesis. Mol. Microbiol. *50*, 1199–1213.

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. *372*, 774–797.

Lewis, K. (2010). Persister cells. Annu. Rev. Microbiol. *64*, 357–372.

Lewis, K. (2013). Platforms for antibiotic discovery. Nat. Rev. Drug Discov. *12*, 371–387.

Lin, C.-Y., Awano, N., Masuda, H., Park, J.-H., and Inouye, M. (2013). Transcriptional repressor HipB regulates the multiple promoters in Escherichia coli. J. Mol. Microbiol. Biotechnol. *23*, 440–447.

Long, F., Vagin, A.A., Young, P., and Murshudov, G.N. (2008). BALBES: a molecular-replacement pipeline. Acta Crystallogr. D. Biol. Crystallogr. *64*, 125–132.

Moyed, H.S., and Bertrand, K.P. (1983). HipA , a Newly Recognized Gene of *Escherichia coli* K-12 That Affects Frequency of Persistence After Inhibition of Murein Synthesis. J. Bacteriol. *155*, 768–775.

Moyed, H.S., and Broderick, S.H. (1986). Molecular cloning and expression of hipA, a gene of *Escherichia coli* K-12 that affects frequency of persistence after inhibition of murein synthesis. J. Bacteriol. *166*, 399–403.

Pei, J., Kim, B.-H., and Grishin, N. V (2008). PROMALS3D: a tool for multiple protein sequence and structure alignments. Nucleic Acids Res. *36*, 2295–2300.

Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A.G., Gajda, M., Gorba, C., Mertens, H.D.T., Konarev, P. V., and Svergun, D.I. (2012). New developments in the ATSAS program package for small-angle scattering data analysis. J. Appl. Crystallogr. *45*, 342–350.

Rambo, R.P., and Tainer, J. a (2013). Accurate assessment of mass, models and resolution by small-angle scattering. Nature *496*, 477–481.

Rotem, E., Loinger, A., Ronin, I., Levin-Reisman, I., Gabay, C., Shoresh, N., Biham, O., and Balaban, N.Q. (2010). Regulation of phenotypic variability by a threshold-based mechanism underlies bacterial persistence. Proc. Natl. Acad. Sci. U. S. A. *107*, 12541–12546.

Scherrer, R., and Moyed, H.S. (1988). Conditional impairment of cell division and altered lethality in hipA mutants of Escherichia coli K-12. J. Bacteriol. *170*, 3321–3326.

Schumacher, M.A., Piro, K.M., Xu, W., Hansen, S., Lewis, K., and Brennan, R.G. (2009). Molecular mechanisms of HipA-mediated multidrug tolerance and its neutralization by HipB. Science *323*, 396–401.

Schumacher, M.A., Min, J., Link, T.M., Guan, Z., Xu, W., Ahn, Y.-H., Soderblom, E.J., Kurie, J.M., Evdokimov, A., Moseley, M.A., et al. (2012). Role of Unusual P Loop Ejection and Autophosphorylation in HipA-Mediated Persistence and Multidrug Tolerance. Cell Rep. *2*, 518–525.

Shao, Y., Harrison, E.M., Bi, D., Tai, C., He, X., Ou, H.-Y., Rajakumar, K., and Deng, Z. (2011). TADB: a web-based resource for Type 2 toxin-antitoxin loci in bacteria and archaea. Nucleic Acids Res. *39*, D606–11.

Svergun, D., Barberato, C., and Koch, M.H.J. (1995). CRYSOL a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. J. Appl. Crystallogr. 768–773.

Theunissen, S., De Smet, L., Dansercoer, A., Motte, B., Coenye, T., Van Beeumen, J.J., Devreese, B., Savvides, S.N., and Vergauwen, B. (2010). The 285 kDa Bap/RTX hybrid cell surface protein (SO4317) of *Shewanella oneidensis* MR-1 is a key mediator of biofilm formation. Res. Microbiol. *161*, 144–152.

Thormann, K.M., Saville, R.M., Shukla, S., Dale, A., Spormann, A.M., Saville, M., and Pelletier, D.A. (2004). Initial Phases of Biofilm Formation in *Shewanella oneidensis* MR-1. J. Bacteriol. *186*, 8096–8104.

Thormann, K.M., Duttler, S., Saville, R.M., Hyodo, M., Shukla, S., Hayakawa, Y., and Spormann, A.M. (2006). Control of Formation and Cellular Detachment from *Shewanella oneidensis* MR-1 Biofilms by Cyclic di-GMP. J. Bacteriol. *188*, 2681–2691.

Vagin, a., and Teplyakov, a. (1997). MOLREP : an Automated Program for Molecular Replacement. J. Appl. Crystallogr. *30*, 1022–1025.

Winter, G. (2010). Xia2: an expert system for macromolecular crystallography data reduction. J. Appl. Crystallogr. *43*, 186–190.

Chapter 3

Structural Insights into HipA-HipB Toxin-Antitoxin System in *Shewanella oneidensis* MR-1

SUPPLEMENTARY MATERIALS

Figure S1. Sequence alignment of *S.oneidensis* MR-1 and *E.coli* K12 HipA which indicates a 28% identity, the orange underlined highlighted sequences show the 5 loops involved in AMPPNP and Mg^{2+} binding.

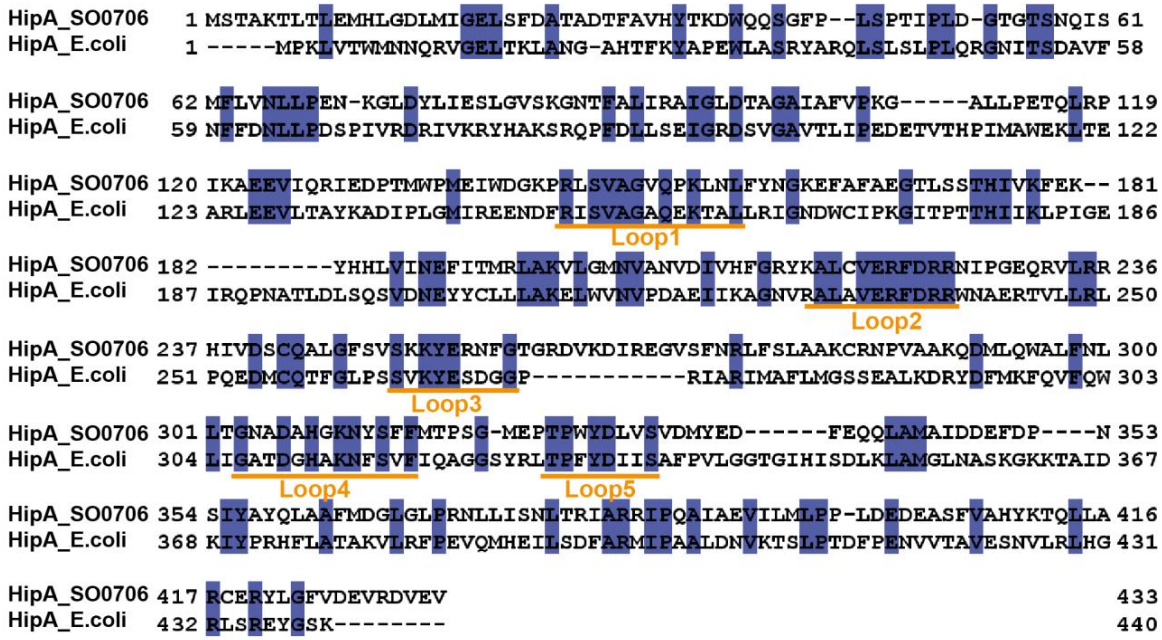


Figure S2. a. Comparison of operator sequences in *S*. *oneidensis* MR-1 and *E*. *coli* HipAB system. The 4 operator sequences were underlined and conserved sequence were in bold. The red highlight indicates the start codon of HipB. **b.** Operon organization of $HipAB_{so}$ from *S*. *oneidensis* MR-1. Agarose gels are stained with ethidium bromide. (A) *S*. *oneidensis* total cDNA template and primers ES31 and ES32. (B) *S*. *oneidensis* total cDNA template and primers ES46 and ES47. (C) *S*. *oneidensis* total cDNA template and primers ES48 and ES49. (Neg. Ctrl. =negative control, using no-RT RNA template to asses DNA contamination in RNA preparations; Pos.Ctrl. = positive control, using *S*. *oneidensis* genomic DNA template). **c.** Growth experiments indicated overexpression of HipAso results in temporary inhibition of growth in *S.oneidensis* MR-1. Left panel: overexpression of HipAso, HipBso and HipABso in the wild type *S.oneidensis* MR-1; right panel: overexpression of HipBso, HipABso in the ΔHipAB mutant *S.oneidensis* MR-1. At least three independent bacteria growth of each strain were evaluated for each experiments, error bars indicate the standard error of mean.

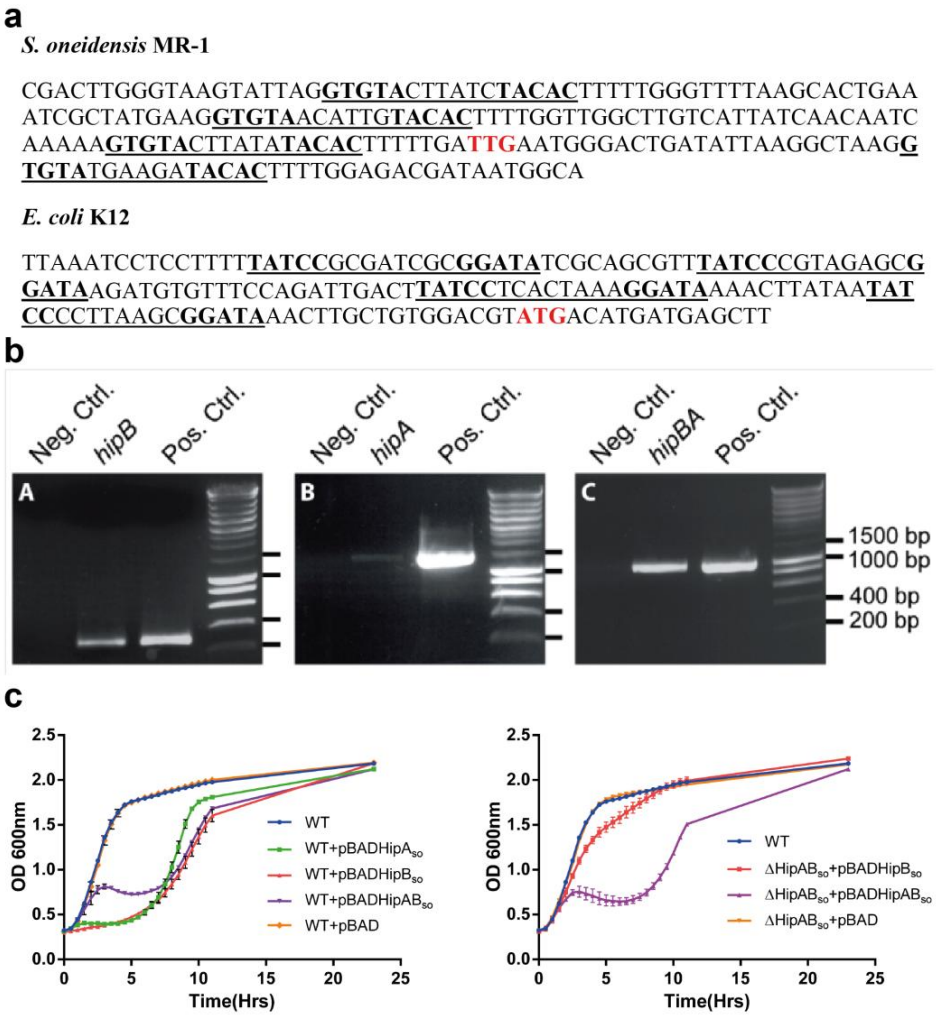


Figure S3.a. Protein:DNA interaction sites in the ternary $HipA_{so}$-$HipB_{so}$:DNA complex plotted by NUCPLOT. **b.** The Two $HipB_{so}$ monomers C terminus GWY interaction with $HipA_{so}$ molecules plotted by Ligplot. Red circles indicate the hydrophobic interaction from $HipA_{so}$. **c.** ITC experiments of HipB with 2 operator DNA (76bp) and Nt truncated HipB with HipA

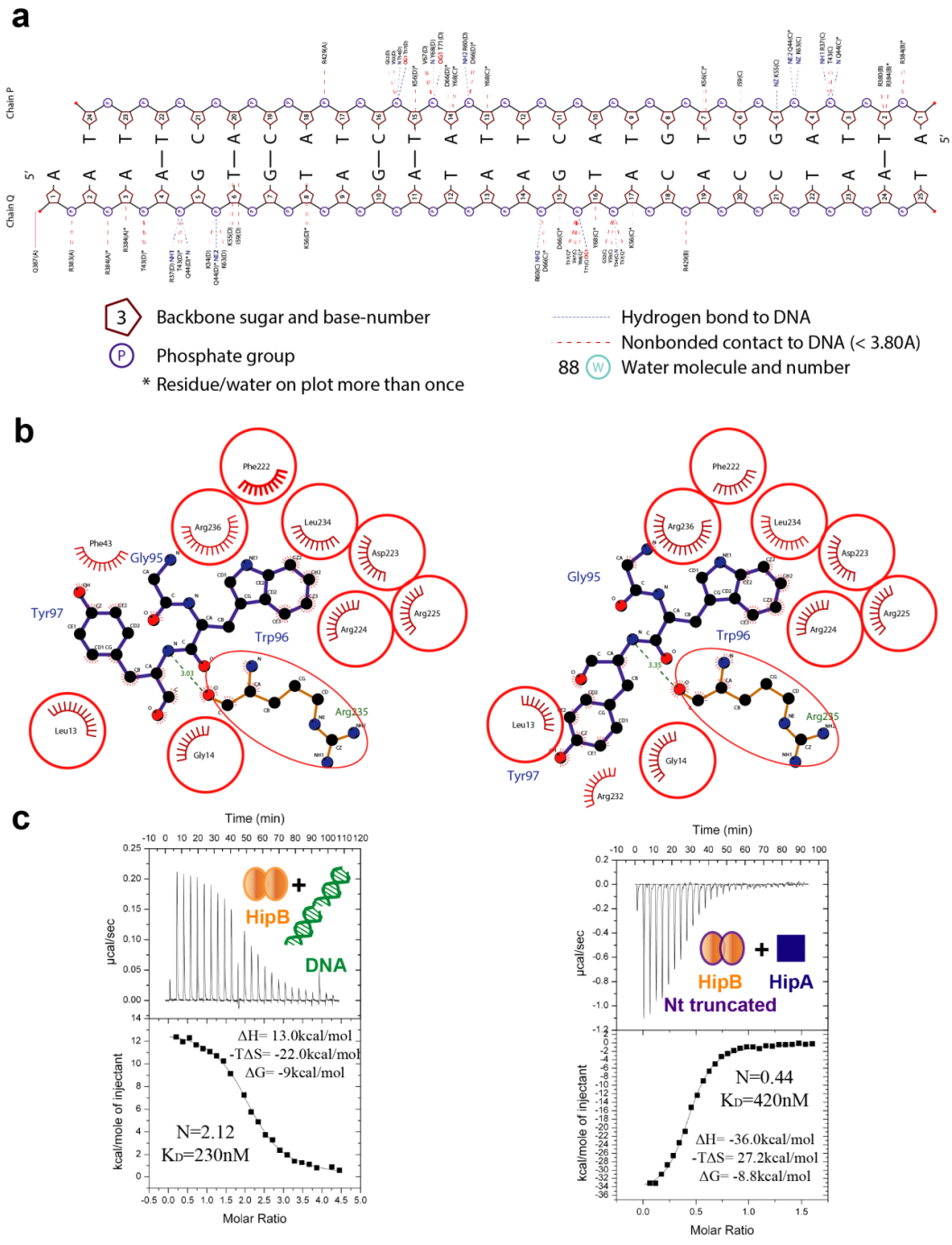


Figure S4. a. Details of the interaction between $HipA_{so}$, AMPPNP, Mg^{2+} and water molecules plotted by Ligplot. The blue balls represent the water molecules involved in the interaction. **b.** Density map of AMPPNP in the $HipA_{so}$-AMPPNP-Mg structure showed in 2Fo-Fc=2σ level, the surface view of the$HipA_{so}$-AMPPNP-Mg structure represents a pocket for the ATP binding. **c.** Alignment of the $HipA_{so}$ AMPPNP Mg bound form and phosphorylated HipA form showed that the phosphorylation loop is ejected from inside to outside. The pLoop is shown in orange with phosphorylation site is shown in atom colored sphere. The alignment of the N terminus also resulted in a 2Å shift of the C terminus helix bundle which is essential for the DNA binding. pHipA in blue, AMPPNP Mg binding HipA in grey.

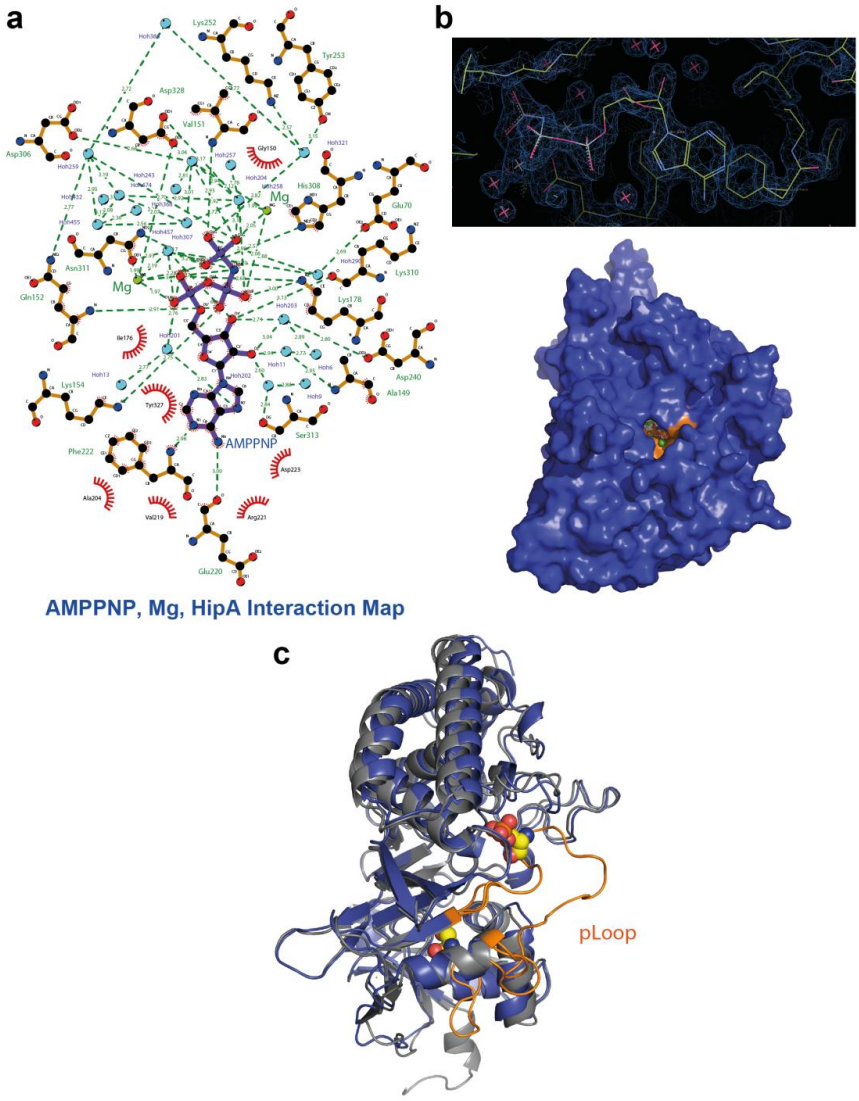


Figure S5. a. Radioactive kinase assay indicated there is no phosphate transfer between *S.oneidensis* MR-1 HipAso and EFTuso. $HipA_{so}$ was used as positive control and $HipA_{so}$D306Q was used as negative control in this experiment. 1 HipAsoD306Q, 2 HipAso, 3 $HipA_{so}$+$EFTu_{so}$-His, 4 $HipA_{so}$+$EFTu_{so}$-GST, 5 $HipA_{so}$+$EFTu_{so}$-His+GDP, 6 $HipA_{so}$+$EFTu_{so}$-GST+GDP, 7 $HipA_{so}$D306Q+$EFTu_{so}$-His+GDP 8 $HipA_{so}$D306Q+$EFTu_{so}$-GST+GDP. SDS PAGE analysis 9 $EFTu_{so}$-His, 10 $HipA_{so}$, 11 $EFTu_{so}$-GST. **b.** ITC experiment reveals that there is no interaction between $HipA_{so}$ and $EFTu_{so}$.

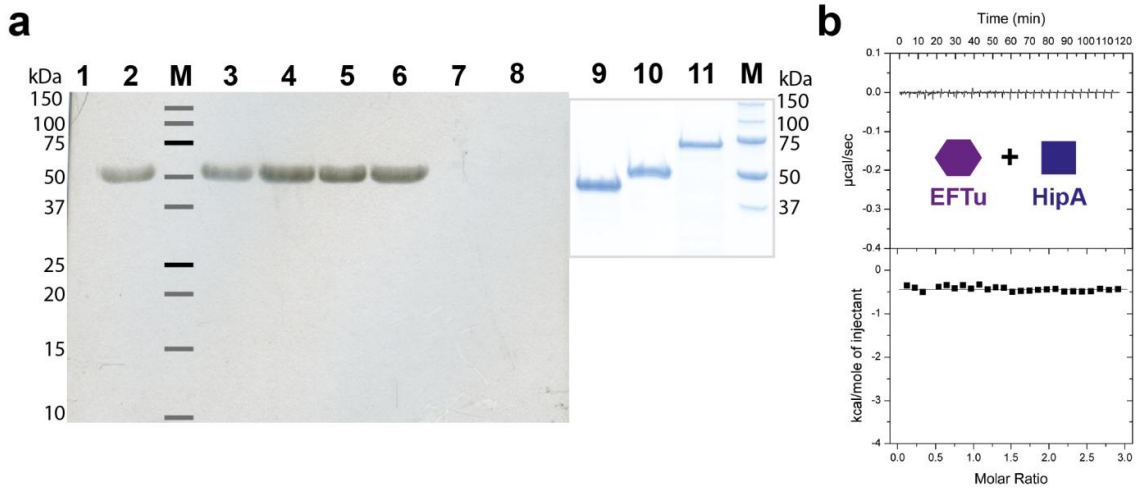


Figure S6. Sequence alignment of SO0706 ($HipA_{so}$); SO3170 and SO0063, the green stars represent the potential Ser phosphorylation site and the red star represent the ATP and Mg^{2+} binding site, those residues are high conserved between those 3 predicted HipA family proteins .

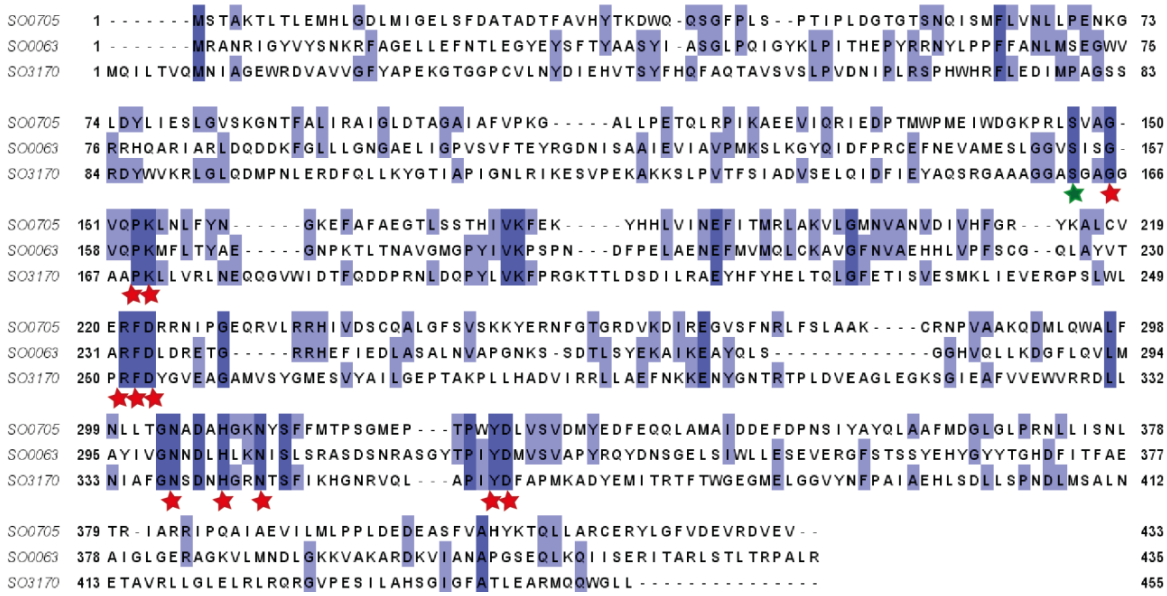


Table S1. Primers that used in this study.

Name	Sequence(5'-3')
MR01	CATATGAGTACAGCTAAAACGCTTACG
MR02	GGATCCCTACACTTCCACATCCCTGAC
MR03	CATATGAATGGGACTGATATTAAGG
MR04	GGATCCTTAATACCAGCCGTTAGTTTC
MR05	ACCGGCAATGCCCAAGCACACGGTAAAAACTACTC
MR06	GAGTAGTTTTTAC CGTGTGCTTGGGCATTGCCGGT
MR07	CATATGGCAAAAGCTAAATTTGAACG
MR08	CTCGAGGCAGCAATGATCTTAGCTACTACAC
MR09	CACCATGGGTAGTACAGCTAAAACGCTTAC
MR10	CACCATGGGTAATGGGACTGATATTAAGGCTAAG
MR11	TTAATACCAGCCGTTAGTTTCG
MR12	CTACACTTCCACATCCCTGAC
MR13	GTGCACGGCGCCGTGATTACGGCGTTGCTG
MR14	CTCGAGTTACTACACTTCCACATCCCTGACC

Table S2. Small Angle X-ray Scattering statistics

Distance Distribution	**Rg**(Å)	**Dmax**(Å)	Estimated MW
HipB	22.812 ± 0.136	79.8	26.0kDa
HipB:DNA	26.013 ± 0.349	88	38.6kDa
HipAB	37.133 ± 0.846	128	132.8kDa
HipAB:DNA	38.584 ± 0.601	132	146.3kDa

Table S3. Summary of thermodynamic parameters obtained by ITC.

Titration cell	Conc. μM	Syringe	Conc. μM	K_D nM	N	ΔH cal mol^{-1}	ΔS cal $mol^{-1}K^{-1}$
HipA	7	HipB	41	490	0.52	-3.429E4	-86.1
DNA (26bp)	6.8	HipB	63.5	290	1.13	1.381E4	76.2
HipAD306Q	18.5	HipB:DNA	150.6	300	0.46	-3.167E4	-76.5
DNA (76bp)	2	HipB	50	230	2.12	1.308E4	74.2
HipA	13	HipB Nt truncated	103	420	0.44	-3.60E4	-91.5

Chapter 4

Characterization of *Shewanella oneidensis* MR-1 HipA-HipB Toxin-Antitoxin System Assembly using Mass Spectrometry

The contents of this chapter adapted from a manuscript:

Yurong Wen, Bart Devreese. Characterization of *Shewanella oneidensis* MR-1 HipA-HipB Toxin-Antitoxin system Assembly using Mass Spectrometry.

Under preparation.

4.1 Introduction

As a powerful and increasingly important approach in the field of integrative structural biology, mass spectrometry based techniques can give crucial insights in the study of supermolecular complex assemblies (See Chapter 2 Introduction section B). After an in-depth study of the molecular mechanism by X-ray crystallography, generating an atomic resolution three-dimensional structure, and other biophysical techniques (See Chapter 3), we here further applied mass spectrometry based techniques, i.e. ion mobility mass spectrometry and chemical crosslinking combined with MS to investigate the assembly and to explore the molecular mechanism behind the regulation of the type II toxin-antitoxin module HipA-HipB in *Shewanella oneidensis* MR-1. *S.oneidensis* MR-1 is a gram negative bacteria that is readily studied for its potential application in bioremediation, microbial fuel cells and nanotechnology (El-Naggar et al., 2010; Gorby et al., 2006; Hau and Gralnick, 2007). All of these applications are highly dependent on the capacity of this bacterium to form biofilms. We previously reported that interruption of the SO0706 gene in *S.oneidensis* MR-1 after transposon mutagenesis resulted in a 50% of decrease of biofilm formation (Theunissen et al., 2010). SO0706 is homologous to the HipA toxin from *Escherichia coli* with 28% identity. HipA, a protein kinase, was demonstrated to be essential for persister formation and multidrug tolerance. While it was originally believed that it halted ribosomal activity by phosphorylating elongation factor Tu (EF-Tu) (Schumacher et al., 2009, 2012), it was recently established that its natural substrate is actually glutamyl-tRNA-synthetase. By phosporylating this enzyme at its ATP binding site, uncharged tRNA(Glu) accumulates and the stringent response is activated (Germain et al., 2013; Kaspy et al., 2013). The *hip*A gene forms an operon with HipB, a labile antitoxin that neutralizes HipA activity by forming a HipAB ternary complex. HipB is a DNA-binding protein that mediates interactions with

operators that are in the upstream of the HipAB operon to reduce HipA expression (Schumacher et al., 2009) (Chapter 3).

We here provide the results of a set of mass spectrometric experiments involving ion mobility mass spectrometry and chemical crosslinking experiments that give further insights in the *S.oneidensis* MR-1 HipA-HipB assembly, its regulation and the HipA kinase activity mechanism. Firstly, IM-MS results indicate that HipA has three different conformation states, i.e. apoHipA, Mg^{2+}-ATP bound HipA and phosphorylated HipA (pHipA). We show that binding of Mg^{2+}-ATP drives the conformational change required for autophosphorylation. Secondly, we show that the *S.oneidensis* MR-1 HipA-HipB assembly is distinct from the homologue from *E.coli* as demonstrated by IMMS and XLMS studies. Finally, our XMLS data give further support to the pLoop ejection mechanism as the basis of autophosphorylation which is consistent with the data observed from atomic resolution structures obtained from X-ray crystal diffraction data.

4.2 Results

4.2.1 HipA has three different conformational states: apoHipA, ATP bound HipA, phosphorylated HipA

The HipA family was reported to belong to the Serine/Threonine kinase family with similarity to the PI 3/4 kinase super family; it can be partially autophosphorylated *in vivo* or under incubation with ATP and Mg^{2+} *in vitro* (Correia et al., 2006). We purified the *S.oneidensis* MR-1 recombinant HipA and applied Ion Mobility MS (IMMS) to characterize it. Surprisingly, ion mobility analysis of freshly purified recombinant HipA revealed two distinct conformational populations (Figure 4.1a). The two populations from the gas phase separation were further extracted using Driftscope analysis (Table 4.1) and the mass

difference between the two populations was about 110 Da, which could correspond to a phosphorylation in addition to binding a small alkali metal. Indeed, it is previously reported that also *E.coli* HipA purification results in a partially phosphorylated population and this is actually the deactivated form which plays a role in growth arrest reversion (Schumacher et al., 2012). We further evaluated a tryptic digest of this sample used for IMMS analysis with Liquid Chromatography Mass Spectrometry (LCMS) and found that recombinant HipA is a mixture of phosphorylated HipA (pHipA) and non-phosphorylated HipA (apoHipA) (data not shown). These results further indicated that there is around 8-10% of pHipA that is phosphorylated at Ser147.

In some preparations of the recombinant HipA, we observed that the phosphorylated form was nearly absent. When analyzing these samples using IM-MS, only one population was observed indicating that the mobility difference is a direct consequence of the phosphorylation (Figure 4.1b). Not surprisingly, phosphorylation of HipA results in one positive charge reduction. The drift time of pHipA and apoHipA 13^{+} ions are 9.6ms and 13.4ms respectively which corresponds to a CCS difference of almost 950$Å^2$ CCS between the forms (Figure 4.1c). This cannot be explained by the presence of the phosphate group alone, and suggests that pHipA has a more compact conformation compared to apoHipA.

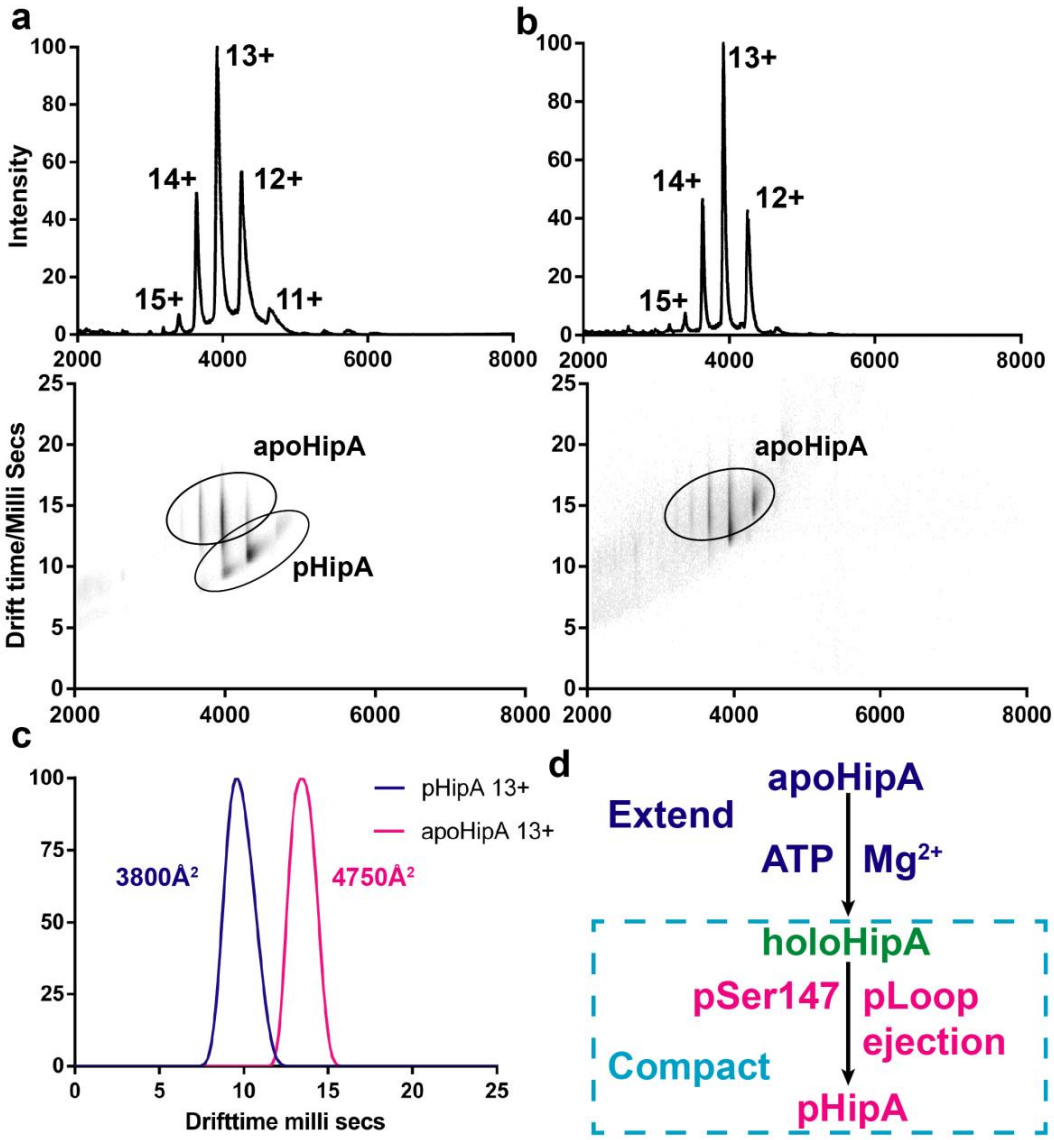


Figure 4.1. Results of IMMS measurements of two different preparations of recombinant HipA. **a.** corresponds to a population containing a mixture of apoHipA and phosphorylated HipA (pHipA), **b.** is the nonphosphorylated HipA alone (apoHipA). Both the m/z-intensity view and the m/z-drift time view are shown. In the latter, the two populations that separated in the gas phase are encircled. **c.** Drift time chromatogram of the charge 13+ pHipA and apoHipA ions with the m/z range from 3880-3960 for apoHipA and 3880-4080 for pHipA respectively. Calculation of the collision cross section indicated a value of 3800Å^2 for 13+ pHipA and 4750 Å^2 13+ apoHipA respectively. **d.** Schematic view of the three different HipA conformations. Upon binding of ATP and Mg^{2+}, apoHipA is phosphorylated transforming the protein to a more collapsed pHipA conformation to pHipA.

We further analyzed the Mg^{2+}-ATP binding effect on the HipA conformation. Freshly purified HipA was incubated with 5mM pH neutralized ATP and 10mM $MgCl_2$. All monomeric HipA shifted to a more compact conformation (Figure S1) (Table 4.1). We observed also some oligomeric HipA possibly due to the Mg^{2+} effect. There is so far no evidence for a physiological role for dimeric HipA. However, phosphorylation of HipA was reported to be a intermolecules kinetic reaction which means that the autophosphorylation strategy is based on one HipA molecules phosphorylating another, requiring the direct contact between two HipA molecules with low affinity. Analysis of the tryptic digested sample by LCMS indicated that after incubation with ATP and Mg^{2+}, the pHipA ratio increased up to approximately 45-50% of the total HipA. Crystal structure analysis indicated that the link between the C-terminal and N-terminal domains of HipA could be flexible and that there are five loops between those two domains involved in the ATP and Mg^{2+} binding to HipA. Those five loops form a pocket that could pull together the N and C domains in order to tightly anchor the Mg^{2+}-ATP binding, hence resulting in a substantial conformational change similar as observed in many kinases.

Table 4.1 Experimental and theoretical collision cross section and molecular weight

	apoHipA	**HipA:ATP:Mg**	**pHipA**	**HipB2***	**HipAB2 ATP:Mg**	**HipA2B2 ATP:Mg**
Experimental CCS(Å²)	4962	3822	3744	2629	5249	8116
PA *S.O.* CCS(Å²)	No data	3628	3576	2585	5273	8143
PA *E.coli* CCS(Å²)	No data	3276	3160	2235	4366	6717
Experimental MW (Da)	50767.6±5.3	51335.7±62.8	50878.1±18.8	21828.6±1.1 25343.7±1.6	76658.8±32.8	127940.2±46.9
Theoretical MW (Da)	50764.3	51319.5	50844.3	21839.2 25340.8	76660.3	127979.8

*The two different experimental and theoretical molecular weight indicate the state with or without His tag

Furthermore, the Mg^{2+}-ATP binding HipA trans-autophosphorylates at S147 residue which leads to an additional conformational change through pLoop ejection (Chapter 2). However, based on X-ray data, this would lead to a CCS difference of only 52 $Å^2$. Our instrument has not the resolution to distinguish such a small conformational change (Figure 4.1d). Therefore, the conformational change between the apoHipA and Mg^{2+}-ATP binding HipA results in an intermediate conformation state which serves for the HipA phosphorylation.

4.2.2 Characterization of HipA, HipB and its operator DNA interactions by Ion Mobility Mass Spectrometry reveals a distinct assembly with respect to *S.oneidensis* MR-1 and *E.coli* HipAB complexes

hipA and its upstream linked gene *hipB* form a toxin-antitoxin module which is recognized as a major factor involved in persistence (Black et al., 1994). HipB is a Helix-Turn-Helix (HTH) DNA-binding protein that functions as the antitoxin. It neutralizes HipA activity by forming a complex sequestering HipA in the nucleoid following binding to the operators upstream of the HipAB operon. Formation of this complex suppresses the transcription of the *hipA* gene (Schumacher et al., 2009, 2012). Here, an IMMS approach was applied to investigate the assembly between HipA, HipB and its operator DNA. Native ESI analysis confirmed that HipB is a dimer with molecular weight 21343.7±1.6 Da. Similarly, it was determined that this dimer interacts with 2 molecules of HipA forming a $HipA_2B_2$ heterotetramer with molecular weight 127940.2±46.9 Da. Several species including the HipB dimer, HipA monomer, $HipAB_2$ and $HipA_2B_2$ were nicely separated in the IM mode (Figure 4.2a). Additionally, we also performed ESI-MS analysis on HipB in complex with a single operator DNA showing that the $HipB_2$ dimer can interact with a single operator DNA fragment (Figure 4.2b). We also tested interaction with a DNA fragment containing 2 operator sites showing binding to 2 $HipB_2$ dimers (Data not shown).

To gain more insight into the HipA and HipB assembly, the experimental CCS restraints were generated and corrected using the equation reported before (Ruotolo et al., 2008). The theoretical CCS of *S.oneidensis* MR-1 and *E.coli* HipAB assembly was calculated by applying the Projection Approximation (PA) mode using the PDB data files (3DNV, 3FBR, 3TPE) containing the atomic resolution crystal structures (Schumacher et al., 2009, 2012). The experimental CCSs of the protein assembly are well correlated in a nice linear relationship with R^2 value 0.9961. This represents a very important value representing the solvent accessible surface area (SASA) in the globular protein with the unit of $Å^2$/Da. The SASA per Da of the apoHipA decreases from 0.098 $Å^2$/Da to 0.073 $Å^2$/Da upon the Mg^{2+}-ATP binding, revealing that the ATP and Mg^{2+} occupied region of the exposed surface is the link between the N and C domain. The experimental CCS and theoretical CCS generated with the PA model from *S.oneidensis* MR-1 crystal structure nicely fit with a tiny deviation of 2-4%. However, the theoretical CCS of the *E.coli* homologue HipA and HipB assembly are considerable different from the *S.oneidensis* MR-1 experimental CCS (Figure 4.2c). This indicated that the toxin antitoxin system HipAB assembly between *E.coli* and *S.oneidensis* MR-1 are distinct as consistent with the data observed from the atomic structure (Chapter 2).

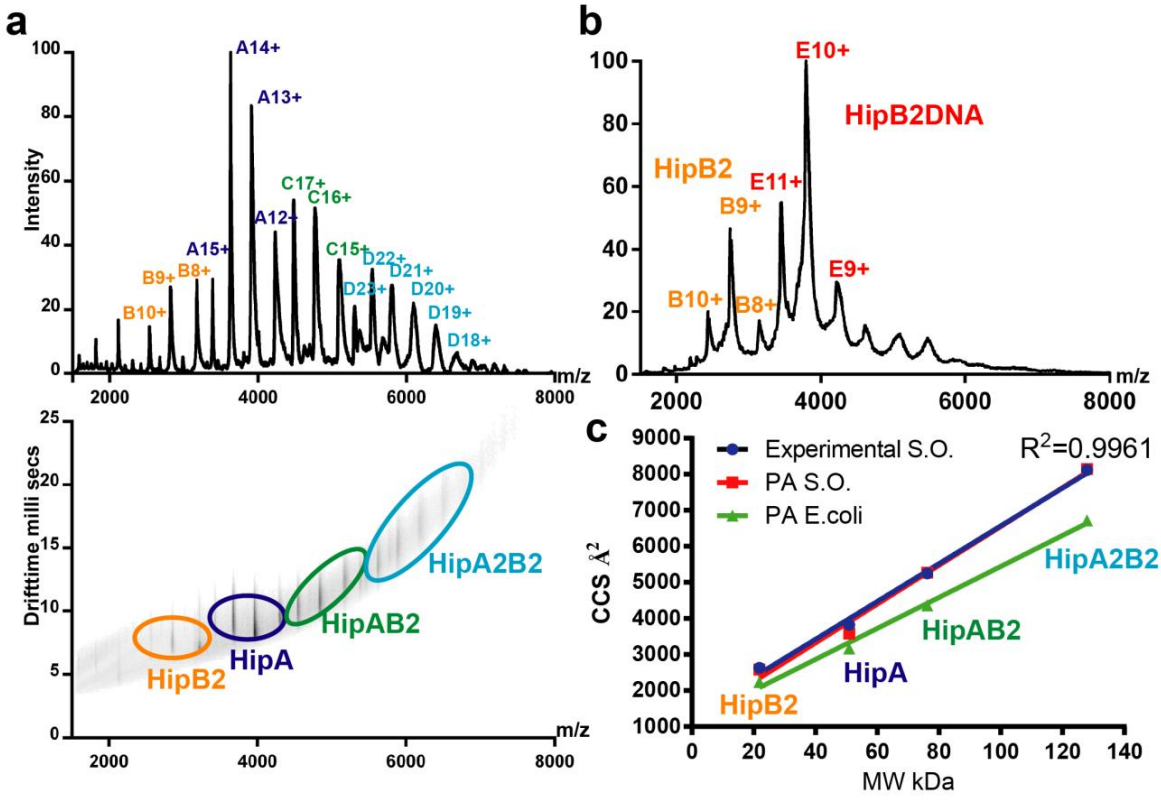


Figure 4.2. a. The IMMS measurement of the HipAB. Both the m/z-intensity view and the m/z-drift time view are shown, the charges of the 4 populations are denoted, B: $HipB_2$, A: HipA, C: $HipAB_2$, D: $HipA_2B_2$. The 4 populations are also circled in the m/z-drift time view. **b.** native ESI-MS measurement of $HipB_2DNA$ complex, the charge of the populations is denoted B: $HipB_2$, E: $HipB_2DNA$. **c.** The experimental and the theoretical PA model CCS are shown with the molecular weight correlation. The experimental $HipB_2$, HipA, $HipAB_2$, $HipA_2B_2$ CCS has significant linear correlation with molecular weight with R^2 0.9961. The *S.oneidensis* experimental CCS has nicely fit the *S.oneidensis* theoretical PA calculation and the dramatically difference between the *S.oneidensis* experimental CCS and *E.coli* PA model CCS represent the distinct assembly between the *S.oneidensis* and *E.coli*.

4.2.3 Probing the HipA kinase conformational changes upon phosphorylation using chemical crosslinking

Chemical crosslinking was used to probe the assembly of *S.oneidensis* MR-1 HipA and HipB complexes. After BS3 crosslinking, tryptic peptides were generated and analyzed using LCMS on a high resolution FTMS instrument. Data were analyzed using pLink software which revealed 6 crosslinked peptides including both inter- and intra-crosslinked peptides (Yang et al., 2012) (Table 4.2).

Table 4.2 Crosslinked peptides between HipA and HipB assembly identified by pLink

No.	**Sequence 1** **X-linked position**	**Sequence 2** **X-linked position**	**Score**	**Calc_M**	**ppm**
1	IEDPTMWPMEIWDGKPR(15) HipA(143)	DVKDIR(3) HipA(264)	7.41 E-04	2982.462	0.30166
2	RKSAALTQDVAAMLCGVTK(2) HipB(39)	KYER(1) HipA(252)	8.57 E-02	2694.42	0.62327
3	VEKGEDVYISTVFK(3) HipB(64)	KTLIR(1) HipB(57)	9.95 E-09	2380.325	0.01259
4	VYEDTLLETIMASPLNQQSLGLLIKER(25) HipB(35)	VEKGEDVYISTVFK(3) HipB(64)	4.36 E-02	4824.54	1.27446
5	VEKGEDVYISTVFK(3) HipB(64)	GSHMNGTDIK(1) HipB(-2)	9.56 E-02	2809.384	-4.66484
6	VYEDTLLETIMASPLNQQSLGLLIKER(25) HipB(35)	GSHMNGTDIK(1) HipB(-2)	8.44 E-01	4270.186	-0.89202

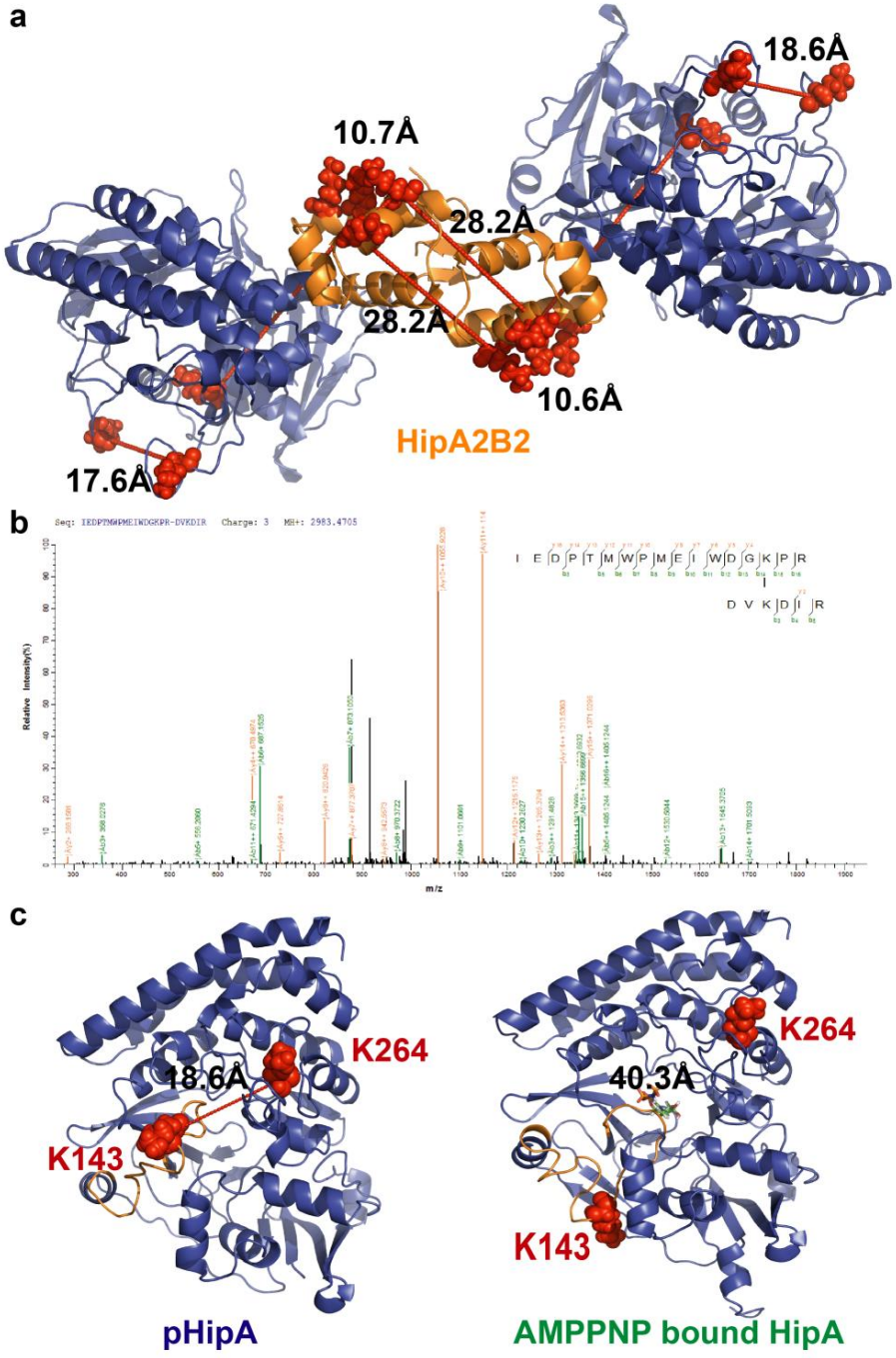


Figure 4.3. Results of crosslinking experiments using BS3. a. The crosslinked peptides are denoted on the *S.oneidensis* MR-1 $HipA_2B_2$ 3D dimensional structure accompanied with the distance restraints. **b** The ms/ms spectrum of the crosslinked peptide K143 and K264 **c.** The K143 and K264 crosslinked peptide in the pHipA and ATP-mimic binding HipA. The distances are 18.6Å and 40.3Å respectively indicating that in the AMPPNP bound HipA, the distance between these residues is too large and cross-linking is further hindered by a helical structure. The observation of the crosslinked K143 and K264 in the pHipA thus probes the pLoop ejection upon of HipA phosphorylation.

The spacer length of the BS3 crosslinker is 11.4Å and an efficient crosslink distance between two lysine residue Cα atoms is reported to be approximately 5-30Å (Leitner et al., 2012; Walzthoeni et al., 2012; Yang et al., 2012). Distance restraints are generated and nicely fit the values reported for BS3 crosslinkers (Figure 4.3a). It is noteworthy that two distinct crosslinked sites (K35 and K64) were observed from the N terminus domain of HipB, which could be due to the flexibility of the HipB N terminus domain. Surprisingly, when we applied the chemical crosslinking approach to a pHipA containing sample, an additional crosslinked peptide was observed between K143 and K264. K143 is positioned inside the reported phosphorylation loop (pLoop) of which repositioning is crucial for the activation of the kinase activity of HipA. Interestingly, this crosslink was not observed in the HipAD306Q mutant which was reported not to undergo phosphorylation. The MS/MS fragmentation spectrum of this crosslinked peptide was highly confident with most of the b and y ions generated (Figure 4.3b). Indeed, from the crystal structure of pHipA and the AMPPNP bound HipA, the distance observed between K143 and K264 are 18.6Å and 40.3Å respectively (Figure 4.3b). Those two residues are too far away to be crosslinked and, moreover, they are separated by a helix in both the apoHipA and in HipA bound to the ATP-mimicking compound AMPPNP. This indicates that the K143 and K264 crosslinked peptide observed from the pHipA describes the phosphorylation pLoop conformational change which we could not observe using ion mobility experiments.

4.3 Discussion

Mass spectrometry combined with ion mobility and chemical crosslinking was dedicated to be powerful in study of protein dynamics such as protein-ligand interaction, conformational change and transition interactions. In this study, three different conformations of recombinant HipA were observed in ion mobility MS experiments, indicating the HipA kinase activation procedure going from the apoHipA, ATP Mg^{2+} binding to autophosphorylation. ESI-IM-MS also specializes at probing the weak non-covalent protein-protein interactions; the observation of dimeric HipA upon ATP and Mg^{2+} could reflect the trans-autophosphorylation between two HipA molecules. Furthermore, the HipA kinase activity loop ejection mechanism was characterized through the chemical crosslinking mass spectrometry approach. The native ion mobility mass spectrometry data represent good correlations with the atomic resolution model provided in the previous chapter and revealed a distinct assembly between *E.coli* and *S.oneidensis* MR-1 HipAB and their DNA binding structure.

Here, the successful characterization of the *S.oneidensis* MR-1 HipA-HipB assembly and mechanism through combination b IMMS and XLMS established an example of gathering information on macromolecular assembly using MS approach. Adding MS to an integrative structural biology approach can complement the crystallography challenge such as non crystallizable and constituents' transient interactions. Recently, these mass spectrometry approaches were summarized to be applicable in pharmaceutical drug discovery and development in applications to study the conformational transition and protein ligand screening (Niu et al., 2013; Pacholarz et al., 2012). The toxin antitoxin systems are demonstrated as an attractive target for new generation antimicrobial drug discovery which may aid in the fight to major challenges under current antibiotic treatment such as persistence, multidrug tolerance and multidrug resistance etc…(Shapiro, 2013; Wen et al., 2014; Williams

and Hergenrother, 2012). The successfully probing of the conformation change between the HipA with its ligand binding and post-translational modification revealed the potential of the mass spectrometry based approach in the investigation of protein ligand complexes and may further lead to the pharmaceutical exploitation of the toxin antitoxin system as drug targets.

4.4 Experimental procedures

4.4.1 Preparation of Protein and Protein-DNA complex Samples

S.oneidensis MR-1 HipA, HipB proteins with an N-terminal His-tag were expressed in *E.coli* BL21 DE3 cells using the pET15b expression vector (Merck Millipore). Stable overexpression of the target proteins was generated by inducing the cultures for 1 h (HipA) or 4h (HipB) with 1mM isopropyl β-D-1-thiogalactopyranoside (IPTG) (Duchefa Biochemie) when cultures were grown to an OD600 of 0.6-0.7 in LB medium supplied with Carbencillin (100µg/ml) at 37ºC. Detailed overexpression and purification protocols were previously described (See Chapter 2). The double stranded operator DNA obtained from annealing two single primers 5'-ATTAGGTGTACTTATCTACACTTTTT-3', 5'-AAAAAGTGTAGATAAGTACACCTAAT-3' for the following purification (IDT). The HipAB and HipB:DNA complex were formed by mixing proper ratios and subsequently purified by Superdex 200 chromatography column (GE). The concentrations of the samples are determined with Nanodrop (Thermo Scientific).

4.4.2 ESI Ion mobility MS Analysis

Purified samples with a protein concentration of 10-20µM were buffer exchanged against 100-200mM ammonium acetate buffer pH 7 using Micro Biospin 6 columns (Bio-Rad) just prior to MS analysis. The HipA-ATP-Mg samples were obtained by adding 5mM ATP

(Sigma) and 10mM MgCl2 to the buffer exchanged sample before the MS measurement. The traveling wave ion mobility MS measurements were performed on a q-IMS-TOF instrument Synapt G1 HDMS, Waters, UK) equipped with a nanomate electrospray ionization source (Advion). The buffer exchanged samples were sprayed using a capillary voltage of 1.65-1.75kV using the type D nano-ESI chip. The IM separator was pressurized with 0.5mbar. The applied traveling wave parameters were: Ion mobility wave velocity 400m/s, wave height 12. Ions underwent TOF analysis with m/z range from 1500 or 2000 to 8000. The TOF tube vacuum pressure maintained at around 9.5×10^{-7} mbar. All the other instrument parameters were tuned by following the protocol reported before (Ruotolo et al., 2008). All the data was analyzed using Masslynx V4.1 and DriftScope V2.3 (Waters) with minimal smoothing and background subtraction.

4.4.3 Chemical Crosslinking proteolysis and LC-MS analysis of crosslinked peptides

Bis[sulfosuccinimidyl] suberate (BS3, Pierce) was the crosslinker used in the crosslink experiment. We prepared a stock concentration of 100mM or 200mM BS3 in water. For the crosslinking conditions optimization, different BS3 concentration 0mM, 0.5mM, 1mM, 2mM and 5mM, the incubation time of 15min, 30min, 45min, 60min and 120min were used. The 4 degree and room temperature condition were also tested and the final crosslinking approach was used and described below. Finally, the crosslinker BS3 concentration of 1mM or 2mM were used with 45-60mins incubation time in the room temperature. After final concentration of 1mM or 2mM of BS3 was added to the purified protein or protein complex and the mixture was incubated at room temperature for 45-60min, the reaction was quenched by the addition of Tris buffer pH7.0 to a final concentration of 50mM. The sample was subsequently digested overnight with trypsin (Porcine sequencing grade, Promega) at 37 degrees using an enzyme to protein ratio of 1:20. The sample was buffer exchanged on a Superdex peptide PC

3.2/30 column and to enrich the peptides, the fractions eluted out from 0.9-1.5ml was collected, dried and dissolved using 0.1% formic acid aqueous solution prior to LCMS analysis. The digests were analyzed by nanoESI LCMS/MS using an Agilent 1200 HPLC at a flow rate 300nL/min. A C18 Acclaim PepMap300 5µm (Thermo Scientific Dionex), trapped the peptides followed by separation on a C18 Acclaim PepMap100 (3µm) 75µm×250mm column (Thermo Scientific Dionex). Peptides were eluted with a gradient of acetonitrile. Buffer A contained 0.1% formic acid in water and buffer B was 0.1% formic acid in acetonitrile. The applied gradient was: 2%B to 10%B in 10min, 10%B to 40%B in 30min, 40% to 80% in 10min, and 80% to 2% in 5min. The column outlet was directly interfaced to a nanoflow electrospray source (Advion) equipped with D-chips connected to an LTQ-FT Fourier transform ion cyclotron resonance Mass spectrometer (Thermo). Data dependent analysis was carried out using a resolution of 100000. MS spectra were acquired over an m/z range of 150-2000 with a charge filter above 1 and the 10 highest peaks were submitted to MS/MS scans using a threshold energy of 35v for collision induced dissociation (CID). The chemical crosslink data analysis was performed using the pLink program (Yang et al., 2012).

4.5 Reference

Black, D.S., Irwin, B., and Moyed, H.S. (1994). Autoregulation of hip, an operon that affects lethality due to inhibition of peptidoglycan or DNA synthesis. J. Bacteriol. *176*, 4081–4091.

Correia, F.F., D'Onofrio, A., Rejtar, T., Li, L., Karger, B.L., Makarova, K., Koonin, E. V, and Lewis, K. (2006). Kinase activity of overexpressed HipA is required for growth arrest and multidrug tolerance in *Escherichia coli*. J. Bacteriol. *188*, 8360–8367.

El-Naggar, M.Y., Wanger, G., Leung, K.M., Yuzvinsky, T.D., Southam, G., Yang, J., Lau, W.M., Nealson, K.H., and Gorby, Y. a (2010). Electrical transport along bacterial nanowires from *Shewanella oneidensis* MR-1. Proc. Natl. Acad. Sci. U. S. A. *107*, 18127–18131.

Germain, E., Castro-Roa, D., Zenkin, N., and Gerdes, K. (2013). Molecular Mechanism of Bacterial Persistence by HipA. Mol. Cell http://dx.doi.org/10.1016/j.molcel.2013.08.045.

Gorby, Y. a, Yanina, S., McLean, J.S., Rosso, K.M., Moyles, D., Dohnalkova, A., Beveridge, T.J., Chang, I.S., Kim, B.H., Kim, K.S., et al. (2006). Electrically conductive bacterial nanowires produced by *Shewanella oneidensis* strain MR-1 and other microorganisms. Proc. Natl. Acad. Sci. U. S. A. *103*, 11358–11363.

Hau, H.H., and Gralnick, J. a (2007). Ecology and biotechnology of the genus Shewanella. Annu. Rev. Microbiol. *61*, 237–258.

Kaspy, I., Rotem, E., Weiss, N., Ronin, I., Balaban, N.Q., and Glaser, G. (2013). HipA-mediated antibiotic persistence via phosphorylation of the glutamyl-tRNA-synthetase. Nat. Commun. *4*, 1–7.

Leitner, A., Joachimiak, L. a, Bracher, A., Mönkemeyer, L., Walzthoeni, T., Chen, B., Pechmann, S., Holmes, S., Cong, Y., Ma, B., et al. (2012). The molecular architecture of the eukaryotic chaperonin TRiC/CCT. Structure *20*, 814–825.

Niu, S., Rabuck, J.N., and Ruotolo, B.T. (2013). Ion mobility-mass spectrometry of intact protein--ligand complexes for pharmaceutical drug discovery and development. Curr. Opin. Chem. Biol. *17*, 809–817.

Pacholarz, K.J., Garlish, R.A., Taylor, R.J., and Barran, P.E. (2012). Mass spectrometry based tools to investigate protein-ligand interactions for drug discovery. Chem. Soc. Rev. *41*, 4335–4355.

Ruotolo, B.T., Benesch, J.L.P., Sandercock, A.M., Hyung, S.-J., and Robinson, C. V (2008). Ion mobility-mass spectrometry analysis of large protein complexes. Nat. Protoc. *3*, 1139–1152.

Schumacher, M.A., Piro, K.M., Xu, W., Hansen, S., Lewis, K., and Brennan, R.G. (2009). Molecular mechanisms of HipA-mediated multidrug tolerance and its neutralization by HipB. Science *323*, 396–401.

Schumacher, M.A., Min, J., Link, T.M., Guan, Z., Xu, W., Ahn, Y.-H., Soderblom, E.J., Kurie, J.M., Evdokimov, A., Moseley, M.A., et al. (2012). Role of Unusual P Loop Ejection and Autophosphorylation in HipA-Mediated Persistence and Multidrug Tolerance. Cell Rep. *2*, 518–525.

Shapiro, S. (2013). Speculative strategies for new antibacterials: all roads should not lead to Rome. J. Antibiot. (Tokyo). *66*, 371–386.

Theunissen, S., De Smet, L., Dansercoer, A., Motte, B., Coenye, T., Van Beeumen, J.J., Devreese, B., Savvides, S.N., and Vergauwen, B. (2010). The 285 kDa Bap/RTX hybrid cell surface protein (SO4317) of *Shewanella oneidensis* MR-1 is a key mediator of biofilm formation. Res. Microbiol. *161*, 144–152.

Walzthoeni, T., Claassen, M., Leitner, A., Herzog, F., Bohn, S., Förster, F., Beck, M., and Aebersold, R. (2012). False discovery rate estimation for cross-linked peptides identified by mass spectrometry. Nat. Methods *9*, 901–903.

Wen, Y., Behiels, E., and Devreese, B. (2014). Toxin-Antitoxin systems: their role in persistence, biofilm formation and pathogenicity. Pathog. Dis. doi: 10.1111/2049–632X.12145.

Williams, J.J., and Hergenrother, P.J. (2012). Artificial activation of toxin-antitoxin systems as an antibacterial strategy. Trends Microbiol. *20*, 291–298.

Yang, B., Wu, Y.-J., Zhu, M., Fan, S.-B., Lin, J., Zhang, K., Li, S., Chi, H., Li, Y.-X., Chen, H.-F., et al. (2012). Identification of cross-linked peptides from complex samples. Nat. Methods *9*, 904–906.

Chapter 5

Modeling *Pseudomonas aeruginosa* Type II Secretion System XcpP and XcpQ Periplasmic Domain Assembly via Mass Spectrometry Based Integrative Structural Biology

The contents of this chapter are part of a published research article and two manuscripts under preparation.

1 Ruben Van der Meeren, **Yurong Wen**, Patrick Van Gelder, Jan Tommassen, Bart Devreese, and Savvas N Savvides. 2013. New insights into the assembly of bacterial secretins: structural studies of the periplasmic domain of XcpQ from *Pseudomonas aeruginosa*. The Journal of Biological Chemistry 288(2): 1214–25.

2 Ruben Van der Meeren*, **Yurong Wen***, Patrick Van Gelder, Bart Devreese, and Savvas N Savvides. Multiple binding interactions between XcpP and XcpQ from the type II secretion system of *Pseudomonas aeruginosa*. Under preparation.

3 Ruben Van der Meeren, **Yurong Wen**, Patrick Van Gelder, Bart Devreese, and Savvas N Savvides. XcpP from the type II secretion system from *Pseudomonas aeruginosa* dimerizes via its transmembrane helix. Under preparation.

5.1 Introduction

Pseudomonas aeruginosa is a gram negative opportunistic human pathogen involved in a wide range of human infections, especially in patients having a compromised host defense mechanism. The bacterium is identified as one of the most prominent pathogens isolated from cystic fibrosis patients and is a frequent cause of nosocomial infections (http://www.pseudomonas.com/). *P.aeruginosa* uses specific structures such as the pili to adhere to the surface of the host epithelial cells such as in human lungs . The successful adhesion and colonization of *P.aeruginosa* on the host tissue, ultimately leading to infection, highly relies on its capacity to secrete diverse virulence factors, such as exotoxins and proteolytic enzymes (Bleves et al., 2010; Davies, 2002; Gerlach and Hensel, 2007). Examples of such well characterized exoproteins in *P.aeruginosa* are Exotoxin A and pseudolysin. The interaction and transportation between the bacteria and the host organism or extracellular space are highly dependent on bacterial secretion systems. Bacterial secretion systems are multicomponent supramolecular assemblies involved in translocation of macromolecules such as effector proteins, nucleic acids and complexes across both the inner and outer membrane of bacteria.

At present, in gram negative bacteria, six different classes of secretion systems have been described, going from so called type Ⅰ secretion system (T1SS) to type VI secretion system (T6SS) (Economou et al., 2006). T2SS and T5SS are demonstrated to use a two-step process with firstly a transfer of the exoprotein into the periplasm and then to the extracellular milieu. The translocation from the cytoplasm to periplasm is established by the inner membrane universal general secretion (Sec) or twin-arginine (Tat) machineries which are responsible for transport of unfolded and folded or partially folded proteins respectively. These two secretion systems are specific by recognizing dedicated cleavable N-terminal signal peptides. T1SS, T3SS, T4SS and T6SS have been shown to use a single step strategy to transport the

exoprotein into the extracellular environment (T1SS) or into the host cell (T3SS, T4SS and T6SS). Instead of having a common signal peptide recognition mechanism, those systems are driven by a secretion signal which differs from one system to another (Chandran, 2013; Douzi et al., 2012; Fronzes et al., 2009; Korotkov et al., 2012; Miyata et al., 2013; Tseng et al., 2009) (Figure 5.1). In addition, Gram positive bacteria share some secretion systems with gram negative bacteriasuch as the ESX system (also defined as T7SS) described in *Mycobacterium tuberculosis* (Simeone et al., 2009).

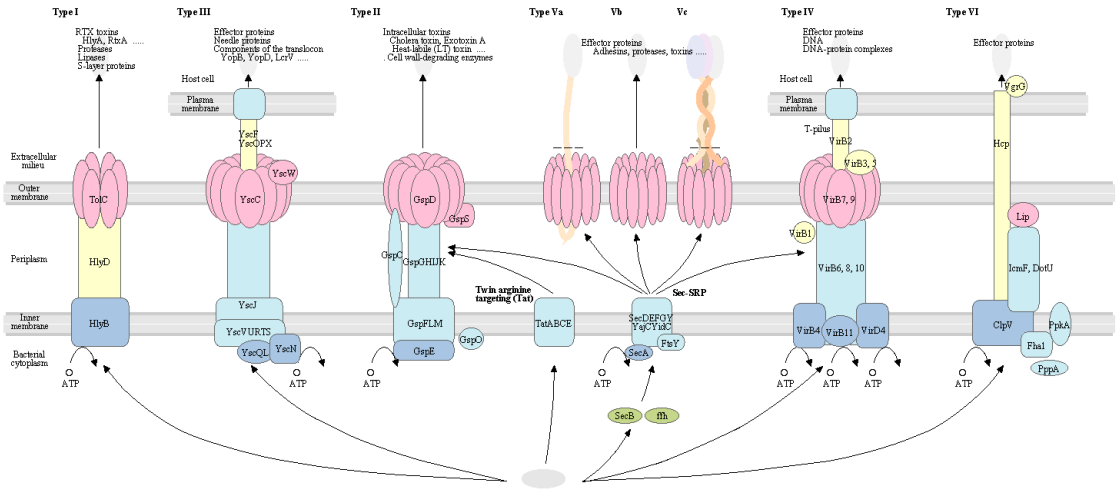


Figure 5.1 Schematic presentation of six bacterial secretion systems in gram negative bacteria. The outer membrane protein secretin is shown in pink, ATPase and chaperones are shown in blue, the pseudopilins are shown in yellow. Adapted from KEGG.

In *P.aeruginosa*, five out of these six gram negative bacterial secretion systems have been identified and characterized, except T4SS. Here we will mainly focus on the T2SS which is one of the most versatile systems used in gram negative bacteria. In general, T2SS is known as the general secretory pathway (Gsp) and has been named as GspC to M and GspO in *Escherichia coli*. However, the nomenclatures of its homologues in different strains have different names which are listed in Table1. Therefore, the *P.auruginosa* specific terms $XcpP_C$,

$XcpQ_D$, $XcpZ_M$ and $XcpA_O$ will be used in this chapter, where the subscript refers to the Gsp annotation.

T2SS usually contains 12-16 different protein subunits which are mostly encoded in a single operon. However, In *P.aeruginosa* the T2SS is encoded in two operons with the one of XcpP and XcpQ being different from that of the others XcpR to XcpZ, whereas the XcpA does not locate in the same cluster as the other subunits (Bleves et al., 2010; Sandkvist, 2001; Sandkvist et al., 2001).

Protein function	ETEC	*Vibrio cholerae*	*Aeromonas hydrophila*	*Dickeya dadantii*‡	*Klebsiella oxytoca*	*Pseudomonas aeruginosa*§	*Xanthomonas campestris*
Peptidoglycan binding		EpsA	ExeA				
Unknown function		EpsB	ExeB	OutB	PulB		
Inner-membrane platform protein, interaction with secretin	GspC	EpsC	ExeC	OutC	PulC	XcpP	XpsC‖
Outer-membrane secretin	GspD	EpsD	ExeD	OutD	PulD	XcpQ	XpsD
Secretion ATPase	GspE	EpsE	ExeE	OutE	PulE	XcpR	XpsE
Inner-membrane platform protein	GspF	EpsF	ExeF	OutF	PulF	XcpS	XpsF
Major pseudopilin	GspG	EpsG	ExeG	OutG	PulG	XcpT	XpsG
Minor pseudopilin	GspH	EpsH	ExeH	OutH	PulH	XcpU	XpsH
Minor pseudopilin	GspI	EpsI	ExeI	OutI	PulI	XcpV	XpsI
Minor pseudopilin	GspJ	EpsJ	ExeJ	OutJ	PulJ	XcpW	XpsJ
Minor pseudopilin	GspK	EpsK	ExeK	OutK	PulK	XcpX	XpsK
Inner-membrane platform protein	GspL	EpsL	ExeL	OutL	PulL	XcpY	XpsL
Inner-membrane platform protein	GspM	EpsM	ExeM	OutM	PulM	XcpZ	XpsM
Unknown function		EpsN	ExeN		PulN		
Prepilin peptidase	GspO	VcpD	TapD	OutO	PulO	XcpA (also known as PilD)	XpsO
Pilotin	YghG¶			OutS	PulS		

Table 5.1 Composition and species specific nomenclature of the T2SS. The general secretory pathway (Gsp) prefix has been suggested for use in all type II secretion systems (T2SSs), but this nomenclature from *E.coli* is not universally used. This table shows the species and systems specific names only from the best characterized bacteria species. Adapted from (Korotkov et al., 2012)

In *P.aeruginosa*, the T2SS machinery can be distinguished into 5 subassemblies: outer membrane secretin ($XcpQ_D$), the pseudopilus ($XcpT_G$, $XcpU_H$, $XcpV_I$, $XcpW_J$, $XcpX_K$), the inner membrane platform ($XcpP_C$, $XcpS_F$, $XcpY_L$ and $XcpZ_M$), the secretion ATPase ($XcpR_E$) and the prepilin peptidase ($XcpA_O$). Recently, based on numerous biochemical and structural studies, several hypothetical action mechanisms for T2SS were proposed using *E.coli* Gsp

system as the model organism. In general, the exoprotein, for example, the cholera toxin AB5 is first transported from cytoplasm through Sec to the periplasm, refolds and assembled in the functional form. The exoprotein is recognized by the secretin or by GspC and a signal is possibly transmitted to the ATPase GspE. The energy from the ATP hydrolysis drives the motions to the pseudopilin coupling protein GspL and results in the growth of the pseudopilin assembly GspG to further push out the exoprotein (Douzi et al., 2012; Korotkov et al., 2012; McLaughlin et al., 2012) (Figure 5.2). Despite plenty of biochemical and functional studies on bacterial T2SS, there exist still quite some gaps to completely understand the molecular basis of this bacterial machinery.

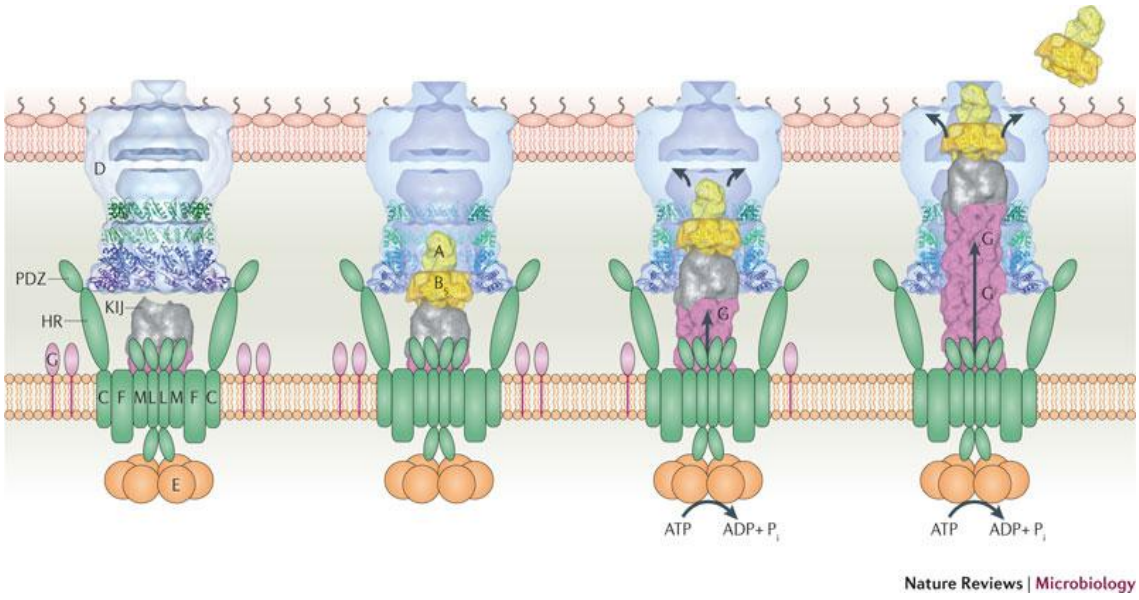


Figure 5.2. A hypothetical mechanism of action for the Type II Secretion System (T2SS). The ATPase GspE is in orange, the inner membrane platform proteins are in green, the pseudopilins GspKIJ is shown in silver and GspG is shown in pink, the outer membrane secretin GspD is in blue, and the exoprotein in this case is cholera toxin is shown in yellow. Adapted from (Korotkov et al., 2012).

Firstly, it is still yet unknown how the substrates are recognized by the T2SS machinery and what is the profile of those substrates; secondly, it is questioned how the secretion ATPase $XcpR_E$ constantly supplies the energy to drive the pseudopilins to push out the substrates;

thirdly, we do not understand how the $XcpP_C$ or secretin $XcpQ_D$ recognize the substrates and how they manage to open and close to release the substrates of secretion. For the latter, the subunit $XcpP_C$ is thought to play a key role as it bridges the inner membrane, periplasmic space and the outer membrane secretin $XcpQ_D$ (Korotkov et al., 2011). Therefore study of the assembly of $XcpP_C$ and $XcpQ_D$ is crucial to understand the transport machinery.

$XcpP_C$ is an inner membrane protein with a large periplasmic domain which can directly interact with the outer membrane secretin (Douzi et al., 2011). It includes an N-terminal transmembrane domain (TM), a homologue region domain (HR), a coiled coil domain (CC) and linkers between those domains, the HR domain which is the most conserved domain contains 6 β strands (β1-β6) (Bleves et al., 1999; Gérard-Vincent et al., 2002). $XcpQ_D$ is the outer membrane secretin, having a 35 residues N-terminal signaling sequence and 4 periplasmic N-terminal domains (N0, N1, N2, N3), and a highly conserved outer membrane C-terminal β barrel domain which is speculated to form a dodecameric cylindrical structure (Reichow et al., 2010) (Figure 5.3). Recently, the crystal structure of the periplasmic domain N0 of enterotoxigenic *E.coli* secretin GspD was also shown to be organized in a dodecameric assembly (Korotkov et al., 2013).

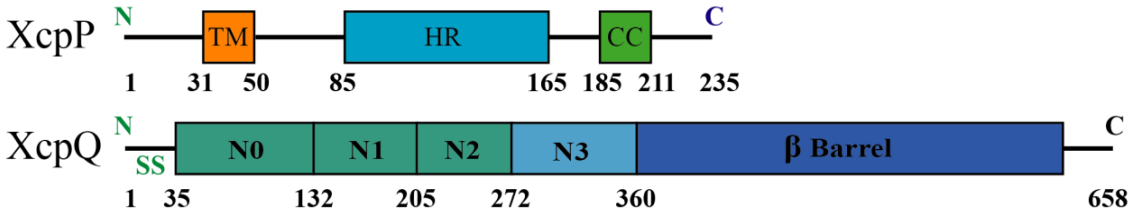


Figure 5.3. Schematic diagram of the *P.aeruginosa* $XcpP_C$ and $XcpQ_D$ protein domain structure

In the enterotoxigenic *E.coli* system, the N0 domain β1 of GspD is involved in the interaction with the inner membrane platform protein GspC HR domain β1 (Korotkov et al., 2011). However, a Surface Plasmon Resonance (SPR) and pull down interaction study has shown

that, in the *P.aeruginosa* Xcp system, the $XcpQ_D$ N-terminus can interact with $XcpP_C$ only when the N3 domain is present (Douzi et al., 2011). More recently, a cysteine scanning mutagenesis and disulfide mapping analysis study indicated that the GspC and GspD interact in multiple transient ways. OutC (GspC) HR domain β1 and β6 was reported to interact with both the N0 and N2N3 domain of OutD (GspD) in *E.coli* and *Erwinia chrysanthemi* (Wang et al., 2012). Furthermore, it was also shown that, in the *E.chrysanthemi* Out system, a 20 residues peptide from the HR region β5 and β6 of the inner membrane protein OutC can interact with two distinct regions of the outer membrane secretin OutD, which located in the N0 and N2N3 domain respectively (Login et al., 2010). In short, how exactly the $XcpP_C$ interact with $XcpQ_D$ in the periplasmic is still a debate. Therefore, we here apply multiple biophysical and biochemical approaches to investigate their assembly and dynamic association.

5.2 Results

5.2.1 Mass Spectrometry Characterization of *P.aeruginosa* $XcpP_C$ and $XcpQ_D$

To probe the interaction between $XcpP_C$ and $XcpQ_D$, multiple constructs were designed and the respective proteins were purified (Figure 5.4a). The quality of the overexpression and purification of a protein product is crucial for structural and functional investigations. In this project, we applied the techniques of peptide mass fingerprinting and intact protein mass analysis for the validation of our samples. The molecular weight of the purified protein was measured using ESI MS. For example, MS analysis of the product of the $XcpP_C$ HRCC construct indicated a good fit between the theoretical and experimental mass (Figure 5.4b). During the process of overexpression and purification, unexpected modifications may be introduced to the samples by post-translational modifications or partial polypeptide cleavage. Highly accurate mass determination of the sample product can definitely benefit further investigation. For example, after a thrombin cleavage of the $XcpQ_D$ N0N3' His-tag protein, a mass of 26548Da was obtained from ESI MS measurement apparently resulting from an unexpected miss cleavage in the N3 domain, simultaneous with cleavage of the His-tag. The difference between the theoretical and experimental mass indicated that the obtained product was actually 5 residues (273-277 TPTAR) longer than the N0N2 (35-272) domain, with 48 (278-325) of the residues miss-cleaved from the C-terminus of the N3 domain, further designated as N3' (Figure 5.4c). Surprisingly, this protein product was further successfully applied for crystallization and structure determination (Van der Meeren et al., 2013).

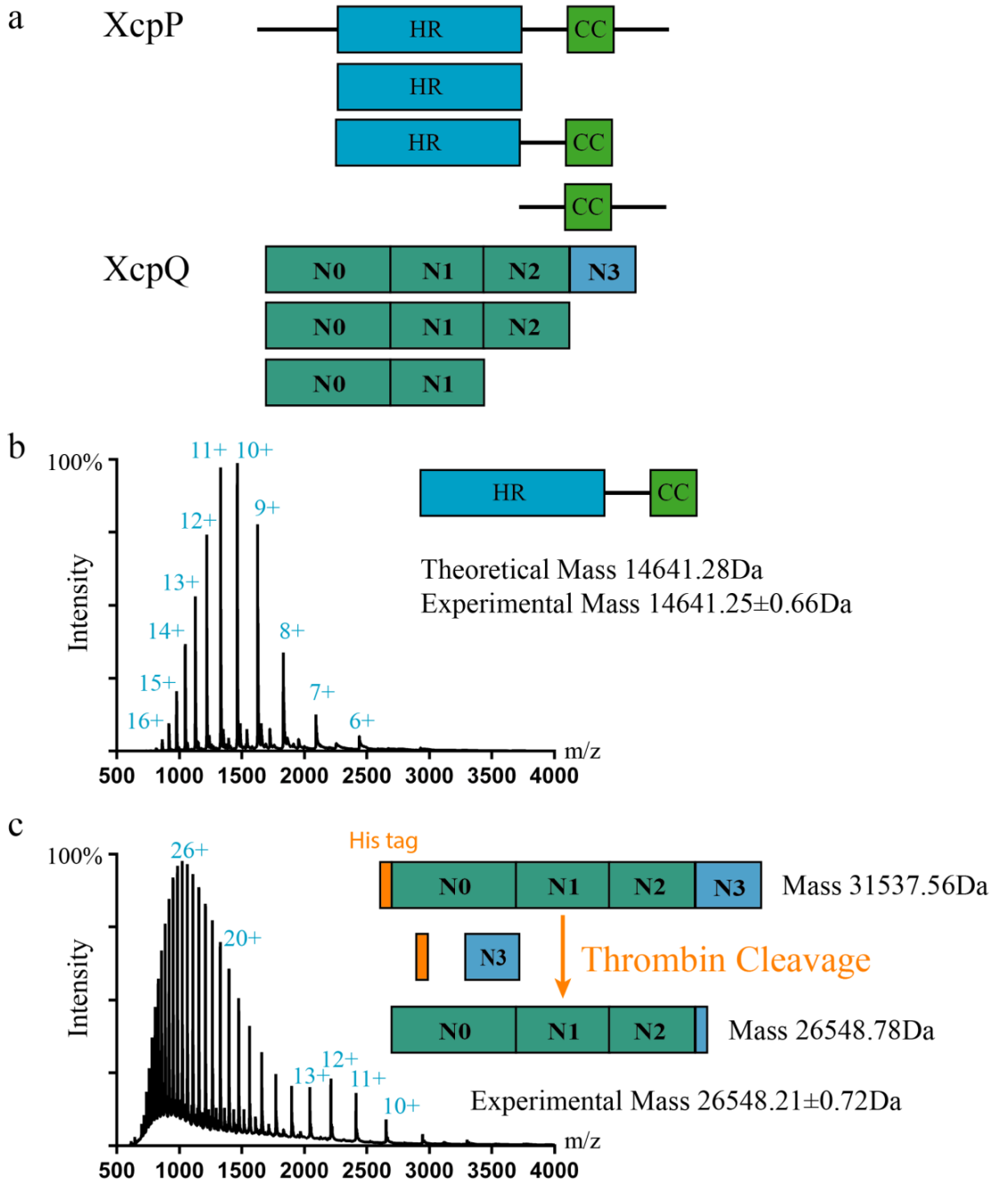


Figure 5.4. a. The schematic diagram of the *P.aeruginosa* $XcpP_C$ and $XcpQ_D$ constructs used in this study. b. Measurement under denature condition of intact $XcpP_C$ HRCC shows a good correlation of the theoretical and experimental mass. c. Measurement under denaturing conditions of intact $XcpQ_D$ N0N3 revealed there was a thrombin miss-cleavage during sample preparation.

Aside from being crucial to verify the quality of the samples that we are working with, intact ESI MS measurements under native conditions can give significant insight in the state of the structural assembly of complex such as the oligomeric state, the binding stoichiometry and even a low resolution model of the complex assembly (See also chapter 3 and 4). Here, we also applied the ESI IM MS approach for the investigation of the $XcpP_C$ and $XcpQ_D$ periplasmic domain assemblies. $XcpQ_D$ is thought to form a dodecameric β barrel structure in the outer membrane, and based on the crystal structure of monomer GspD in complex with nanobody protein Nb7, the periplasmic domain of GspD in ETEC was reported to form a C12 symmetry assembly (Korotkov et al., 2009). However, the *P.aeruginosa* XcpQN0N3' protein crystallizes in a dimeric form and it was shown that the Nb7 could have blocked the interface of the dimerization of the reported GspD structure (Van der Meeren et al., 2013). The wild type of $XcpQ_D$ N0N3' exists in a mixture of monomer and dimer (Figure 5.5a). Furthermore, it was demonstrated that the $XcpQ_D$ N0N3' can form a dimer in a concentration dependent way as proved by native ESI-IM-MS, Size Exclusion Chromatography (SEC) and Small Angle X-ray Scattering (SAXS). The mutation S210C in XcpQN0N3' which resulted in a disulfide bridge formation is also only detected in dimeric form in solution and further crystallizes in the same packing as the wild type (Figure 5.5b). In conclusion, we have demonstrated that the *P.aeruginosa* T2SS secretin $XcpQ_D$ periplasmic domain is a dimeric protein which most likely oligomerizes into a functional dodecameric assembly with C6 symmetry (Van der Meeren et al., 2013).

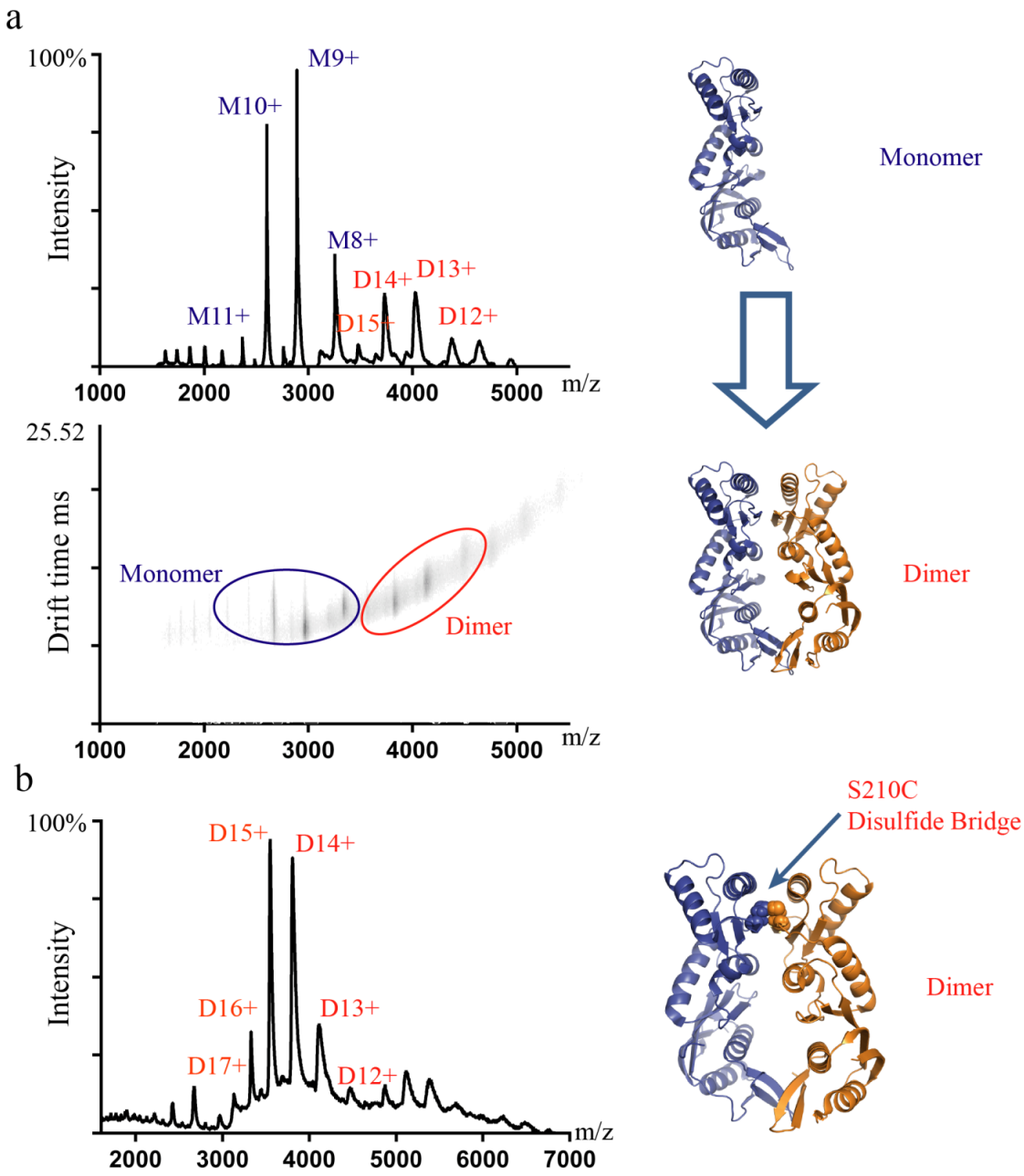


Figure 5.5. a. The *P.aeruginosa* $XcpQ_D$ N0N3' is a mixture of monomer and dimer in solution indicated in both m/z intensity and m/z drifttime view. The M and D denote the charge of the monomer and dimer b. The S210C mutant of *P.aeruginosa* $XcpQ_D$ N0N3' forms a dimer in solution, the S210C mutation site is shown in sphere between the interfaces of the dimerization.

5.2.2 Study of the *P.aeruginosa* XcpP and XcpQ Assembly using Chemical Crosslinking and Mass Spectrometry

Chemical crosslinking combined with mass spectrometry has become one of most successful approaches in the characterization of macromolecular assemblies (Chapter 3). Here we applied different crosslinkers to investigate the interaction between the $XcpP_C$ and $XcpQ_D$ periplasmic domains. Initially, formaldehyde crosslinking experiments with different constructs of *P.aeruginosa* $XcpP_C$ and $XcpQ_D$ were performed. It is shown that $XcpP_C$ HR, HRCC and $XcpQ_D$ N0N2 but not the $XcpP_C$ CC can form higher oligomeric states lane 2, 4, 5 and 3 respectively. Furthermore, $XcpP_C$ HR, $XcpP_C$ HRCC, but not $XcpP_C$ CC could interact with $XcpQ_D$ N0N2 using formaldehyde crosslinking lane 6, 8 and 7 respectively. All the crosslinked bands were further verified by peptide mapping fingerprint (Figure 5.6).

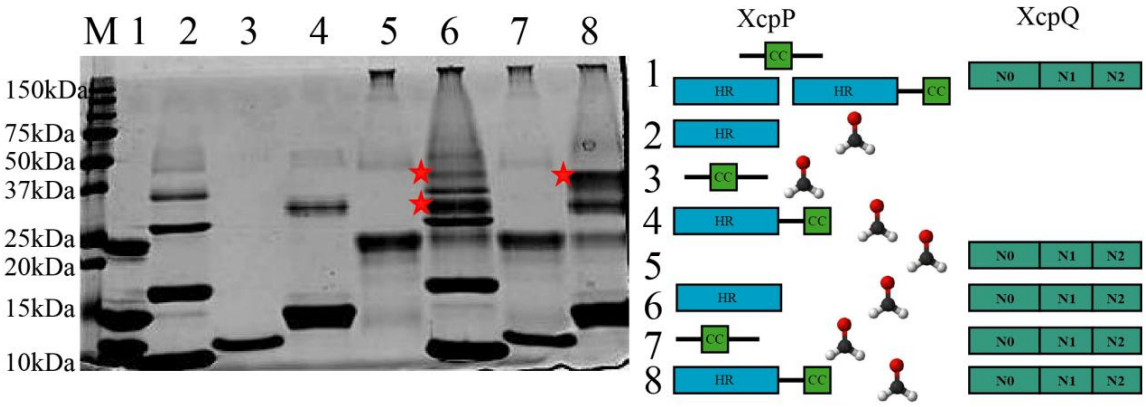


Figure 5.6. The formaldehyde crosslinking experiment with different constructs of *P.aeruginosa* $XcpP_C$ and $XcpQ_D$. The right panel denotes the product used for the crosslinking experiment. The red stars indicate the bands corresponding to the crosslinked products.

The formaldehyde crosslinking experiment reveals that the $XcpP_C$ and $XcpQ_D$ periplasmic domains can exist in an oligomeric state in solution, which also fits the results obtained by native mass spectrometry (Data not shown for $XcpP_C$). It is also shown that the $XcpP_C$ using

its HR domain but not its CC domain can interact with the $XcpQ_D$ N0N2 domain in the absence of the N3 domain. This is surprising as a previous report stated that the $XcpP_C$ interacts with $XcpQ_D$ only when the full N3 domain is present (Douzi et al., 2011).

After successful observation of the interaction between *P.aeruginosa* $XcpP_C$ and $XcpQ_D$ periplasmic domains from the formaldehyde crosslink experiments, we further refined the approach to identify the crosslink site, and to build the assembly network. Building on the recent development of algorithms for large scale crosslinked peptide identification and improvement of false discovery rates, we applied BS3 and DSS as the crosslinkers for further investigation (See also chapter 3). The spacer lengths of the BS3 and DSS crosslinkers are 11.4Å and the efficient crosslink distance between two lysine residue Cα atoms are reported approximately from 5-30Å (Leitner et al., 2012; Walzthoeni et al., 2012; Yang et al., 2012). The main procedures that are involved in the chemical crosslink mass spectrometry approach are featured in Figure 5.7. We have carried out an *in vitro* crosslink strategy using the different purified protein constructs as well as an *in vivo* crosslink strategy which directly applies the crosslinker to the whole *P.aeruginosa* PAO1 cell and the cell membrane. Here we will only present the results of the *in vitro* crosslink, since the *in vivo* crosslink is still under way.

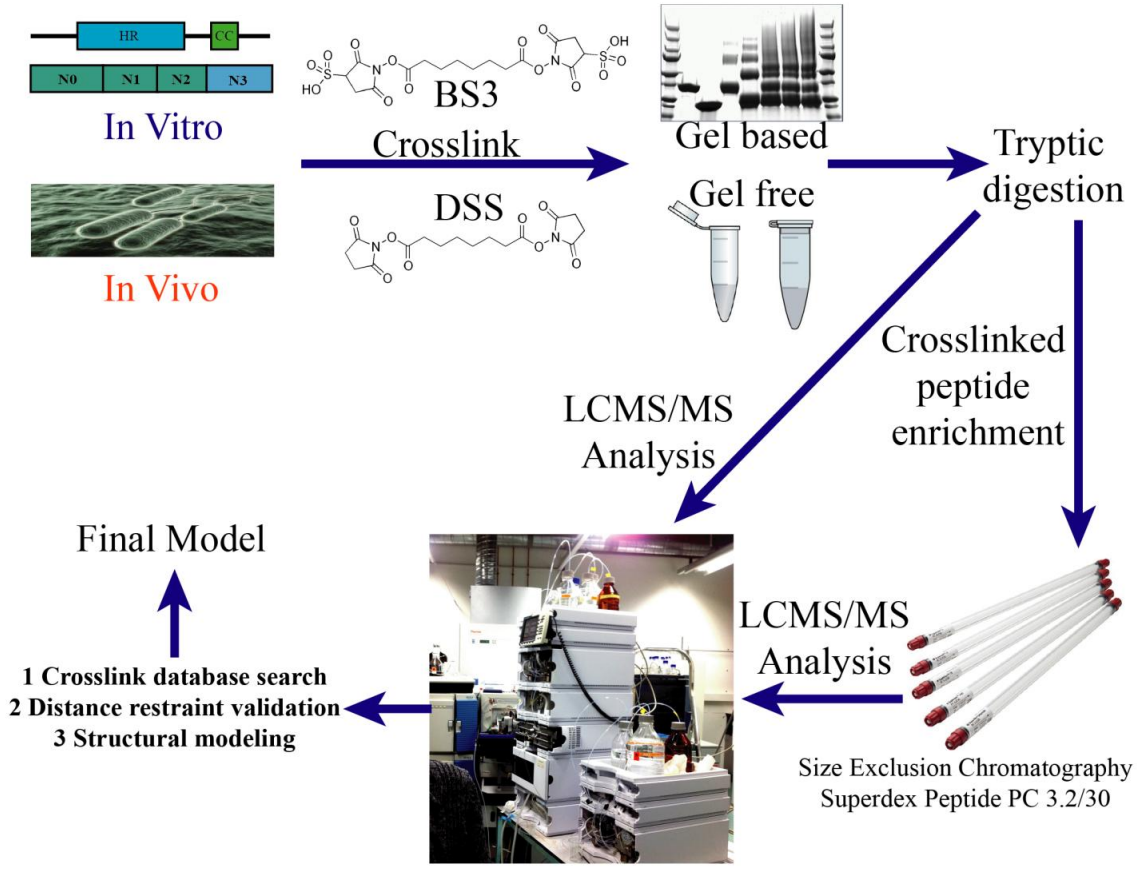


Figure 5.7. The schematic view of the chemical crosslinking mass spectrometry approach.

The crosslink database search was mainly performed using pLink and pFind software platforms and the homology models were generated using the Swiss-Model workspace (http://swissmodel.expasy.org/). A total of 987 MS/MS spectra containing crosslinked peptides ms/ms spectrum where obtained and could be grouped in 33 peptides from the *P.aeruginosa* $XcpP_C$ and $XcpQ_D$ periplasmic domains using BS3 as the crosslinker. Those crosslinked peptides were further screened and verified, with 16 of the most significant (FDR<5%) ones being used for the structural modeling (Table 5.1).

Protein Position	Sequence 1 (X-linked Residue No.)	Sequence 2 (X-linked Residue No.)	Domains
XcpQ intra and inter crosslinked peptides			
XcpQ(292)-XcpQ(263)	HNDAKTLAETLGQISEGMK(5)	AKLVQLAQSLDTPTAR(2)	N3-N2
XcpQ(263)-XcpQ(263)	AKLVQLAQSLDTPTAR(2)	AKLVQLAQSLDTPTAR(2)	N2-N2
XcpQ(263)-XcpQ(236)	AKLVQLAQSLDTPTAR(2)	GQAKGAAGAQVIADAR(4)	N2-N2
XcpQ(236)-XcpQ(236)	GQAKGAAGAQVIADAR(4)	GQAKGAAGAQVIADAR(4)	N2-N2
XcpQ(236)-XcpQ(204)	GQAKGAAGAQVIADAR(4)	QLDQKGSHDYSVINLR(5)	N2-N2:N1*
XcpQ(131)-XcpQ(84)	IVPNAEAKTEAGGGQSAPDRLETR(8)	VKGQVSVVSK(2)	N0-N0
XcpQ(84)-XcpQ(84)	VKGQVSVVSK(2)	VKGQVSVVSK(2)	N0-N0
XcpP intra and inter crosslinked peptides			
XcpP(143)-XcpP(128)	YAVGGEISDGVKLHAVYR(12)	QGDKPHR(4)	HR-HR
XcpP(143)-XcpP(116)	YAVGGEISDGVKLHAVYR(12)	GSFVQSDPKLSSAIIQR(9)	HR-HR
XcpP(128)-XcpP(128)	QGDKPHR(4)	QGDKPHR(4)	HR-HR
XcpP(116)-XcpP(128)	GSFVQSDPKLSSAIIQR(9)	QGDKPHR(4)	HR-HR
XcpP(116)-XcpP(128)	GSFVQSDPKLSSAIIQR(9)	LSSAIIQRQGDKPHR(12)#	HR-HR
XcpP(116)-XcpP(128)	LFGTSAQDPNAPPPATNLDLVLKGSFVQSDPK(23)#	QGDKPHR(4)	HR-HR
XcpP and XcpQ inter crosslinked peptides			
XcpQ(263)-XcpP(128)	AKLVQLAQSLDTPTAR(2)	QGDKPHR(4)	N2-HR
XcpQ(236)-XcpP(128)	GQAKGAAGAQVIADAR(4)	QGDKPHR(4)	N2-HR
XcpQ(84)- XcpP(128)	VKGQVSVVSK(2)	QGDKPHR(4)	N0-HR

Table 5.2 the crosslinked peptides between *P.aeruginosa* XcpP and XcpQ periplasmic domain

represents the detection of one miss cleavage site in the peptide

* In the hinge domain that links XcpQN2 and XcpQN1

The *P.aeruginosa* XcpQN0N2 crystal structure (PDB: 4E9J) and the XcpPHR domain structure model obtained from the Swiss-Model workspace using the *E.coli* GspC structure (PDB: 3OSS or 2LNV) as the template are used for the structural modeling. The XcpP HR domain, predicted to contain a 6 β strand structure is the most conserved region in this T2SS inner membrane protein and is the crucial domain for binding the outer membrane secretin XcpQ periplasmic domain (Figure 5.8).

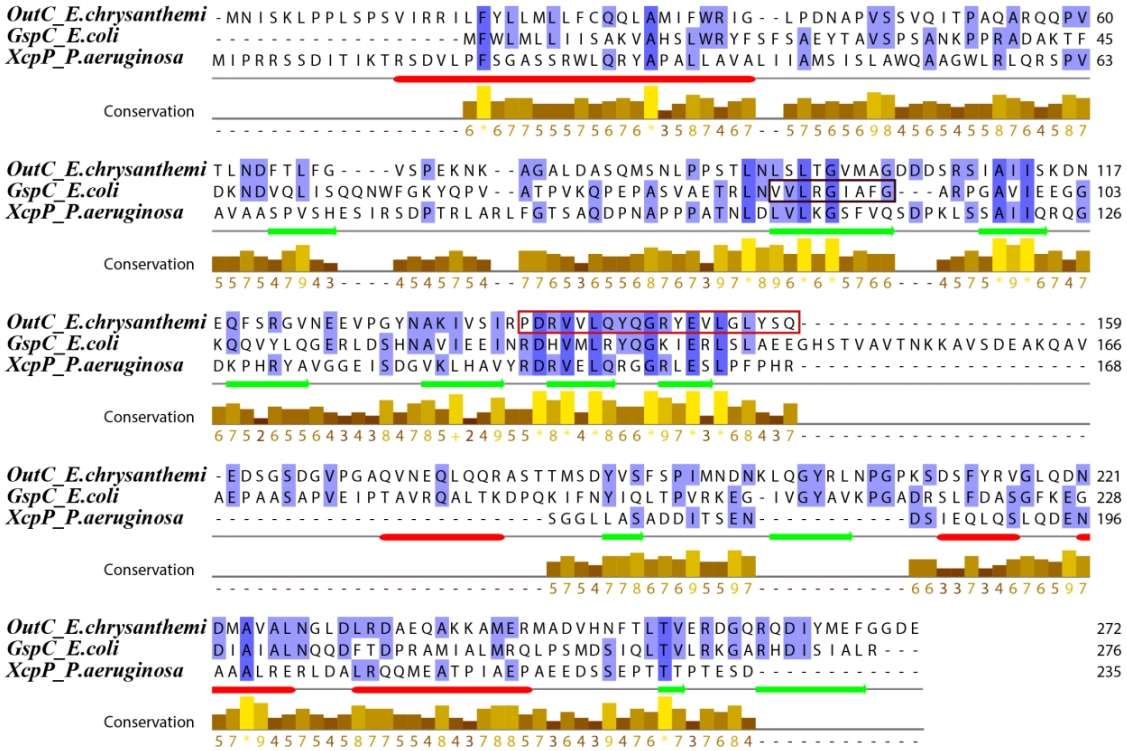


Figure 5.8. Sequence alignment of *E.chrysanthemi* OutC, *E.coli* GspC and *P.aeruginosa* XcpP. The red square box from *E.chrysanthemi* OutC represents the secretin interaction segment (139-159, β5 and β6) that interacts with two distinct regions with OutD N0 and N2N3. The black square box from *E.coli* GspC HR domain β1 represents the interface with GspD N0 domain β1.

The $XcpP_C$ HR domain region β3 K128 is observed to be crosslinked with both residues from the $XcpQ_D$ N0 and N2N3 domain. This crosslink site could manage the interaction interface

to access the fragment that is highly homologous with the OutC 20 residue secretin interaction peptide (SIP) (139-159, β5 and β6) which was reported to be essential to interact with two distinct domains of OutD (Login et al., 2010) as well as the GspC HR domain β1 to interact with GspD N0 domain (Korotkov et al., 2011) (Figure 5.8).

The crosslinked sites were first validated with the known structural models of the $XcpP_C$ HR domain and $XcpQ_D$ N0N2 domains. The intramolecular cross linked residues from $XcpQ_D$ are at two interfaces of the $XcpQ_D$ N0N2 dimer. and all the crosslinked lysine residues are at a distance within the range amenable for crosslinking according to reported values (Hall et al., 2012; Leitner et al., 2012) (Figure 5.9a). The crosslink sites network is further shown in the overall domain view, with both inter- and intramolecules crosslink sites of $XcpP_C$ and $XcpQ_D$ being denoted. Especially noteworthy are the 3 intermolecular crosslinked sites, the ones between $XcpP_C$ HR K128 and, respectively, the $XcpQ_D$ N0 domain K84 and XcpQ N2N3 domain K236 and K263 (Figure 5.9b). Since it is impossible for the $XcpP_C$ HR domain to interact with two distinct domains of $XcpQ_D$ simultaneously, this suggests that there are two different interaction interfaces for the periplasmic $XcpP_C$ and $XcpQ_D$. This is indicative for the two states of binding states (Figure 5.9c). Examples of some crosslinked peptides MS/MS spectra generated from pLink are denoted (Figure 5.9d) (See also Chapter two: Introduction Section B).

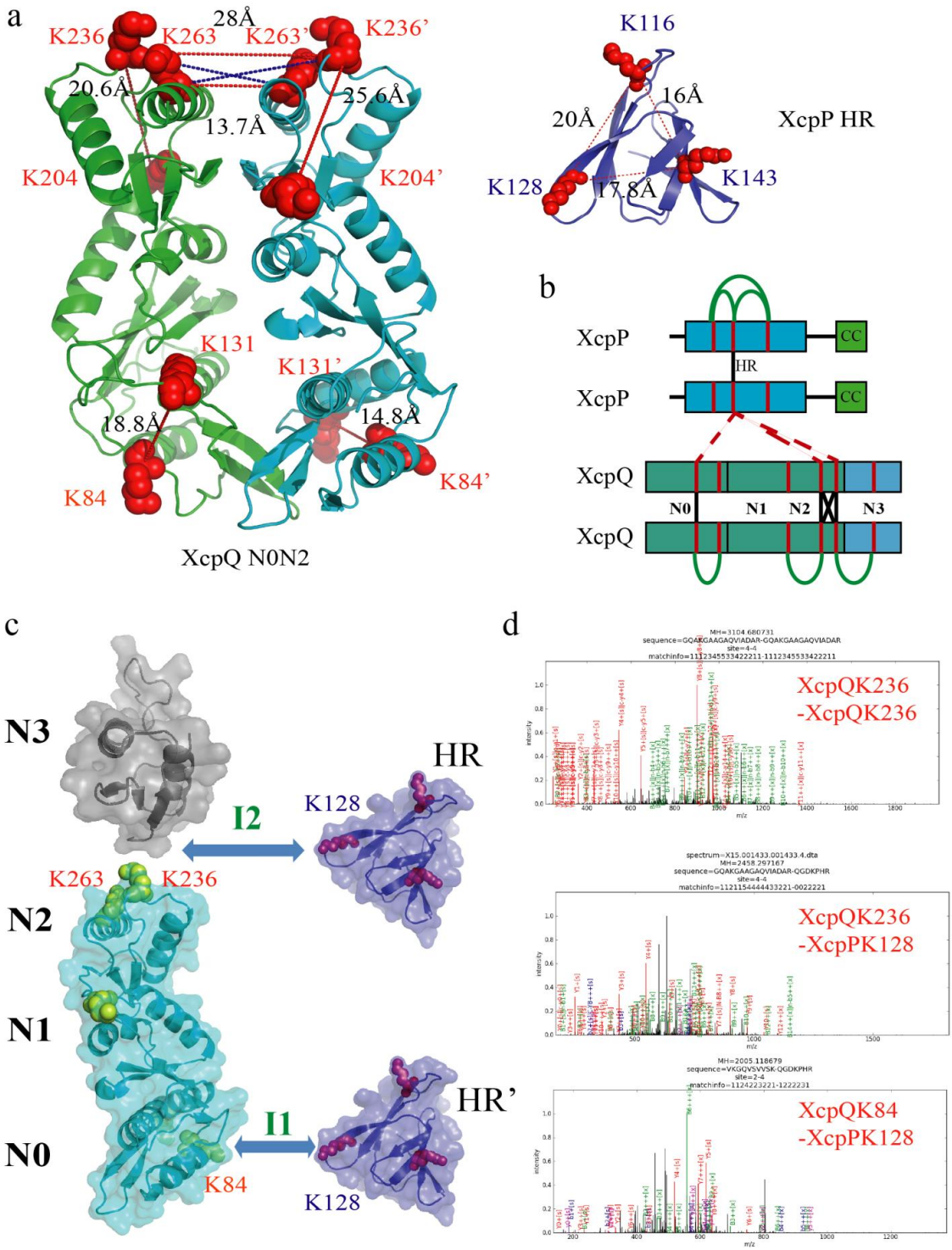

a
28Å
K236
K263
K263'
K236'
20.6Å
25.6Å
13.7Å
K204
K204'
K131
K131'
18.8Å
14.8Å
K84
K84'
XcpQ N0N2
K116
16Å
20Å
XcpP HR
K128
17.8Å
K143
b
XcpP
CC
HR
XcpP
CC
XcpQ
N0
N1
N2
N3
XcpQ
c
N3
HR
I2
K128
K263
K236
N2
N1
N0
I1
HR'
K84
K128
d
XcpQK236
-XcpQK236
XcpQK236
-XcpPK128
XcpQK84
-XcpPK128

Figure 5.9. a. Validation of the crosslinked peptide distance restraints with the known crystal structure of $XcpQ_D$ N0N2 and the homologue model $XcpP_C$ HR, the lysine residues involved in the crosslinking are shown in red sphere accompanied with the distances. **b.** Schematic view of the crosslinked sites, the black straight lines indicate the inter crosslink in $XcpP_C$ or $XcpQ_D$, the red dash lines indicate the inter crosslink between $XcpP_C$ and $XcpQ_D$. **c.** Two distinct interaction interfaces of the $XcpP_C$ HR domain to the $XcpQ_D$ N0 domain and $XcpQ_D$ N2N3 domain are indicated in green I1 and I2. **d.** Examples of the crosslinked peptides MS/MS spectra generated with pLink.

5.3 Discussion and Perspective

We have carried out a mass spectrometry based approach to study the assembly of *P.aeruginosa* $XcpP_C$ and $XcpQ_D$. Although the analysis is not comprehensive, multiple features are observed concerning the machinery assembly. The crosslinking data suggested multiple interaction interfaces in the periplasmic domains between $XcpP_C$ and $XcpQ_D$. This is compatible with the result that we observed from other affinity and mutation studies (Ruben Van der Meeren, Pers. information).

In the T2SS, the secreted proteins are first transported from the cytoplasm to the periplasm through the Tat or Sec inner membrane machineries. The correctly folded and assembled exoproteins are then further transported out to the extracellular space via the Xcp machinery. Via recognition through the inner membrane platform $XcpP_C$ or the secretin $XcpQ_D$, the exoproteins have to be delivered from the periplasmic space inside the cylindrical architecture of the outer membrane $XcpQ_D$ (Douzi et al., 2012). However, how the assembly between $XcpP_C$ and $XcpQ_D$ transits in order to translocate the exoprotein from the "out" to "in" of the cylindrical architecture is still a matter of debate.

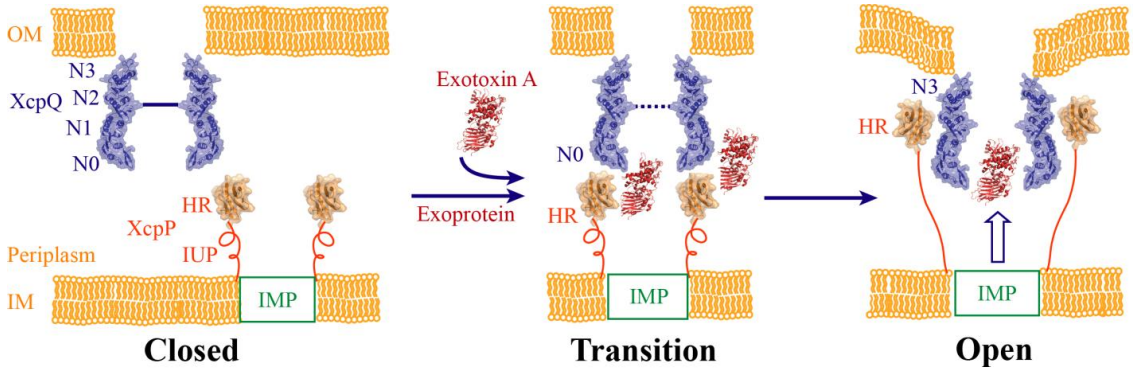


Figure 5.10. The transition interaction model of type II secretion system periplasmic domain in *Pseudomonas aeruginosa*. The secretin XcpQ channel is closed in the absence of XcpP binding, upon the exoprotein appropriate folded in the periplasmic space in example with Exotoxin A(PDB: 1IKQ), it drives the binding of XcpP with XcpQ N0 domain together with exoprotein, further the channel of secretion XcpQ is open with the XcpP binding to the N3 domain, the exoprotein is transported out with the signal of ATPase and the push force from the pseudopilins or the retraction of outer membrane secretin XcpQ.

We hypothesize that the $XcpP_C$ and $XcpQ_D$ periplasmic assembly could consist of "open" and "closed" states of the cylindrical architecture to let the exoprotein in. The exoproteins reported to interact with both the $XcpP_C$ and $XcpQ_D$ could be followed by the trigger of the assembly of the inner membrane platform and the outer membrane secretin. The two distinct

interaction interfaces of the $XcpP_C$ HR domain to $XcpQ_D$ N0 and $XcpQ_D$ N2N3 determined by crosslinking may then represent the "open" and "closed" state of T2SS in the periplasmic (Figure 5.10). However, how the HR domain β1 and β6 are involved in this two states binding as described for the *E.chrysanthemi* Out protein systems remains elusive. Given that the XcpQ N1-N2-N3 domains are connected with a very flexible "hinge" structure, the interaction between $XcpP_C$ HR and $XcpQ_D$ N2N3 can be achieved through the extension of the $XcpQ_D$ "hinge" between the N1-N2-N3 domains or a stretch of the XcpP intrinsically unstructured protein (IUP) domain between TM and HR domain and the switch binding between β1 or β6 of the HR domain caused rotation. The transient interactions are very challenging to investigate only with crystallography, especially associated with membrane proteins. Molecular Dynamics (MD) simulation and macromolecular docking is a widely used technique to obtain information about the time dependent evolution of conformational changes and transition (Adcock and McCammon, 2006). We will apply MD simulation and macromolecular docking (such as HADDOCK) and integrate the distance restraints that we generated from the mass spectrometry based approach to model the possible surface interaction of the $XcpP_C$ HR domain with $XcpQ_D$ N0 domain and $XcpQ_D$ N2N3 domain respectively in order to gain further insight about their assemblies (See Chapter two: Introduction Section B).

The in vivo *P.aeruginosa* crosslink experiments are ongoing and may generate more precise physiological level information about the T2SS machinery.

5.4 Experimental Procedures

5.4.1 Protein overexpression and purification

The *P.aeruginosa* XcpP and XcpQ protein overexpression and purification were mainly carried out by Ruben Van der Meeren with assistance from the author of this thesis. For more details, see (Van der Meeren et al., 2013). In general, the DNA of XcpP and XcpQ domains were amplified by PCR from genomic DNA using appropriate primers. The PCR product was ligated into a pET15b+ expression vector (Invitrogen) using the NdeI and BamHI restriction sites. For expression, *E. coli* BL21(DE3) transformed with pET containing the target DNA sequences were grown in LB medium supplied with carbenicillin (100µg/mL) at 37ºC and were induced at OD600nm of 0.6-1.0 with 1mM isopropyl β-D-1-thiogalactopyranoside (IPTG) (Duchefa Biochemie) followed by growth for 5 hours and harvesting by centrifugation. The ensuing cell pellet was resuspended (500mM NaCl, 20mM Tris pH 8.0) in the presence of protease inhibitors (Complete®, Roche), and were lysed by sonication. The cell debris was pelleted by centrifugation at 75,000g for 30 minutes and the supernatant was filtered using a syringe filter cap (0.22µm). The clarified lysate was loaded onto a Ni-NTA column (Qiagen) pre-equilibrated with buffer A (20mM Tris-HCl pH 8, 500mM NaCl, 10mM Imidazole), washed with buffer B (20mM Tris-HCl pH 8, 500mM NaCl, 50mM Imidazole) and eluted with buffer C (20mM Tris-HCl pH 8, 500mM NaCl, 250mM Imidazole). To remove the His6-tag, fractions of purified protein were pooled and concentrated on a Vivaspin 15R column 10,000 MWCO (Sartorius Stedim) to ~1mL before diluting the sample 10 times with thrombin digestion buffer (150mM NaCl, 20mM Tris pH 8 and 2.5mM CaCl2). One unit of biotinylated thrombin (Novagen) was added per mL of diluted sample and the cleavage reaction was allowed to continue for overnight in the dark at room temperature until protein was cleaved completely as evaluated by SDS-PAGE. Biotinylated thrombin was removed from solution by adding strepavidin agarose (Novagen)

followed by a centrifugation step (10 minutes, 4000 x g) and filtration using a syringe filter cap (0.22µm). The sample was subsequently subjected to size exclusion chromatography (SEC) on a Superdex-75 column (GE Healthcare) equilibrated with 150mM NaCl, 20mM HEPES pH 7.5. Fractions containing pure protein (>95% purity as judged by SDS-PAGE), were pooled and concentrated for further study. The concentrations of the samples are determined with Nanodrop (Thermo Scientific).

5.4.2 Chemical Crosslinking, Proteolysis and FTICR

Crosslinker BS3 or DSS (Pierce) was prepared at a stock concentration of 100 or 200mM in water or DMSO respectively. For the crosslinking conditions optimization, different BS3 concentration 0mM, 0.5mM, 1mM, 2mM and 5mM, the incubation time of 15min, 30min, 45min, 60min and 120min were used. The 4 degree and room temperature crosslinking conditions were also tested and the final crosslinking approach was used as described below. A final concentration of 1mM or 2mM of BS3 or DSS was added to the purified protein or protein complex (1-2mg/ml) with the final DMSO volume ratio being lower than 2%(for DSS). The reaction mixture was incubated at room temperature for 45-60min and then the reaction was quenched by the addition of a final concentration of 50mM Tris buffer. The sample was subsequently digested overnight at 37 degrees with trypsin (Porcine sequencing grade, Promega) using an enzyme to protein ratio of 1:20. The sample was buffer exchanged on a Superdex peptide PC 3.2/30 column and the peptides were dissolved in 0.1% FA prior to liquid chromatography mass spectrometric (LC-MS) analysis. The digests were analyzed by nano LC-MS/MS using an Agilent 1200 HPLC at a flow rate 300nL/min. A C18 Acclaim PepMap300 5µm column (Thermo Scientific Dionex), trapped the peptides, which was followed by a separation on a C18 Acclaim PepMap100 3µm, 75µm×250mm column (Thermo Scientific Dionex). Peptides were eluted applying a gradient of acetonitrile with

buffer A (0% v/v acetonitrile in 0.1% formic acid) and buffer B (100% v/v acetonitrile in 0.1% formic acid). The gradient went from 2%B to 10%B in 10min, from 10%B to 40%B in 30min, from 40%B to 80%B in 10min, and back from 80%B to 2%B in 5min. The column outlet was directly interfaced via a nanoflow electrospray source (Advion) to an LTQ-FT Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR-MS, Thermo). Data dependent analysis was carried out using a resolution of 100000 followed by a 10 MS/MS spectra scan. MS spectra were acquired over an m/z range of 150-2000, with a charge filter above 1; MS/MS scans were applied using threshold energy of 35 eV for collision induced dissociation (CID). The chemical crosslink data was submitted to the pLink program for analysis of cross-linked peptides (Yang et al., 2012).

5.4.3 Ion Mobility Mass Spectrometry Analysis

Purified samples with a protein concentration of 10-20μM were buffer exchanged into 100-200mM ammonium acetate buffer using Micro Biospin 6 columns (Bio-Rad) just prior to MS analysis. The denatured condition was generated with a final concentration of 50% acetonitrile in 0.1% formic acid. The ion mobility and mass measurements were carried out on a quadrupole ion mobility (IM) TOF MS instrument (Synapt G1 HDMS, Waters, UK) equipped with a Nanomate source (Advion). The buffer exchanged samples were sprayed using a capillary voltage of 1.65-1.75 kV. The IM separator was pressurized at 0.5mbar. Data were acquired using the traveling waves: ion mobility wave velocity 400-600m/s and a wave height 12-15. Ions underwent TOF analysis with an m/z separation range from 1000 to 8000, under a vacuum pressure maintained at around 9.5×10^{-7} mbar. All the other instrument parameters were tuned following the protocol reported before (Ruotolo et al., 2008). All the data was analyzed under Masslynx V4.1 and DriftScope V2.3 with minimal smoothing and background subtraction.

5.5 References

Adcock, S. A, and McCammon, J.A. (2006). Molecular dynamics: survey of methods for simulating the activity of proteins. Chem. Rev. *106*, 1589–1615.

Bleves, S., Gérard-vincent, M., Lazdunski, A., Filloux, A., Bleves, S., Ge, M., and Filloux, A. (1999). Structure-Function Analysis of XcpP , a Component Involved in General Secretory Pathway-Dependent Protein Secretion in *Pseudomonas aeruginosa*. J. Bacteriol. *181*, 4012–4019.

Bleves, S., Viarre, V., Salacha, R., Michel, G.P.F., Filloux, A., and Voulhoux, R. (2010). Protein secretion systems in *Pseudomonas aeruginosa*: A wealth of pathogenic weapons. Int. J. Med. Microbiol. *300*, 534–543.

Chandran, V. (2013). Type IV secretion machinery: molecular architecture and function. Biochem. Soc. Trans. *41*, 17–28.

Davies, J.C. (2002). *Pseudomonas aeruginosa* in cystic fibrosis : pathogenesis and persistence. Paediatr. Respir. Rev. *0550*, 128–134.

Douzi, B., Ball, G., Cambillau, C., Tegoni, M., and Voulhoux, R. (2011). Deciphering the Xcp *Pseudomonas aeruginosa* type II secretion machinery through multiple interactions with substrates. J. Biol. Chem. *286*, 40792–40801.

Douzi, B., Filloux, A., and Voulhoux, R. (2012). On the path to uncover the bacterial type II secretion system. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *367*, 1059–1072.

Economou, A., Christie, P.J., Fernandez, R.C., Palmer, T., Plano, G. V, and Pugsley, A.P. (2006). Secretion by numbers: Protein traffic in prokaryotes. Mol. Microbiol. *62*, 308–319.

Fronzes, R., Christie, P.J., and Waksman, G. (2009). The structural biology of type IV secretion systems. Nat. Rev. Microbiol. *7*, 703–714.

Gérard-Vincent, M., Robert, V., Ball, G., Bleves, S., Michel, G.P.F., Lazdunski, A., and Filloux, A. (2002). Identification of XcpP domains that confer functionality and specificity to the *Pseudomonas aeruginosa* type II secretion apparatus. Mol. Microbiol. *44*, 1651–1665.

Gerlach, R.G., and Hensel, M. (2007). Protein secretion systems and adhesins: the molecular armory of Gram-negative pathogens. Int. J. Med. Microbiol. *297*, 401–415.

Hall, Z., Politis, A., and Robinson, C. V (2012). Structural modeling of heteromeric protein complexes from disassembly pathways and ion mobility-mass spectrometry. Structure *20*, 1596–1609.

Korotkov, K. V, Pardon, E., Steyaert, J., and Hol, W.G.J. (2009). Crystal structure of the N-terminal domain of the secretin GspD from ETEC determined with the assistance of a nanobody. Structure *17*, 255–265.

Korotkov, K. V, Johnson, T.L., Jobling, M.G., Pruneda, J., Pardon, E., Héroux, A., Turley, S., Steyaert, J., Holmes, R.K., Sandkvist, M., et al. (2011). Structural and functional studies on the interaction of GspC and GspD in the type II secretion system. PLoS Pathog. *7*, e1002228.

Korotkov, K. V, Sandkvist, M., and Hol, W.G.J. (2012). The type II secretion system: biogenesis, molecular architecture and mechanism. Nat. Rev. Microbiol. *10*, 336–351.

Korotkov, K. V, Delarosa, J.R., and Hol, W.G.J. (2013). A dodecameric ring-like structure of the N0 domain of the type II secretin from enterotoxigenic *Escherichia coli*. J. Struct. Biol. *183*, 354–362.

Leitner, A., Joachimiak, L. a, Bracher, A., Mönkemeyer, L., Walzthoeni, T., Chen, B., Pechmann, S., Holmes, S., Cong, Y., Ma, B., et al. (2012). The molecular architecture of the eukaryotic chaperonin TRiC/CCT. Structure *20*, 814–825.

Login, F.H., Fries, M., Wang, X., Pickersgill, R.W., and Shevchik, V.E. (2010). A 20-residue peptide of the inner membrane protein OutC mediates interaction with two distinct sites of the outer membrane secretin OutD and is essential for the functional type II secretion system in *Erwinia chrysanthemi*. Mol. Microbiol. *76*, 944–955.

McLaughlin, L.S., Haft, R.J.F., and Forest, K.T. (2012). Structural insights into the Type II secretion nanomachine. Curr. Opin. Struct. Biol. *22*, 208–216.

Van der Meeren, R., Wen, Y., Van Gelder, P., Tommassen, J., Devreese, B., and Savvides, S.N. (2013). New insights into the assembly of bacterial secretins: structural studies of the periplasmic domain of XcpQ from *Pseudomonas aeruginosa*. J. Biol. Chem. *288*, 1214–1225.

Miyata, S.T., Bachmann, V., and Pukatzki, S. (2013). Type VI secretion system regulation as a consequence of evolutionary pressure. J. Med. Microbiol. *62*, 663–676.

Reichow, S.L., Korotkov, K. V, Hol, W.G.J., and Gonen, T. (2010). Structure of the cholera toxin secretion channel in its closed state. Nat. Struct. Mol. Biol. *17*, 1226–1232.

Ruotolo, B.T., Benesch, J.L.P., Sandercock, A.M., Hyung, S.-J., and Robinson, C. V (2008). Ion mobility-mass spectrometry analysis of large protein complexes. Nat. Protoc. *3*, 1139–1152.

Sandkvist, M. (2001). Type II Secretion and Pathogenesis. Infect. Immun. *69*, 3523–3535.

Sandkvist, M., Cross, A.R., and Branch, C. (2001). Biology of type II secretion. Mol. Microbiol. *40*, 271–283.

Simeone, R., Bottai, D., and Brosch, R. (2009). ESX/type VII secretion systems and their role in host-pathogen interaction. Curr. Opin. Microbiol. *12*, 4–10.

Tseng, T.-T., Tyler, B.M., and Setubal, J.C. (2009). Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. BMC Microbiol. *9 Suppl 1*, S2.

Walzthoeni, T., Claassen, M., Leitner, A., Herzog, F., Bohn, S., Förster, F., Beck, M., and Aebersold, R. (2012). False discovery rate estimation for cross-linked peptides identified by mass spectrometry. Nat. Methods *9*, 901–903.

Wang, X., Pineau, C., Gu, S., Guschinskaya, N., Pickersgill, R.W., and Shevchik, V.E. (2012). Cysteine scanning mutagenesis and disulfide mapping analysis of arrangement of GspC and GspD protomers within the type 2 secretion system. J. Biol. Chem. *287*, 19082–19093.

Yang, B., Wu, Y.-J., Zhu, M., Fan, S.-B., Lin, J., Zhang, K., Li, S., Chi, H., Li, Y.-X., Chen, H.-F., et al. (2012). Identification of cross-linked peptides from complex samples. Nat. Methods *9*, 904–906.

Chapter 6

General Discussion:

Conclusion and Perspectives

In this PhD thesis, using an integrative structural biology approach, we have provided new insights in the molecular mechanism of the *Shewanella oneidensis* MR-1 HipA-HipB toxin antitoxin system and the *Pseudomonas aeruginosa* type 2 secretion system which are involved in bacterial phenomena like persistence and virulence. Despite these efforts, a number of questions remain to be clarified.

6.1 What needs still to be done in the study of *Shewanella oneidensis* MR-1 HipA-HipB toxin-antitoxin system molecular mechanism?

(1) The toxin HipA is a Serine/Threonine kinase and in this work, it is proven to have at least 3 different conformational states, namely apoHipA, ATP Mg^{2+} bound HipA and phosphorylated HipA (pHipA). *In vivo* verification revealed that *E. coli* HipA is only partially phosphorylated (Correia et al., 2006). We now showed that the intermediate state of ATP Mg^{2+} bound HipA could have a similar conformation as pHipA regardless of phosphorylation loop ejection. Both states were observed in the complex of HipA-HipB with their operator DNA (ATP Mg^{2+} bound HipA in *E.coli* and pHipA in *Shewanella oneidensis*). Recently, concentration threshold-based generation of variability in the toxin-antitoxin modules is reported to explain the bistability of persister formation (Cataudella et al., 2013; Feng et al., 2014; Gelens et al., 2013; Rotem et al., 2010). The identification of the effective population of HipA that is involved in the HipB binding is the key to further develop the bistability model. It is crucial to verify which conformational state of HipA is really involved in forming the heterocomplex with operator DNA binding thus further results in transcriptional suppression. This hypothesis can be further verified with experiments including: affinity measurement experiments with pHipA, ATP Mg^{2+} bound HipA, γ-phosphatase treated HipA and the HipAD306Q mutant using ITC or SPR techniques;

phosphorylation transition rate measurements from apoHipA to ATP Mg bound HipA and further phosphorylated HipA using a radioactivity assay.

(2) We have strong evidence that the EF-Tu is also not the substrate of toxin HipA in *Shewanella oneidensis* similar as in *E.coli*, where recently GltX has been proposed as the target of HipA in *E.coli* as a result of high throughput genetic screening (Germain et al., 2013; Kaspy et al., 2013). The *Shewanella oneidensis* and *E.coli* HipA share only 28% identity but act rather similar via a phosphorylation loop ejection mechanism. GltX from *Shewanella oneidensis* and *E.coli* are highly conserved with more than 90% of identity. It is essential to verify whether the GltX is also the substrate of HipA in *Shewanella oneidensis*. HipA overexpression inhibits not only macromolecular synthesis, but also DNA replication and transcription (Bokinsky et al., 2013; Korch and Hill, 2006; Korch et al., 2003). Is there any other alternative or species-specific substrate rather than GltX? This can be verified by producing the *Shewanella oneidensis* GltX and combine ITC SPR binding experiment, a chemical crosslinking mass spectrometry approach, radioactivity assay and structural biology techniques such as X-ray crystallography, SAXS.

(3) Besides SO0705-SO0706 studied in depth in this PhD project, the SO0062-SO0063 and SO3169-SO3170 are also identified as HipA-HipB toxin antitoxin modules. Hence the *Shewanella oneidensis* MR-1 is the first bacteria reported to have multiple HipA-HipB toxin antitoxin modules. It is interesting to study what are the roles of those toxin antitoxin modules, to reveal a number of questions: Are there cross reaction between the toxin and antitoxin from different modules? Is there redundancy between the three HipA-HipB toxin antitoxin systems in terms of biofilm formation and persistence?

6.2 Can the HipA-HipB TA system be an antimicrobial drug target?

Iterative cycles of antimicrobial drug discovery are urgently need due to the circumvent the almost inevitable selection for MDR and MDT bacteria that arise in clinically settings at some point after widespread (ab)use of new antibiotics (Palumbi, 2001; Walsh and Wencewicz, 2013). TA modules are an attractive potential target for antimicrobial therapy. The speculative strategies of toxin antitoxin systems involved in antimicrobial drug discovery has been reported by artificially activating the toxin antitoxin system (Shapiro, 2013; Williams and Hergenrother, 2012). However, a limitation to directly target the toxin antitoxin systems may be hindered the redundancy of the systems. Therefore, to development broad spectrum antimicrobial drugs may be better focusing to upstream and downstream paths of the TA systems regulation. For example, most of the type 2 TA systems, the antitoxins are reported to be degradated by Clp or Lon protease, hence the Clp or Lon protease could be considered as the direct drug target. Indeed, recently it was demonstrated that a combination of conventional antibiotics with a Acyldepsipeptide (ADEP) that could induce the activity of ClpP protease, kills both antibiotic sensitive and dormant persister cells (Conlon et al., 2013; Gerdes and Ingmer, 2013). Another potential TA module target can be the toxin itself. A peptide based on the CcdB toxin which targets the DNA gyrase showed also inhibition effect to the bacterial topoisomerase (Trovatti et al., 2008). Further investigation of the molecular mechanism with which HipA phosphorylates GltX could also be a source of information to perform a rational drug design approach (Lewis, 2013).

6.3 How far can we go with a mass spectrometry based integrative structural biology approach in the study of type 2 secretion system?

Mass spectrometry based integrative structural biology has developed as a powerful approach in the study of macromolecular assembly including membrane associated proteins (Marcoux and Robinson, 2013). Most recently, the development of detergent-free mass spectrometry technology was established, where intact membrane protein complexes are released in the gas phase from amphipols, bicelles and nanodiscs (Hopper et al., 2013). T2SS is a complicated multi-subunit membrane associated macromolecular assembly. At present, the structural and functional investigations of such system are not achievable through a single approach; hence integrative structural biology endorses its power to study such structures. The mass spectrometry based approach has a large advantage in terms of sensitivity and throughput and could give different level information to bridge the other technologies. The T2SS can be divided into 5 platforms, and native mass spectrometry coupled with Ion Mobility could first give an overview of the interaction partners and stoichiometry of the different platforms. The further integrative approach that we established for the inter membrane XcpP and outer membrane secretin XcpQ using cross-linking mass spectrometry could definitely be applied on the other platforms and generate distance restraints both inside each platform and later between each platform. The combination of all the information from the mass spectrometry platform and data from other techniques such as low resolution Electron Microscopy, Small angle X-ray scattering of platform complex and high resolution X-ray crystallography of individual subunit, will give a comprehensive structural and functional picture of the T2SS.

6.4 References

Bokinsky, G., Baidoo, E.E.K., Akella, S., Burd, H., Weaver, D., Alonso-Gutierrez, J., García-Martín, H., Lee, T.S., and Keasling, J.D. (2013). HipA-triggered growth arrest and β-lactam tolerance in *Escherichia coli* is mediated by RelA-dependent ppGpp synthesis. J. Bacteriol. *195*, 3173–3182.

Cataudella, I., Sneppen, K., Gerdes, K., and Mitarai, N. (2013). Conditional cooperativity of toxin - antitoxin regulation can mediate bistability between growth and dormancy. PLoS Comput. Biol. *9*, e1003174.

Conlon, B.P., Nakayasu, E.S., Fleck, L.E., LaFleur, M.D., Isabella, V.M., Coleman, K., Leonard, S.N., Smith, R.D., Adkins, J.N., and Lewis, K. (2013). Activated ClpP kills persisters and eradicates a chronic biofilm infection. Nature *503*, 365–370.

Correia, F.F., D'Onofrio, A., Rejtar, T., Li, L., Karger, B.L., Makarova, K., Koonin, E. V, and Lewis, K. (2006). Kinase activity of overexpressed HipA is required for growth arrest and multidrug tolerance in *Escherichia coli*. J. Bacteriol. *188*, 8360–8367.

Feng, J., Kessler, D.A., Ben-Jacob, E., and Levine, H. (2014). Growth feedback as a basis for persister bistability. Proc. Natl. Acad. Sci. U. S. A. *111*, 544–549.

Gelens, L., Hill, L., Vandervelde, A., Danckaert, J., and Loris, R. (2013). A general model for toxin-antitoxin module dynamics can explain persister cell formation in *E.coli*. PLoS Comput. Biol. *9*, e1003190.

Gerdes, K., and Ingmer, H. (2013). Antibiotics: Killing the survivors. Nature *503*, 347–349.

Germain, E., Castro-Roa, D., Zenkin, N., and Gerdes, K. (2013). Molecular mechanism of bacterial persistence by HipA. Mol. Cell *52*, 248–254.

Hopper, J.T.S., Yu, Y.T.-C., Li, D., Raymond, A., Bostock, M., Liko, I., Mikhailov, V., Laganowsky, A., Benesch, J.L.P., Caffrey, M., et al. (2013). Detergent-free mass spectrometry of membrane protein complexes. Nat. Methods *10*, 1206–1208.

Kaspy, I., Rotem, E., Weiss, N., Ronin, I., Balaban, N.Q., and Glaser, G. (2013). HipA-mediated antibiotic persistence via phosphorylation of the glutamyl-tRNA-synthetase. Nat. Commun. *4*, 3001.

Korch, S.B., and Hill, T.M. (2006). Ectopic overexpression of wild-type and mutant hipA genes in *Escherichia coli*: effects on macromolecular synthesis and persister formation. J. Bacteriol. *188*, 3826–3836.

Korch, S.B., Henderson, T. a., and Hill, T.M. (2003). Characterization of the hipA7 allele of *Escherichia coli* and evidence that high persistence is governed by (p)ppGpp synthesis. Mol. Microbiol. *50*, 1199–1213.

Lewis, K. (2013). Platforms for antibiotic discovery. Nat. Rev. Drug Discov. *12*, 371–387.

Marcoux, J., and Robinson, C. V (2013). Twenty years of gas phase structural biology. Structure *21*, 1541–1550.

Palumbi, S.R. (2001). Humans as the world's greatest evolutionary force. Science *293*, 1786–1790.

Rotem, E., Loinger, A., Ronin, I., Levin-Reisman, I., Gabay, C., Shoresh, N., Biham, O., and Balaban, N.Q. (2010). Regulation of phenotypic variability by a threshold-based mechanism underlies bacterial persistence. Proc. Natl. Acad. Sci. U. S. A. *107*, 12541–12546.

Shapiro, S. (2013). Speculative strategies for new antibacterials: all roads should not lead to Rome. J. Antibiot. (Tokyo). *66*, 371–386.

Trovatti, E., Cotrim, C.A., Garrido, S.S., Barros, R.S., and Marchetto, R. (2008). Peptides based on CcdB protein as novel inhibitors of bacterial topoisomerases. Bioorg. Med. Chem. Lett. *18*, 6161–6164.

Walsh, C.T., and Wencewicz, T.A. (2013). Prospects for new antibiotics: a molecule-centered perspective. J. Antibiot. (Tokyo). *67*, 7–22.

Williams, J.J., and Hergenrother, P.J. (2012). Artificial activation of toxin-antitoxin systems as an antibacterial strategy. Trends Microbiol. *20*, 291–298.

Summary

Antibiotics have provided an enormous contribution to humans' fight against infectious microorganisms resulting in a dramatic decrease in mortality due to microbial infections around the globe. However, the widespread and uncontrolled use of antibiotics, particularly as growth promoters in feedstock caused an environmental stress for bacterial evolution and Multi Drug Resistance (MDR) and Multi Drug Tolerance (MDT) are two of the major challenges for further eradication of the bacterial infections. In recent years, this antibiotics crisis has become one of the most serious public health concerns and threats by renewed awareness. Therefore, how the bacteria could survive under the situation of antimicrobial treatment and other environmental stress, and how the bacteria infect the host organism are essential to develop novel antibacterial strategies. Finding the right strategies such as discovery of new antimicrobial drugs are urgent to fight with the current crisis of antibiotic therapy.

In this PhD research, we picked up two major process concerning bacterial infections which are bacterial persistence and virulence respectively. Among others, we investigated Toxin-antitoxin systems of which the recent discovery of new families is one of the most exciting molecular microbial phenomena crucial to understand the bacterial behavior. Toxin-antitoxin systems are widespread two or three component systems encoded in the bacterial genome and highlighted to be involved in the persistence, biofilm formation, pathogenicity, programmed cell death and many other cellular processes in bacteria. Due to those characteristics, it also has potential as attractive targets in antimicrobial drug discovery. We applied multiple structural biology, biochemical and biophysical techniques to explore the molecular mechanism of the HipAB toxin-antitoxin system coded in *S.oneidensis* MR-1 genome, involved in biofilm and persistence. Multiple atomic structures including individual HipB as

well as complexes of HipA with non-hydrolysable ATP analogues, Mg^{2+} and HipA-HipB with its operator DNA were solved. We showed that phosphorylated HipA can engage in an unexpected ternary complex with HipB and double stranded operator DNA that is distinct from the prototypical counterpart complex from *E.coli*. The structure of HipB in complex with its operator DNA reveals a flexible C-terminus that is sequestering HipA in the ternary complex, indicative of its role in binding HipA to abolish its function in persistence. The structure of a HipA mutant devoid of phosphorylation capacity in complex with a non-hydrolyzable ATP analogue uncovers that HipA autophosphorylation is coupled to an unusual conformational collapse of its pLoop. However, HipA is unable to phosphorylate the translation factor Elongation factor Tu (EF-Tu), contrary to previous reports but in agreement with more recent findings. Our studies suggest that the phosphorylation state of HipA is an important factor in persistence and that the structural and mechanistic diversity of HipAB modules as regulatory factors in bacterial persistence is broader than previously thought.

Aside from the investigation of bacterial persistence, we also study the bacterial secretion systems which play a central role in modulating the interactions between the bacteria and its host organisms or extracellular environment by secreting diverse virulence factors. We focused on the opportunistic human pathogen *Pseudomonas aeruginosa* type 2 secretion system which is one of the most versatile systems used in gram negative bacteria. In *P.aeruginosa*, the T2SS machinery contains 12 protein subunits which can be distinguished into 5 subassemblies. Among those 12 protein subunits, XcpP is thought to be tightly involved in a procedure that bridges the inner membrane and the outer membrane secretin XcpQ; the assembly of XcpP and XcpQ is crucial to understand the transport machinery. We carried out multiple mass spectrometry experiments and tightly collaborated with a PhD student from Prof. Savvas Savvides' group, Ruben Van der Meeren focusing on the biophysical, biochemical and functional analysis of the assembly between XcpP and XcpQ.

We demonstrated that the *P.aeruginosa* T2SS secretin XcpQ periplasmic domain is a dimeric protein which oligomerize into a functional dodecameric assembly with C6 symmetry instead of the previous reported dodecameric assembly with C12 symmetry. We proposed the two transition states for XcpP and XcpQ periplasmic interaction which consist of the "open" and "closed" state of the Type 2 secretion system to transport the exoprotein from the periplasm to the extracellular milieu. Overall, we have successfully applied mass spectrometric based integrative structural biology to study the bacterial persistence and virulence.

The approach combining the biochemical, structural biological, biophysical techniques and especially the mass spectrometric based techniques that could bridge those techniques are powerful for the structural and functional investigation of supermolecular assembly.

Samenvatting

Antibiotica hebben zeer sterk bijdragen aan de strijd tegen infectieuze micro-organismen en de de wereldwijde daling van de mortaliteit tegen gevolge van microbiële infecties is hoofdzakelijk aan hun toepassing te danken. Het wijdverspreide en ongecontrolleerde gebruik van deze antibiotica, voornamelijk als groeipromotoren in de dierenvoeding, heeft echter een omgevingsstress teweeggebracht die een evolutieve druk als gevolgt heeft waardoor multi-drug resistentie (MDR) of tolerantie (MDT) steeds vaker de kop op steken. Dit fenomeen, samen met het uitblijven van nieuwe generatie geneesmiddelen, vormt een bedreiging voor het verder uitroeien van bacteriële infecties. Deze zgn. antibioticacrisis is een van de belangrijkste zorgen voor de volksgezondheid.

Het is daarom van groot belang om te begrijpen hoe bacteriën kunnen overleven onder antimicrobiële behandeling of andere omgevingsfactoren, alsook hoe de infectie van het gastorganisme precies plaatsgrijpt. Dit kan bijdragen tot de ontwikkeling van nieuwe strategieën voor antibacteriële behandelingen of het ontwikkelen van nieuwe geneesmiddelen zelf.

Tijdens mijn doctoraatsonderzoek hebben we ons geconcentreerd op twee belangrijke processen die betrekking hebben op dergelijke bacteriële infecties, in casu bacteriële persistentie en virulentie. Wij onderzochten ondermeer toxine-antitoxinesystemen waarvan de recente ontdekking van een paar nieuwe families tot een van de belangrijkste ontdekkingen in de moleculaire microbiologie van de laatste jaren is. Toxine-antitoxinesystemen zijn wijdverspreide twee-of driecomponentsystemen die door het bacteriële genoom gecodeerd worden. Ze zijn betrokken bij diverse fenomenen waaronder persistentie, biofilmvorming, pathogeniciteit, geprogrammeerde celdood en andere cellulaire processen. Gezien hun belang voor de levenswijze van de bacterie worden gezien als ze een

attractieve alternatieve target in antimicrobiele geneesmiddelenontwikkeling. Wij pasten een waaier van structureel biologische, biochemische en biofysische methoden toe om het moleculaire mechanisme van HipAB toxine-antitoxine systeem gecodeerd in genoom van het modelorganisme *Shewanella oneidensis* MR-1 genome te begrijpen. We losten verschillende atomaire structuren op, waaronder die van het individuele HipB, gefosforyleerd HipA alsook complexen van HipA met niet-hydrolyseerbare ATP analogen en met HipB zelf. Ook de structuur van het HipA-HipB complex met het operator DNA werd opgehelderd. We toonden aan dat gefosforyleerd HipA kan betrokken worden in een overwacht ternair complex met HipB en dubbelstrengig operator DNA dat afwijkt van het prototypische model bekomen bij de structuuropheldering van het *E.coli* HipAB. De structuur van het HipB in complex met operator DNA geeft een flexibele C-terminus weer die HipA kan vasthouden in een ternair complex. Dit bevestigt de rol van HipB dat HipA bindt om te verhinderen dat het zijn functie in persistentie kan uitvoeren. De structuur van een HipA mutant die niet in staat is om te autofosforyleren, in complex met niet-hydrolyseerbare ATP analogen geeft aan dat HipA autofosforylatie is gekoppeld aan een ongewone conformationele veraandering van haar pLoop. We toonden aan dat, in tegenstelling tot wat oorspronkelijk beschreven was voor E.coli HipA dat het S.oneidensis eiwit niet in staat is om de elongatiefactor TU (EF-TU) te fosforyleren. Onze studie suggereert verder dat de gefosforyleerde toestand van HipA belangrijk is voor persistentie en dat de structurele en mechanistische diversiteit van HipAB modules als regulatore van persistentie groter is dan oorspronkelijk gedacht.

Naast de studie van de moleculaire mechanismen van de inductie van bacteriële persistentie hebben wij ook een studie gedaan van de moleculaire componenten die betrokken zijn bij bacteriële secretie. Deze secretiesystemen spelen een central rol in het moduleren van de interacties tussen de bacterie en haar gastheerorganisme of de extracellulaire omgeving en exporteren diverse virulentiefactoren naar de omgeving. We hebben ons toegespitst op de

opportunistische humane pathogen *Pseudomonas aeruginosa* en haar type 2 secretiesysteem (T2SS) dat model staat voor een van de meest voorkomende systemen in gram negatieve bacteriën. In *P.aeruginosa* bevat de T2SS machinerie 12 eiwitsubeenheden die kunnen onderverdeeld worden in 5 subassemblages. Bij deze twaalf subeenheden is XcpP een belangrijke subeenheden omdat die betrokken is bij de overbrugging tussen de binnenste en de buitenste membraan waar het interageert met het secretine XcpQ. Een goed begrip van het complex tussen XcpP en XcpQ is cruciaal om de transportmachinerie te begrijpen. We hebben diverse massaspectrometrische experimenten uitgevoerd, en hierbij nauw samengewerkt met een andere doctoraatstudent, Ruben Van der Meeren uit de groep van Prof. Savvas Savvides. We hebben aangetoond dat het T2SS secretine XcpQ periplasmatische domein een dimeer is, die oligomerizeert tot een functionele dodecameer met C6 symmetrie in plaats van een dodecameer met C12 symmetrie zoals voorheen gerapporteerd. We stelden een model op dat die twee transititetoestanden voor de interactie van de periplasmatische domeinen van XcpP en XcpQ, die een “open” en een “gesloten” toestand van het T2SS secretiesysteem inhouden.

Globaal gezien hebben wij voornamelijk gebruik gemaakt van massaspectrometrische technieken, en tonen aan hoe deze methoden in het portfolio van biochemische, structureel biologische en biofysische methoden voor de karakterisatie van eiwitstructuren kunnen worden ingezet. Dit draagt bij tot een integratieve en multidisciplinaire aanpak van het onderzoek naar de werking van eiwitten in het algemeen.

Acknowledgements

Finally I come to the very last part of my thesis, oh yeah. For me, the PhD is not just a doctoral degree but more like a procedure of tracing the wisdom. When I looked back to the whole full year PhD, I would like to thank your all that guide the direction or gave me hands on this path.

First of all, I would like to express my gratitude to my promoter Bart who gave me this opportunity to start a new trip far away from my hometown, you are always kind and enthusiasm to help especially at the very beginning to make my never feel alone. You are always there to discuss and have given the most freedom in the research to let me try so many different things. Thank you very much again for all the guide, advice, support and revisions.

I am very grateful to Savvas who led me into the structural biology world and enlightened my PhD, Thanks a lot for so many experience, drove me to the synchrotron, talked about your university life and speech skills, sit with me in front of the computer to solve the crystal structure, encouraged me when I am frustrated…You were always trying to make things different and to explorer the student's potential.

I really appreciate and will never forget what Bart and Savvas have done for me as the mentor and friend during my PhD.

I want to thank Prof.Jozef Van Beeumen and Christine for your kind help and invitation to your home to enjoy the beautiful traveling stories, special thanks to Prof.Jozef Van Beeumen for carefully reading my thesis and gave a lot very valuable comments.

I would like to address my appreciation to Ruben; who taught me the whole procedure of structural biology, worked together in so many week days and weekends (thanks to Jozefin, Merel and Hazel for their understanding), always gave the fastest feedback concerning collaborations, so many times drove me to the synchrotron and still remembered the 180km/h speed on the highway back from Hamburg, we have the best collaboration.

I absolutely have to mention all the lovely colleague from the LProBE, thank Ester for the fundamental work of the project, thank Isabel for the great collaboration in the bachelor practicum, I still remembered the moment at the first day I came to Gent, Isabel and Ingrid were waiting in the lobby of the Gent Sint Pieters station, thank Gonzales for all the help with Synapt, thank Isaak for all the help with my PhD defense and thank Isaak, Agata, Kasia for all happy Lunch chatting, thank for all the wonderful and happy activity of "Sint Niklass", excursion and parties. Without all of your, Isabel, Isaak, Gonzales, Agata, Kasia, Laurence, Sara, Griet, Simon, Silke, Pablo, Ingrid, Andreia, Isabel, Andreia, Rita, the life wouldn't be the same for my PhD. I would like to thank the colleagues from the structural unit, Jan for the tremendous help with SAXS modeling, Yehudi, the expert engineer in the AKTA troubleshooting, Kenneth, patient and valuable explanation of structure solving, Bjorn, help with ITC and SPR and always very good criticism for the biological issues, thank Jonathan, Erwin, Ola, Sammy, Wouter, Loes for all the patient help with all the different scientific problems. Special thanks to Ruben, Yehudi, Kenneth, Bjorn, Erwin and Ola for the all the wonderful experience of synchrotron and conference trips, those are some of my happiest time in Belgium.

Of course I have to mention my family, my Grandma, my father, my mother, my sister and my brother, thank all for your support and understanding in the last 26 years. At the end, I want to express appreciation to my girlfriend Fang, you are always there with me at all the moment no matter happy, sad or frustrated, I love you and you are the most beautiful existence in my PhD life.

Thank you all those make my life colorful in the past, now and in the future.

Yurong Wen

February 2014 in Gent

Curriculum Vitae

PERSONAL DETAILS

Name	Yurong Wen
Date of Birth	27th April 1987
Gender	Male
Nationality	Chinese
Email	Yurong.Wen@Hotmail.com
	Yurong.Wen@Ugent.be
Telephone	+32487629796

EDUCTIONAL PROFILE

Mar 2010-Feb 2014	PhD candidate in Biochemistry and Biotechnology, Department of Biochemistry and Microbiology, Ghent University, Belgium PhD thesis: An Integrative Structural Biology Approach in the Study of Bacterial Persistence and Virulence
Sep 2008-Feb 2010	Master candidate in Biochemistry and Molecular Biology, Health Science Center, Xi'an Jiaotong University, China Master thesis: Screening Specific Autoimmune Antigen of Pristane Induced Arthritis Mouse Model Rat Cartilage by Proteomics Method.
Sep 2004- Jul 2008	Bachelor in Biological Science and Engineering, School of Life Science and Technology, Xi'an Jiaotong University, China

SCIENTIFIC CONTRIBUTIONS

- **Yurong Wen**, Ester Behiels, Jan Felix, Jonathan Elegheert, Bjorn Vergauwen, Bart Devreese*, Savvas Savvides*. Structural Insights into Toxin-Antitoxin System HipA-HipB in *Shewanella oneidensis* MR-1. Under submission.

- **Yurong Wen**, Bart Devreese. Characterization of *Shewanella oneidensis* MR-1 HipA-HipB Toxin-Antitoxin system Assembly using Mass Spectrometry. Under preparation.

- Ruben Van der Meeren*, **Yurong Wen***, Patrick Van Gelder, Bart Devreese, and Savvas N Savvides. Multiple binding interactions between XcpP and XcpQ from the type II secretion system of *Pseudomonas aeruginosa*. Under preparation.

- Ruben Van der Meeren, **Yurong Wen**, Patrick Van Gelder, Bart Devreese, and Savvas N Savvides. XcpP from the type II secretion system from *Pseudomonas aeruginosa* dimerizes via its transmembrane helix. Under preparation.

- Johan Desmet*, Kenneth Verstraete*, Yehudi Bloch, Eric Lorent, **Yurong Wen**, Bart Devreese, Karen Vandenbroucke, Stefan Loverix, Thore Hettmann, Sabrina Deroo, Klaartje Somers, Paula Henderikx, Ignace Lasters[#], Savvas N. Savvides[#] Structural basis of IL-23 antagonism by an Alphabody^{TM} protein scaffold. Submitted

- **Yurong Wen**, Ester Behiels, Bart Devreese. (2014). Toxin-Antitoxin systems: their role in persistence, biofilm formation and pathogenicity. Pathogens and Disease. doi: 10.1111/2049–632X.12145.

- Jan Felix, Jonathan Elegheert, Irina Gutsche, Alexander V Shkumatov, **Yurong Wen**, Nathalie Bracke, Erwin Pannecoucke, Isabel Vandenberghe, Bart Devreese, Dmitri I Svergun, Ewald Pauwels, Bjorn Vergauwen, Savvas N Savvides. (2013). Human IL-34 and CSF-1 establish structurally similar extracellular assemblies with their common hematopoietic receptor. Structure 21, 528–539.

- Ruben Van der Meeren, **Yurong Wen**, Patrick Van Gelder, Jan Tommassen, Bart Devreese, and Savvas N Savvides. (2013). New insights into the assembly of bacterial secretins: structural studies of the periplasmic domain of XcpQ from *Pseudomonas aeruginosa*. Journal of Biological Chemistry 288(2): 1214–25.

Participation in International or National Conference or Summer School

Oral Presentation and Poster Presentation

- Eurobiofilms 2013 Conference. "Structural basis of *Shewanella oneidensis* MR-1 HipAB with its operator DNA reveals new insight in persister mechanism." Gent, Belgium September 9th-12th 2013

Poster Presentation

- Instruct Biennial Structural Biology Meeting. "Characterization of Persister Factor HipAB Molecular Mechanism in *Shewanella Oneidensis* MR-1." Heidelberg, Germany May 22nd-24th 2013
- 7th Belgian Crystallography symposium. "New insights in the molecular mechanism of the persister Factor HipAB from *Shewanella oneidensis* MR-1." Brussels, Belgium 2012
- 2nd Young scientist day Belgium Biophysical Society. "Protein Interaction study of the Type II secretion system (T2SS) from *Pseudomonas aeruginosa* using Ion Mobility Mass Spectrometry." 2011

Summer School

- "Future of Biophysics" Stanford University. Erice-Sicily, Italy 9th-19th June 2013

Assistant in the Educational Event.

- Bachelor practicum course "Electrospray Ionization, Liquid Chromatography and Mass Spectrometry" 2012-2013 semester and 2013-2014 semester
- Master Thesis Andreia Henriques "Characterization Toxin-Antitoxin HipA-HipB interaction of *Shewanella oneidensis* Metal Reducing 1, by crosslinking reactions using mass spectrometry" 2013