AN INVESTIGATION OF CRP FUNCTION IN THE DEVELOPMENT AND SECONDARY METABOLISM OF *STREPTOMYCES COELICOLOR*

By

CHAN GAO, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

DOCTOR OF PHILOSOPHY (2012) McMaster University

(Biology) Hamilton, Ontario

TITLE: An investigation of Crp function in the development and secondary metabolism of *Streptomyces coelicolor*

AUTHOR: Chan Gao, B.Sc. (Nankai University, China)

SUPERVISOR: Dr. Marie A. Elliot

NUMBER OF PAGES: xvii, 214

EXPLANATORY NOTES

This thesis is composed of two parts, which detail the work I did in both Dr. Marie Elliot's lab (Part I) and Dr. Christian Baron's lab (Part II). The reason for having two different projects is that I initially started as a Master's candidate in Dr. Baron's lab, doing research on type IV secretion systems (T4SS) in Gram negative bacteria. Following Dr. Baron's departure from McMaster (after I had transferred to the Ph.D. program), I switched to Dr. Elliot's group, where I started to work with the Gram positive bacterium *Streptomyces coelicolor*. In Part I, I characterized the phenotype of a *S. coelicolor crp* deletion mutant strain in terms of morphology, life cycle development and antibiotic production and moved the investigation forward to a molecular level by uncovering Crp association sites in the chromosome of *S. coelicolor* and identifying the downstream Crp regulon. This was followed by bioassays and metabolic profiling analyses with the aim of enhancing the synthesis of known antibiotics or stimulating the production of novel secondary metabolites via overexpressing Crp in wild-type *S. coelicolor* species. In Part II, I focused on binary protein interactions, specifically, probing interaction sites using biochemical approaches and evaluating the importance of protein-protein interactions (mainly VirB4-VirB8) in type IV secretion system assembly and substrate translocation. The NTPase activity of VirB4 was also addressed.

ABSTRACT

Cyclic AMP receptor protein (Crp) is a transcription regulator controlling diverse cellular processes in many bacteria. Deletion of the *crp* gene of *Streptomyces coelicolor* significantly delays spore germination but accelerates sporulation, while overexpression of *crp* arrests morphological development at the vegetative stage. In this study, I set out to dissect Crp function during spore formation and germination by manipulating its expression at different developmental stages using a strain carrying the *crp* gene cloned downstream of a thiostrepton inducible promoter (*tipA-crp*). The impact of Crp on the synthesis of actinorhodin, undecylprodigiosin and calcium-dependent antibiotic was assessed by performing quantitative antibiotic production assays. Chromatin immunoprecipitation-microarray assays were applied to uncover ~ 400 Crp association sites distributed across the entire genome of *S. coelicolor*. Genes associated with these sites are involved in a remarkable variety of functions including metabolism, morphogenesis and transcription regulation, with secondary metabolism clearly overrepresented. This result correlates with our transcriptional profiling experiments, which revealed that the majority of the genes affected by Crp induction fall in the functional categories of secondary metabolism, morphogenesis and transcription factors, amongst which the most striking ones are those located in the antibiotic biosynthetic clusters: *act*, *red*, *cda*, and *cpk*. Finally, multiple *Streptomyces* species have been engineered to overproduce Crp and have been subjected to comparative metabolic profiling to identify the up- and down-regulation of secondary metabolite production elicited by Crp overexpression.

ACKNOWLEDGEMENTS

Looking back on the first thirty years of my life, completing my PhD degree is definitely the most challenging experience I have ever had. This long yet fulfilling journey has been filled with the most joyous and the most difficult moments that were shared with many people. First and foremost, I want to thank Dr. Marie Elliot for accepting me into her lab and for guiding me out of the 'darkness' when I was lost. She has always been tremendously supportive, upstanding, inspirational, and caring. I could not have asked for a better mentor and role model.

I would also like to express my heartfelt gratitude to my supervisory committee members, Dr. Turlough Finan and Dr. André Bédard, for providing encouraging and constructive feedback regarding my work, and for always keeping me on the right track.

Throughout the past several years, in addition to my scientific accomplishments, I have also made some awesome friends. It has been a great pleasure for me to work with the Elliot group. I want to thank all of my lab members, in particular: Andrew Duong, for being such a great 'baby' brother and for always being so considerate; Dr. Dave Capstick, for being my BFF (best friend forever) and for his professional editing services and emotional support; Hindra, for showing me different techniques and sharing his scientific expertise; Julia Swiercz, for cheering me up and exchanging girl stories; Matt Moody, for tolerating my jokes and editing my resume; Renée St-Onge, for her warm hugs in the lab and also right before my defense; Dr. Mary Yousef, for always being so caring and for her invaluable advice on both life and graduate studies; and also Dr. Henry Haiser for his assistance with experiments.

Many thanks go to our collaborators: David Mulder from Dr. Brian Coombes's lab, for his help with data processing, and Charles Yin, for his contribution on figure creation.

In the end, I would like to thank my parents for their unconditional love and devotion in the past thirty years, especially the sacrifice they made by allowing their only child to travel to the other side of the world for her dream. I also want to thank my husband, Dr. Kai Wang, for always being extremely understanding, patient, and loving, and for our interesting discussion about DNA electrophoresis on our first date. My doctoral journey would not have been so enjoyable without him. The last but not the least, I also want to thank my kitty, A Gui, who has always been such a sweet baby boy and whom I can always find comfort from. I truly hope that this thesis will become the best gift I have ever presented to all of them.

最后，我要特别感谢我的父母（王可文女士和高文信先生）对我无条件的付出和爱，特别是他们为了成全独生女儿到地球的另一端追寻自己的梦想而所做出的牺牲。我还要感谢我的丈夫，王凯博士。谢谢他一直以来极大的耐心，理解与关爱。也要感谢他在我们第一次约会时与我讨论DNA电泳。没有他，我的博士之路不会

如此令人愉快。最后，同样重要的是，我要感谢我的小猫，阿贵。他那么地贴心那么地乖，让我总能从他身上得到安慰。我衷心地希望这部博士论文能够成为我呈送给他们所有人的最好礼物。

TABLE OF CONTENTS

LIST OF FIGURES

Part I

Part II

LIST OF TABLES

Part I

Part II

LIST OF ABBREVIATIONS AND SYMBOLS

A	Adenine
Å	Angstrom
aa	Amino acid
ABC	ATP-binding cassette
ABMM	AB minimal medium
Ala or A	Alanine
Arg or R	Arginine
AS	Acetosyringone
Asp or D	Aspartic acid
Asn or N	Asparagine
ATP	Adenosine triphosphate
ATPase	Adenosine triphosphatase
bp	Base pair
C	Cytosine
cAMP	3'-5'-cyclic adenosine monophosphate
c-di-GMP	Cyclic diguanylate
$CaCl_2$	Calcium chloride
Cys	Cysteine
cDNA	Complementary DNA
cm	Centimetre
C-terminus	Carboxyl terminus
2-D	Two dimensional
dH_2O	Distilled water
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
DPA	Dipicolinic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
FAD	Flavin adenine dinucleotide
FMN	Flavin mononucleotide
g	Gram or
g	Gravity
G	Guanine
GFP	Green fluorescent protein
Gln or Q	Glutamine
Glu or E	Glutamic acid
Gly or G	Glycine
GTP	Guanosine triphosphate
h	Hour
HCl	Hydrochloric acid

HF buffer	High fidelity PCR reaction buffer
His or H	Histidine
IgA	Immunoglobulin A
IgG	Immunoglobulin G
Ile	Isoleucine
Inc	Incompatibility group
IPTG	Isopropyl-β-d- thiogalactopyranoside
kb	Kilo bp
KCl_2	Potassium chloride
kDa	Kilo Daltons
KOH	Potassium hydroxide
KH_2PO_4	Monopotassium phosphate
kV	Kilovolts
L-Val	L-Valine
Leu or L	Leucine
Lys or K	Lysine
M	Molar
Mb	Mega bp
mg	Milligram
$MgCl_2$	Magnesium chloride
$MgSO_4$	Magnesium sulphate
min	Minute
ml	Millilitre
mm	Millimetre
mM	Millimolar
mRNA	Messenger RNA
m/z	mass/charge
NADPH	Nicotinamide adenine dinucleotide phosphate
NaH_2PO_4	Monosodium phosphate
NaCl	Sodium chloride
NaOAc	Sodium acetate
ng	Nanogram
nm	Nanometre
nt	Nucleotide
NTP	Nucleoside triphosphate
N-terminus	Amino terminus
ONPG	Ortho-nitrophenyl-β-galactoside
ORF	Open reading frame
oriC	Origin of replication
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
pH	Potential of hydrogen
pI	Isoelectirc point
pmole	Picomole

PMSF	Phenylmethanesulfonyl fluoride
ppGpp	Guanosine pentaphosphate
PVDF	Polyvinylidene fluoride
psi	Pounds per square inch
RNA	Ribonucleic acid
RNase	Ribonuclease
rRNA	Ribosomal RNA
S	Svedberg
SALP	SsgA-like proteins
SARP	*Streptomyces* antibiotic regulatory proteins
SDS	Sodium dodecyl sulfate
s	Second
Ser or S	Serine
T	Thymine
Taq	DNA polymerase from *Thermus aquaticus*
TES buffer	N-Tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid
Thr or T	Threonine
TM	Transmembrane domain
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer RNA
Trp or W	Tryptophan
Tyr or Y	Tyrosine
U	Uridine
U	Enzyme unit
UV	Ultra-violet
V	Volts
V or Val	Valine
v/v	Volume/Volume
w/v	Weight/Volume
Wbl	WhiB-like
X-Gal	5-bromo-4-chloro-indolyl-β-D-galactopyranoside
µg	Microgram
µl	Microlitre
µm	Micrometre
µM	Micromolar

PART I

CHAPTER 1:
GENERAL INTRODUCTION

1.1 THE ACTINOMYCETES

The actinomycetes are a group of Gram positive bacteria that share a high genomic GC (guanine and cytosine) content. They constitute the order *Actinomycetales* and are found in virtually all terrestrial and aquatic environments (Waksman, 1967). These organisms were originally believed to be fungi due to their filamentous growth and sporogenic lifestyle, and were consequently given the name "actino" based on the Greek words "*akti*s" (ray) and "*mykes*" (fungus) (Hopwood, 1999). For more than six decades thereafter, these organisms were frequently regarded as intermediates between fungi and bacteria, but were finally recognized as true bacteria in the 1950's based on their size, chemical composition and biochemical activities (Waksman, 1967).

Most actinobacteria are heterophytes that feed on dead and decaying materials in the environment (Chater *et al*, 2010). Some are also pathogenic, like *Mycobacterium tuberculosis* and *M. leprae*, which cause tuberculosis and leprosy, respectively, and *Corynebacterium diphtheriae*, which causes diphtheria (Stackebrandt & Woese, 1981). Many others, in particular those belonging to the genus *Streptomyces*, have also attracted enormous attention for their prolific production of small molecule natural products with application in agriculture and the pharmaceutical industry (Berdy, 2005).

1.1.1 The Genus *Streptomyces*

Streptomyces is the largest genus of all actinomycetes with over 500 known species (Kampfer, 2006). As the most ubiquitous and abundant soil-dwelling bacteria, they play a central role in carbon recycling because of their ability to break down insoluble organic materials such as lignocellulose and chitin (Hodgson, 2000). *Streptomyces* are also characterized by a complex secondary metabolism, giving rise to a wide range of bioactive compounds including anti-fungal agents, immunosuppressants and antibiotics. The first antibiotic 'actinomycin' was isolated from *Actinomyces antibioticus* by Selman Waksman in 1940, and ushered in the 'golden era' of antibiotic discovery. However, after reaching a peak during the 1960's, new antibiotic development has dramatically declined since that time (Waksman & Woodruff, 1940; Wright, 2007).

To date, *S. coelicolor* A3(2) is the most thoroughly genetically studied member of the *Streptomyces* (Hopwood, 1999). The genome of *S. coelicolor* comprises an 8.7 Mb linear chromosome with a centrally positioned *oriC*, with some strains also containing a 350 kb linear plasmid SCP1 and a circular plasmid SCP2 (Bibb *et al*, 1977). Complete genome sequencing of this species was completed in 2002, revealing 7,825 predicted genes with an overal GC content of ~72% (Bentley *et al*, 2002). Based on the types of the genes, the chromosome is divided into three segments: a central core region extending from 1.5 Mb

to 6.4 Mb, and two arms of uneven length (1.5 Mb and 2.3 Mb) (Bentley *et al*, 2002). The core shows significant synteny to other actinomycetes genomes, whereas the arms contain genes that are more unique to individual organisms (Bentley *et al*, 2002). The majority of the genes that are essential for survival in most bacteria (*e.g.* DNA replication, transcription, and cell division) lie in the central core, while the ones involved in non-essential functions such as secondary metabolism, hydrolytic exoenzymes and 'gas vesicle' proteins are more commonly found in the arms (Bentley *et al*, 2002). Another prominent feature that emerged from the genome sequence was that 12.3% of the *S. coelicolor* genome is dedicated to transcription regulation; this is much higher than the 5.8% observed for *Escherichia coli*, 5.3% for *Bacillus subtilis* (another soil-dwelling, Gram positive spore forming bacterium) and 3.0% for *M. tuberculosis* (Stover *et al*, 2000). This extraordinarily larger number of regulatory genes suggests that a sophisticated regulatory system must be required to control the complex life cycle and secondary metabolism of *Streptomyces*.

1.2 LIFE CYCLE OF *STREPTOMYCES COELICOLOR*

As a multicellular bacterium, *S. coelicolor* undergoes a complex life cycle that starts with the germination of a single spore when it encounters favourable conditions (Elliot *et al*, 2007; Flardh & Buttner, 2009) (Fig. 1.1). Undefined environmental signals trigger the influx of water into spores, resulting in spore swelling and the emergence of germ tubes, which develop into filamentous substrate hyphae. These vegetative cells are multigenomic, and only occasionally are cross-walls laid down; however, these cross-walls do not support cytokinesis. The substrate hyphae grow through tip extension and branching to form a tangled network referred to as the substrate mycelium. This vegetative growth continues until some sort of adverse condition is encountered (frequently presumed to be nutrient depletion), at which point a second type of filamentous cell termed the "aerial hypha" appears. These aerial hyphae escape the aqueous environment of the substrate mycelium and grow up into the air. In contrast to the substrate hyphae, aerial hyphae do not branch but are synchronously subdivided into long chains of up to 50-100 prespore compartments, each containing a single chromosome. These 'prespores' subsequently round up and metamorphose into mature spores, and deposit a grey polyketide pigment on the spore surface that turns the colony from white to grey (Elliot *et al*, 2007; Flardh & Buttner, 2009). The different developmental stages will be discussed in the following sections and the genes regulating morphogenesis are summarized in Table 1.1.

1.2.1 Germination and Vegetative Growth

Preceding germ tube emergence, *Streptomyces* spores swell, resulting in a 23% increase in surface area. This is accompanied by the differentiation of what appears to be a homogeneous spore wall into double layers (Sharples & Williams, 1976). The outer layer gets ruptured at the point of outgrowth while the inner layer develops into the future germ tube cell wall. At this stage, chromosomes and ribosomal particles are clearly

visible under an electron microscope (Jyothikumar *et al*, 2008; Sharples & Williams, 1976) (Fig. 1.2). Currently, very little is known about the triggers of germination and what genes are involved in its regulation. In *B. subtilis*, spore germination commences when nutrients such as single amino acids, sugars, purine nucleotides, or combinations of these, are sensed (Setlow, 2003). These nutrients are termed 'germinants' and they bind germinant receptors embedded in the inner membrane of the spore, eliciting the release of monovalent cations and dipicolinic acid (DPA), the influx of water, and eventually the hydrolysis of the spore cortex (Setlow, 2003). This germination mechanism is not likely to be shared by *S. coelicolor*, as *Streptomyces* spores do not contain DPA, and although germination can be stimulated by amino acids like L-Val, no germinant receptors have been identified (Ensign, 1978).

After germination has initiated and germ tubes have emerged, vegetative growth then proceeds at the hyphal tips. This can be visualized by staining the mycelia with fluorescent vancomycin, where vancomycin specifically interacts with the un-crosslinked peptide cross-bridges found in nascent peptidoglycan (Daniel & Errington, 2003). Tip-associated growth of streptomycetes and other actinomycetes like *Mycobacterium* and *Corynebacterium* distinguishes them from most other bacteria, particularly rod-shaped bacteria like *E. coli* and *B. subtilis* (Carballido-Lopez, 2006). In rod-shaped microorganisms, newly synthesized peptidoglycan is evenly inserted into the lateral wall, resulting in cells that grow length-wise, thereby acquiring a rod shape. This process is directed by the bacterial actin homolog MreB, which assembles to form a helical structure along the cytoplasmic face of the cell membrane (Carballido-Lopez, 2006). In *S. coelicolor*, however, tip-associated cell wall biosynthesis does not require MreB, which is important during sporulation rather than vegetative growth (Mazza *et al*, 2006); instead, apical growth depends on DivIVA, a protein that in *B. subtilis* is a key regulator of accurate cell division at the midpoint of the cell (Edwards & Errington, 1997). In *S. coelicolor*, DivIVA assembles into discrete foci at the hyphal tip, and along the lateral wall of growing hyphae, marking the sites of new branch emergence and leaving each growing tip associated with DivIVA (Flardh, 2003). Overproduction of DivIVA in germinating spores results in swollen and rounded spores that are incapable of outgrowth, while overexpression of this protein in an established substrate mycelium leads to frequent, aberrant branching (Hempel *et al*, 2008). These observations suggest a role for *S. coelicolor* DivIVA as a landmark protein for defining new tips during germ tube emergence and vegetative growth.

1.2.2 Aerial Hyphae Differentiation

So far, two classes of genes have been found essential for the morphogenesis of *S. coelicolor*, the *bld* genes and the *whi* genes. The *whi* genes are primarily involved in cell division and sporulation (as discussed in Section 1.2.3) while the *bld* genes are important for aerial development. The '*bld*' (for 'bald') designation refers to the lack of a fuzzy colony appearance in strains containing mutated *bld* genes, where fuzziness is usually attributed to the presence of aerial hyphae (Elliot *et al*, 2007). An extracellular

complementation study between pairs of *bld* mutants demonstrated a hierarchical cascade: *bldJ* < *bldK*/*bldL* < *bldA*/*H* < *bldG* < *bldC* < *bldD*/*bldM* (Nodwell *et al*, 1999; Willey *et al*, 1993). The mutants following this scheme can extracellularly complement mutants lower in the hierarchy, but not higher. Although it remains to be fully understood how these genes orchestrate morphological development, it was proposed that an intercellular signaling cascade exists and each *bld* gene regulates a specific step of the pathway.

A candidate extracellular signal receiver is the ATP-binding cassette transporter encoded by the *bldK* locus (Nodwell *et al*, 1996; Nodwell *et al*, 1999). This protein complex putatively imports an oligopeptide signaling molecule capable of initiating aerial growth. A gene higher in the hierarchy, *bldA*, encodes the only tRNA able to efficiently translate the rare Leu codon UUA (Leskiw *et al*, 1991). The lack of aerial growth in *bldA* mutants is primarily attributed to the UUA codon in the transcript of another developmental regulatory gene *bldH* (Nguyen *et al*, 2003; Takano *et al*, 2003). Apart from *bldA* and *bldK*, most of the other *bld* genes encode regulatory proteins. BldG and its downstream coexpressed protein SCO3548 are homologous to anti-anti-sigma factor and anti-sigma factor proteins respectively, and act antagonistically as a switching pair to control morphogenesis (Parashar *et al*, 2009). *bldN* encodes a sigma factor required for development but is dispensable for antibiotic production (Bibb *et al*, 2000). It also directly activates the transcription of *bldM*, which encodes an "orphaned" response regulator (Molle & Buttner, 2000). While occupying the highest position in the extracellular complementation hierarchy, BldD encodes a small DNA-binding protein that directly regulates the transcription of many key developmental genes, including *bldA*, *bldM*, *bldN*, *whiB*, *whiG*, and *ftsZ* (the last three are implicated in sporulation) (den Hengst *et al*, 2010; Elliot *et al*, 2001). The main function of BldD appears to be to repress the expression of these genes during vegetative growth (den Hengst *et al*, 2010). In addition to morphological differentiation, a number of *bld* mutants are also impaired in antibiotic production, which usually coincides with the onset of aerial hyphae erection, supporting a regulatory connection between morphogenesis and secondary metabolism.

All of the *bld* genes are required for the production of two substances promoting aerial hyphae erection, SapB and the chaplins (Flardh & Buttner, 2009). High-resolution scanning electron microscopy (SEM) revealed a basketwork-like ultrastructure of paired rodlets on the surface of aerial hyphae and spores (Claessen *et al*, 2002). The major constituents of these rodlets are the chaplin proteins, a family of highly conserved proteins sharing a characteristic hydrophobic chaplin domain (Claessen *et al*, 2003; Elliot *et al*, 2003). *S. coelicolor* encodes eight chaplins (Chp) including ChpA-C, which contain two chaplin domains, and ChpD-H, which have a single chaplin domain. These proteins are secreted to the aerial surfaces and self-assemble into amyloid filaments (Capstick *et al*, 2011), where they are then proposed to be organized into paired rodlets by two rodlin (Rdl) proteins, RdlA and RdlB. This leads to the formation of a hydrophobic layer coating the aerial hyphae surface, and greatly facilitates the hyphal escape from the aqueous vegetative mycelium (Claessen *et al*, 2004). While the chaplins are sufficient for promoting aerial growth on poor carbon sources, on rich medium, production of a

surfactant peptide SapB is also required to efficiently break surface tension (Willey *et al*, 1991). SapB synthesis is mediated by the *ramCSAB* operon, where the *ramS* gene product goes through extensive posttranslational modification by RamC before being cleaved to yield the mature SapB peptide (Kodani *et al*, 2004). The convergently transcribed *ramR* gene encodes a response regulator that directly activates the *ramCSAB* operon (O'Connor *et al*, 2002; Nguyen *et al*, 2002); a direct regulator of the chaplin genes has yet to be identified although they do require the *bld* genes for expression.

1.2.3 Cell Division and Chromosome Segregation

Streptomycetes are the only microorganisms where cell division is not required for bacterial viability and growth, however, it is mandatory for completing sporulation (McCormick & Losick, 1996). As mentioned above, the *whi* genes play important regulatory roles in various aspects of sporulation. The '*whi*' (for 'white') name arose because mutations in these genes block spore formation, and therefore the mutant colonies appear white rather than grey like the wild-type colonies (Flardh & Buttner, 2009). Like the *bld* genes, most *whi* genes encode regulatory proteins. These can be divided into two groups since their mutations block differentiation at different stages. *whiA*, *whiB*, *whiG*, *whiH*, *whiI*, and *whiJ* are recognized as early *whi* genes with more significant roles during the initial stages of sporulation, while the late *whi* genes such as *whiD*, *whiE*, *whiF*, *whiL*, *whiM*, and *whiO* are more important later during spore maturation (Chater, 2001; Elliot *et al*, 2007). WhiG encodes a sigma factor that acts at the earliest stage to commit aerial hyphae to sporulation (Chater *et al*, 1989). Defects in *whiG* result in long aerial hyphae incapable of coiling or progressing into the septation stage (Chater, 1975). As a repressed target of BldD, σ^{WhiG} controls the transcription of *whiH* and *whiI*, which encode DNA-binding transcription factors whose direct targets have not been identified (Ainsa *et al*, 1999; den Hengst *et al*, 2010; Ryding *et al*, 1998). Small proteins WhiB and WhiD both belong to the actinomycete-specific 'Wbl' (for 'WhiB-like') protein family (Molle *et al*, 2000; Soliveri *et al*, 2000). They both contain four invariant Cys residues for coordinating an oxygen-sensitive [4Fe-4S] cluster, implying that they may contribute to development by sensing redox changes (Jakimowicz *et al*, 2005). *whiJ* belongs to a gene cluster paralogous to the *abaA* locus, which affects antibiotic production. WhiJ represses aerial growth and sporulation, and the repression can be relieved by interacting with the product of the gene directly upstream of *whiJ* (Ainsa *et al*, 2010). The regulatory protein encoded by *whiA* affects the transcription of itself and the *parAB* genes, which are required for proper chromosome segregation later during spourlation (Kaiser & Stoddard, 2011). There is also data suggesting that it interacts with WhiG, leading to inhibition of WhiG-dependent transcription (Kaiser & Stoddard, 2011). In contrast to the early *whi* genes, the late *whi* genes are not as well characterized. The *whiE* locus contains eight genes and specifies the synthesis of the grey type II polyketide spore pigment (Flardh & Buttner, 2009; Kelemen *et al*, 1998). Transcription of *whiE* depends on a sigma factor encoded by *sigF* (Kelemen *et al*, 1998; Potuckova *et al*, 1995). SigF is one of nine stress response sigma factors encoded by *S. coelicolor* (Potuckova *et al*, 1995). The *sigF* mutants exhibit deformed thin-walled spores

with uncondensed chromosomes (Potuckova *et al*, 1995). *whiL*, *M* and *O* have not been fully mapped and thus have not been characterized.

As aerial hyphae grow through tip extension, they lay down occasional vegetative cross walls in a way similar to substrate hyphae (Elliot *et al*, 2007). At a later stage, the apical compartment at the tip of each aerial hypha differentiates into a large sporogenic cell encompassing 50 or more copies of the chromosome (Ruban-Osmialowska *et al*, 2006). The sporogenic cell subsequently ceases extension and undergoes large scale, synchronous cell division (Schwedock *et al*, 1997). Upon strong upregulation of the bacterial tubulin homolog FtsZ in the aerial hyphae, FtsZ defines the precise sites of division in a *whi* gene-dependent manner and recruits other cell-division proteins to the same sites (Flardh *et al*, 2000; Grantcharova *et al*, 2005). Initially, FtsZ polymerizes into helical filaments along the length of the sporogenic cell, and these filaments are ultimately remodeled into regularly spaced Z rings that direct septation, before MreB drives the construction of a thick, lysozyme-resistant cell wall (Mazza *et al*, 2006). A recent study revealed that FtsZ itself is recruited to the future division sites by two SsgA-like proteins (SALPs), SsgA and SsgB (Willemse *et al*, 2011). SsgA appears to be positioned first, followed by SsgB, which then recruits FtsZ to the same sites and promotes the ensuing Z-ring formation. SALPs are small proteins found exclusively in the streptomycetes (Noens *et al*, 2005). *S. coelicolor* encodes seven SALPs, all of which affect development. Although SsgA and SsgB act early on during sporulation, the other SALPs appear to act at later sporulation stages, affecting processes such as spore wall thickening and spore dispersal (Noens *et al*, 2005).

As septation initiates, partitioning of the previously uncondensed nuclear material into individual nucleoids occurs. Correct positioning and segregation of chromosomes depends on ParAB and FtsK (Flardh & Buttner, 2009). ParB binds to specific sites around the chromosomal *oriC* and assembles into nucleoprotein complexes to ensure the placement of a single chromosome in each future prespore compartment. This process is putatively driven by the ATPase activity of ParA (Kim *et al*, 2000). The segregation process is completed when the DNA translocase FtsK functions to clear unsegregated nucleoids from closing sporulation septa so as to prevent the guillotining of chromosomal termini in newly formed prespores (Wang *et al*, 2007). Chromosome partition and segregation are also governed by the *whi* gene products (Flardh & Buttner, 2009).

1.3 SECONDARY METABOLISM OF *S. COELICOLOR*

Approximately two-thirds of known antimicrobials with application in both medicine and agriculture are made by streptomycetes. In their natural habitats, these antibiotics are thought to help defend streptomycete colonies from other competitors. Chromosomal sequencing of *S. coelicolor* A3(2) uncovered 22 gene clusters devoted to secondary metabolism (Bentley *et al*, 2002), four of which have been very well characterized. The products of these clusters are four chemically distinct antibiotics: actinorhodin (Act) (Wright & Hopwood, 1976a), undecylprodigiosin (Red) (Malpartida *et al*, 1990; Rudd &

Hopwood, 1980), calcium-dependent antibiotic (CDA) (Chong *et al*, 1998), and yellow cryptic polyketide (yCPK) (Gottelt *et al*, 2010; Pawlik *et al*, 2007). The plasmid SCP1 also harbours a cluster that directs the synthesis of methylenomycin (Mmy) (Chater & Bruton, 1985; Wright & Hopwood, 1976b). SCP1 is, however, absent in the wild-type strain M145 used for this study (Kieser, 2000) and thus will not be discussed further. Act is the best studied polyketide synthesized by *S. coelicolor*. It is initially produced intracellularly, but is then secreted to the outside of the cells where the majority of Act is converted to blue-pigmented γ-Act (lactone form) (Bystrykh *et al*, 1996). This secretion and lactone conversion occurs only when the pH is above 7.0; otherwise, Act remains inside the cells in its red-pigmented form (Bystrykh *et al*, 1996). By contrast, the intracellular undecylprodigiosin looks red at acidic and neutral pH (hence its 'Red' designation), but turns yellow at alkaline pH (Anderson *et al*, 1999). Please refer to Table 1.2 for a summary of the genes involved in secondary metabolism.

1.3.1 Antibiotic Biosynthesis Clusters and Regulatory Genes

Genes mediating antibiotic synthesis are usually arranged in contiguous clusters that vary in size from a few to over 100 kb (Bibb, 2005). Located within these clusters are open reading frames (ORFs) encoding biosynthetic enzymes, resistance determinants, and pathway-specific regulators (Bibb, 2005) (Fig. 1.3). ActII-ORF4 (encoded by *SCO5085*), the pathway-specific regulator of the *act* cluster, binds to specific sequences upstream of at least two biosynthetic operons within the *act* cluster and positively regulates Act production (Arias *et al*, 1999). RedD (encoded by *SCO5877*) is homologous to ActII-ORF4 and activates the expression of the *red* cluster. Transcription of *redD* itself depends on the activity of a response regulator RedZ (encoded by *SCO5881*), which is required for the activation of RedD-independent *red* cluster genes such as *redW*, *redX*, and *redY* (Guthrie *et al*, 1998; Huang *et al*, 2001; White & Bibb, 1997). CdaR (encoded by *SCO3217*) is a specific regulator of the *cda* cluster and is presumed to promote the transcription of the *cda* genes although there is no direct evidence showing this at present (Ryding *et al*, 2002). ActII-ORF4, RedD, and CdaR all belong to a novel family of *Streptomyces* antibiotic regulatory proteins (SARPs), and share an OmpR-like DNA binding fold in their N-termini. OmpR is the response regulator of the EnvZ/OmpR two-component system controlling osmoregulation in *E. coli* (Wietzorrek & Bibb, 1997). Another SARP family pathway-specific regulator identified in *S. coelicolor* is encoded by the *cpkO* gene of the *cpk* cluster. This cluster was discovered in recent years to direct the synthesis of a yellow pigmented cryptic type I polyketide (Pawlik *et al*, 2007). Expression of *cpkO* is negatively regulated by a γ-butyrolactone receptor protein encoded by *scbR (SCO6265)* (Gottelt *et al*, 2010), and this repression can be relieved when ScbR binds the γ-butyrolactone SCB1 synthesized by ScbA (encoded by *SCO6266*) (Gottelt *et al*, 2010). The genes encoding ScbA and ScbR both lie within the *cpk* cluster.

Apart from the pathway-specific regulators, secondary metabolic gene clusters are additionally controlled by pleiotropic regulators that affect the expression of multiple gene clusters (Table 1.1). These pleiotropic regulators typically exert their regulatory

effects by modulating the transcription of pathway-specific regulatory genes, the products of which in turn stimulate the transcription of the structural genes in their cognate clusters (Fig. 1.4). The AbsA1/A2 two-component system encoded by the *absA* locus (*SCO3225/3226*) was one of the first pleiotropic regulators identified in *S. coelicolor* (Brian *et al*, 1996) (Fig 1.5). The *absA* locus is embedded within the *cda* cluster. Interestingly, single site mutations in this locus almost completely abolished the production of Act, Red, and CDA, while the deletion of one or both genes conferred antibiotic hyperproduction (Anderson *et al*, 1999). Investigation showed that the response regulator AbsA2, when phosphorylated by the sensor kinase AbsA1, directly interacts with the promoter regions of *actII-ORF4*, *redZ*, and *cdaR* to reduce their transcription. AbsA1 has dual kinase-phosphatase activity and can also eliminate this repression by dephosphorylating AbsA2-P (McKenzie & Nodwell, 2007). Therefore, the AbsA1/A2 system affects antibiotic production by altering the phosphorylation state of the repressor AbsA2. Although the majority of SARPs appear to function as pathway-specific regulators, *S. coelicolor* produces at least one pleiotropic SARP, AfsR (encoded by *SCO4426*) (Floriano & Bibb, 1996). This protein contains a helix-turn-helix domain typical of SARP members in its N-terminus, two ATP-binding sites in the center, and requires phosphorylation by the membrane-associated kinase AfsK (encoded by *SCO4423*). AfsK autophosphorylates and subsequently transfers its phosphates to the Thr and Ser residues of AfsR, which in turn activates the transcription of the *afsS* gene (*SCO4425*) (Floriano & Bibb, 1996; Petrickova & Petricek, 2003). *afsS* encodes a small sigma factor-like protein that acts as a positive regulator of the *act*, *red*, and *cda* pathways (Vogtli *et al*, 1994). This signaling pathway can be blocked by KbpA (encoded by *SCO4422*), an AfsK kinase inhibitor (Umeyama & Horinouchi, 2001) (Fig.1.6).

Most secondary metabolic gene clusters (and their associated pathway specific regulatory genes) are situated in the arms of the chromosome, whereas many pleiotropic antibiotic regulators are encoded in the central core (Bentley *et al*, 2002; McKenzie & Nodwell, 2007). Included amongst the core-encoded regulators are: AtrA (encoded by *SCO4118*), which activates Act production through direct interaction with the *actII-ORF4* promoter sequence (Uguru *et al*, 2005); CprA (encoded by *SCO6312*) and CprB (encoded by *SCO6071*), which are homologs of the A-factor receptor protein (ArpA) in *S. griseus* and whose null mutants display differential effects on antibiotic production (CprA is a positive regulator of Act and Red, while CprB is a negative regulator of Act) (Onaka *et al*, 1998). The *S. coelicolor* genome encodes 67 two-component systems, each comprising a histidine sensor kinase and a response regulator (Yepes *et al*, 2011). In addition to the canonical AbsA1/A2, other two-component systems involved in secondary metabolism include CutR/S (encoded by *SCO5862/*3), which selectively represses Act production (Chang *et al*, 1996), EcrA1/A2 (encoded by *SCO2517/2518*), which positively regulates the *red* cluster (Li *et al*, 2004), AfsQ1/Q2 (encoded by *SCO4906/4907*), whose deletion significantly decreases Act, Red, and CDA under certain growth conditions (Ishizuka *et al*, 1992), and the most recently identified AbrA1/A2 (encoded by *SCO1744/1745*), and AbrC1/C2/C3 (encoded by *SCO4596/4597/4598*), which act as negative and positive regulators of Act, Red and CDA production, respectively (Yepes *et*

al, 2011). Although the *absB* locus was defined along with *absA*, it does not encode a two-component system; instead, its gene product is a homolog of *E. coli* RNase III, a double strand-specific RNase (Price *et al*, 1999). Disruption of this gene resulted in abolishment of Act, Red, and CDA production and accumulation of 30S rRNA precursors, suggesting that AbsB may exert its effect through the posttranscriptional regulation of genes involved in secondary metabolism (Price *et al*, 1999).

As a consequence of the advancement in genome mining, the number of secondary metabolism regulatory genes has been expanding. In theory, pathway-specific regulators exert their primary effects on a single product while pleiotropic regulators have more global roles in controlling the synthesis of multiple products. However, the distinction between the two is becoming blurred, as some classical pathway-specific regulators (*e.g.* RedZ) have been shown to affect the production of other antibiotics (Huang *et al*, 2005). Likewise, some global regulatory genes such as the *absA* locus lie within a specific biosynthetic cluster, and yet seem to have a strong regulatory impact on multiple clusters. Taken together, these observations suggest that there is likely significant cross-talk between different pathways (Huang *et al*, 2005).

1.3.2 Impact of Bacterial Growth and Nutrients on Antibiotic Production

In addition to the genetic regulation of secondary metabolism in *S. coelicolor*, it is also coordinated with bacterial growth. In liquid culture, secondary metabolism initiates during stationary phase, whereas in surface-grown cultures, it coincides with the onset of morphological differentiation (Bibb, 2005). Although not directly proven, growth rate slow-down or growth arrest caused by nutrient deprivation is deemed an important trigger for morphological development and secondary metabolism. As a crucial nutrient, carbon source can significantly affect antibiotic production. Glucose is preferentially utilized by most bacteria as the major carbon source and inhibits the usage of other sugars. In *S. coelicolor*, growth on glucose has adverse effects on Act production due to the repression of *afsS* transcription (Kim *et al*, 2001; Lee *et al*, 2009). By contrast, phosphorylated N-acetylglucosamine (GlcNAc), an alternative carbon and nitrogen source, can stimulate antibiotic production by inhibiting the DNA binding activity of the global regulator DasR, where DasR otherwise serves to repress the expression of *actII-ORF4* and *redD* (Rigali *et al*, 2008). GlcNAc is most likely obtained through degradation of chitin polymers from the environment or breakdown of the substrate hyphal cell wall before the onset of aerial growth (Rigali *et al*, 2008).

Besides carbon and nitrogen, phosphate is also an important player in antibiotic production. In *E. coli*, the highly phosphorylated guanine nucleotide ppGpp participates in growth rate control and stationary phase gene regulation (Gentry *et al*, 1993). It is synthesized by the ribosome-associated synthetase RelA. In *S. coelicolor*, ppGpp accumulates transiently in response to nitrogen starvation and inhibits rRNA synthesis (Chakraburtty & Bibb, 1997). A *relA* deletion mutant fails to produce Red and Act under conditions of nitrogen limitation, suggesting either that ppGpp functions as a signaling

molecule to trigger antibiotic production, or that the lack of antibiotic production is an indirect consequence of reduced growth rate resulting from defects in rRNA synthesis (Chakraburtty & Bibb, 1997). A later study further resolved this issue by showing increased transcription of *actII-ORF4* following induction of ppGpp synthesis without affecting growth rate (Hesketh *et al*, 2001). Despite an important role for ppGpp in stimulating antibiotic production under conditions of nitrogen starvation, it serves no obvious function under phosphate limiting growth conditions. In multiple *Streptomyces* species, the production of Act and Red is inversely correlated with increased levels of phosphate (Chouayekh & Virolle, 2002; Kang *et al*, 1998). This involves the two-component system PhoR-PhoP, the null mutation of which caused overexpression of the *act* and *red* clusters (Sola-Landa *et al*, 2003). In *S. coelicolor*, the response regulator PhoP negatively affects secondary metabolism by competing with AfsR for the binding site upstream of the *afsS* gene (Santos-Beneit *et al*, 2011; Santos-Beneit *et al*, 2009). Another direct target of PhoP is the *ppk* gene encoding polyphosphate kinase (PPK), disruption of which in *S. lividans* results in a remarkable increase in Act production and elevated transcription of *actII-ORF4*, *redD*, and *cdaR* (Chouayekh & Virolle, 2002). PPK catalyzes poly-phosphate formation from inorganic monophosphate when the latter is abundant; the reservoir of poly-phosphate can be hydrolyzed into monophosphate again in response to phosphate depletion (Kornberg *et al*, 1999). The activation of antibiotic production by the deletion of the *ppk* gene supports a role for inorganic phosphate as a suppressing agent of secondary metabolism since inactivation of PPK prevents the formation of such a phosphate pool and therefore reduces the endogenous phosphate level.

1.3.3 Signaling Molecules

Apart from nutritional cues, extracellular signaling molecules also contribute to the regulation of secondary metabolism. γ-butyrolactones are produced by many streptomycetes, among which the most extensively studied is A-factor (2-isocapryloyl-3*R*-hydroxymethyl-γ-butyrolactone) made by *Streptomyces griseus* (Horinouchi & Beppu, 1992) (Fig. 1.7). It is synthesized by the gene product of *afsA* and binds to its cytoplasmic receptor protein ArpA, dislodging the latter from the promoter region of the *adpA* gene. This leads to the derepression of *adpA* transcription. AdpA, in turn, stimulates the transcription of *strR*, the gene encoding the pathway-specific regulator of the streptomycin biosynthesis pathway, and a number of other genes implicated in morphological development (Kato *et al*, 2004). In addition to secondary metabolism, A-factor of *S. griseus* also influences morphogenesis and the homolog of AdpA in *S. coelicolor*, BldH, is an important regulator of aerial hyphae differentiation (Ohnishi *et al*, 2005; Takano *et al*, 2003).

By contrast, the main γ-butyrolactone in *S. coelicolor*, SCB1 [(2R, 3R, 1'R)-2-(10-hydroxy-6-methylheptyl)-3-hydroxymethylbutanolide], and its cognate binding protein, ScbR, seem to function differently. The gene encoding the homolog of AfsA, *scbA*, is embedded in the *cpk* cluster along with the *scbR* gene (Takano *et al*, 2001) (Fig. 1.3). The

two genes are adjacent to each other and are transcribed divergently (Takano *et al*, 2001). In the absence of SCB1, ScbR binds to the promoter of its own gene and to that of *cpkO* (the pathway-specific activator gene of *cpk* cluster), blocking their transcription while at the same time activating expression of *scbA* (Takano *et al*, 2001). Disrupting *scbR* results in a delay in Act and Red production, whereas deleting *scbA* impairs SCB1 production but leads to precocious Act and Red production through increased transcription of the *act* and *red* clusters (D'Alia *et al*, 2011; Takano *et al*, 2001). This enhanced expression must be an indirect effect as ScbR does not directly bind operator sequences upstream of *act* and *red* genes. It is conceivable that the activation of the *act* and *red* clusters is achieved via an intermediate regulator. An additional contribution to the enhanced antibiotic production may come from the increased levels of common precursor molecules such as malonyl-CoA, stemming from the silencing of the *cpk* cluster in the absence of SCB1.

1.3.4 Overlap of the Regulatory Pathways for Antibiotic Production and Morphological Development

Considering the unique features of the *S. coelicolor* life cycle, the temporal correlation between antibiotic production and morphogenesis dictates that the two processes share some common elements in their genetic control, which adds additional complexity to the system (Bibb, 2005). In accordance with the notion that secondary metabolism occurs in substrate mycelia before the onset of aerial hyphae formation, mutations in aerial differentiation regulatory genes *bldA*, *bldB*, *bldC*, *bldD*, *bldG*, *bldH*, and *bldJ* all impair antibiotic production (Elliot *et al*, 2007). *bldA* is considered a pleiotropic regulatory gene in some literature since its deletion renders the strain deficient in Act, Red, and CDA production (Lawlor *et al*, 1987). The dependence of Act and Red production on *bldA* is due to the presence of a TTA codon in their pathway specific regulatory genes (Fernandez-Moreno *et al*, 1991; White & Bibb, 1997). However, *cdaR* does not contain any TTA codons and it is unknown how its synthesis is controlled by *bldA*. BldD also indirectly influences antibiotic production through its target gene *cdgA*, which encodes an enzyme catalyzing the synthesis of a signaling molecule c-di-GMP (cyclic diguanylate) (den Hengst *et al*, 2010).

Secondary metabolism of *S. coelicolor* is governed in an intricate fashion, involving a vast number of pathways. While we do not yet have a complete understanding of the network that integrates various environmental and developmental signals to generate changes in the expression of regulatory and ultimately biosynthetic genes, over the next few years, the availability of entire genome sequences, high-density microarrays covering the whole genome, RNA-Sequencing (RNA-Seq) and improved genome mining technology will help us unravel the complex regulatory cascade and, more importantly, facilitate full exploitation of the biosynthetic potential of *Streptomyces*.

1.4 GENE REGULATION BY cAMP AND ITS RECEPTOR PROTEIN CRP

1.4.1 cAMP-Crp Complex in Carbon Metabolism

3'-5'-cyclic adenosine monophosphate (cAMP) is one of the most important signaling molecules in both eukaryotic and prokaryotic cells. It is synthesized intracellularly by adenylate cyclase (encoded by the *cya* gene) and is degraded by phosphodiesterase (Ullmann & Danchin, 1983). In *E. coli*, more than 98% of cAMP is secreted to the outside of cells after synthesis (Ullmann & Danchin, 1983).

More than two decades ago, the regulatory complex of cAMP and its receptor protein Crp (for cAMP receptor protein) was discovered through efforts to explain the carbon catabolite repression phenomenon in *E. coli*. This phenomenon has been observed in many microorganisms (Kolb *et al*, 1993), and it enables bacteria to adapt quickly to a preferred carbon source (more readily metabolizable) by inhibiting the synthesis of enzymes required for the use of other less desirable carbon sources. This has been occasionally misnamed the "glucose effect" since glucose was first shown to be selectively used over lactose by *E. coli* (Cohn & Monod, 1953). In the presence of abundant glucose, Enzyme II A (EIIA) from the phosphotransferase system is unphosphorylated, and is thus not able to activate adenylate cyclase; cAMP levels are therefore low inside the bacterial cells. At the same time, EIIA also interacts with and represses components of transporters for alternative carbon sources. Once glucose is depleted, EIIA is phosphorylated and adenylate cyclase is activated, resulting in higher levels of cAMP. cAMP binds to its receptor protein Crp and together, they stimulate the expression of genes encoding enzymes required to metabolize additional carbon sources (*e.g.* lactose) (Deutscher, 2008). Since its discovery in *E. coli*, Crp has become the paradigm for studying the regulatory mechanism of transcription factors (Fic *et al*, 2009).

1.4.2 Transcription Regulation by *E. coli* Crp

Like many transcription factors, *E. coli* Crp acts as a homodimer, with each monomer consisting of two domains: an N-terminal domain involved in dimerization and interaction with the effector molecule cAMP, and a C-terminal domain containing a DNA binding helix-turn-helix motif. Interaction with cAMP allosterically changes the conformation of the C-terminal domain and promotes DNA binding at specific sites in or near the promoter sequences of target genes (Busby & Ebright, 1999; Fic *et al*, 2009). Crp recognizes an imperfect palindrome sequence $TGTGA(N)_6TCACA$, with each subunit occupying half of this sequence (Busby & Ebright, 1999). This specific interaction is greatly enhanced by the presence of micromolar concentrations of the ligand, but is reduced when cAMP levels reach millimolar range (Fic *et al*, 2009). Upon DNA binding, Crp usually causes a sharp bend in the DNA, and typically helps to recruit RNA polymerase (RNAP) to the promoter regions of target genes (Berg & von Hippel, 1988; Fic *et al*, 2009).

E. coli Crp not only plays a key role in carbon catabolite repression, but also controls a vast number of genes implicated in diverse cellular functions such as motility and stress resistance (Busby & Ebright, 1999; Fic *et al*, 2009). Crp-dependent promoters can be

grouped into three classes (Busby & Ebright, 1999; Fic *et al*, 2009) (Fig. 1.9). Class I and II promoters both have a single Crp binding site and require direct Crp-RNAP contacts for transcriptional activation. This means that both Crp and RNAP have to bind on the same face of the DNA helix. The differences are: i) in class I promoters, Crp binds upstream of the -35 sequence (usually centered at around -93, -83, -72, or -62), while in class II promoters, Crp binding overlaps the -35 promoter element and is centered at position -42; ii) activation of class I promoters involves interaction between one subunit of the Crp dimer and the C-terminal domain of the RNAP α subunit, but in class II promoters, the upstream subunit of the Crp dimer interacts with the C-terminal domain of the RNAP α subunit, while the downstream Crp subunit contacts the N-terminus of the RNAP α subunit (Busby & Ebright, 1999). Class III promoters contain more than one Crp binding site, separated by varying distances. Despite the diverse architectures associated with class III promoters, activation is simply accomplished through additive combinations of class I and class II-based mechanisms (Fig. 1.8). When acting in conjunction with other regulators, Crp can bind further upstream and synergistically activate transcription, with both regulators contacting different surfaces of RNAP (Busby & Ebright, 1999). Although Crp was originally characterized as a positive regulator, it can also repress a remarkable number of genes as well (Kolb *et al*, 1993). It can either exert direct repression through promoter occlusion (occupying the RNAP or other activator protein binding sites) and RNAP progression blockage, or indirectly inhibit transcription by influencing the positioning of RNAP at other overlapping promoters (Kolb *et al*, 1993). It may also repress some genes by stabilizing the repression loops mediated by other proteins (Kolb *et al*, 1993).

1.4.3 Importance of cAMP and Crp in *S. coelicolor*

In wild-type *S. coelicolor* cells, cAMP is synthesized at low levels throughout the life cycle, with peak accumulation at transition phase in liquid cultures or during germination and aerial hyphae differentiation in surface-grown cultures (Ruiz *et al*, 2010; Susstrunk *et al*, 1998). The open reading frame *SCO3571* has been annotated as the only *crp*-like gene in the *S. coelicolor* genome (Bentley *et al*, 2002). It encodes a protein of 224 amino acids. Unlike the classic model established in *E. coli*, the cAMP-Crp signal transduction system is believed to be nonessential for relieving glucose-mediated catabolite repression in *S. coelicolor*, because cAMP levels do not change when the carbon source is changed from glucose to galactose, and moreover, a *cya* deficient mutant can grow and differentiate on media containing various carbon sources (Hodgson, 1982; Susstrunk *et al*, 1998). However, disruption of the *cya* gene has a significant impact on spore germination: more than 85% of the spores are unable to germinate at all, and those that do germinate are extremely delayed relative to wild-type. When cultured on unbuffered medium, the *cya* mutant exhibits a "bald" phenotype, and is unable to make antibiotics (Susstrunk *et al*, 1998). This is proposed to be due to the inability of this strain to neutralize the acids generated during vegetative growth. All these defects can be suppressed by adding high concentrations of exogenous cAMP early during growth, transfer to buffered medium, or cultivating in close proximity to a wild-type strain (Susstrunk *et al*, 1998).

Deletion of the *crp* gene brings about very similar phenotypes to those of the *cya* mutant, including severely reduced spore viability, delayed germination, markedly smaller colony size, and diminished Act production (Derouaux *et al*, 2004b). However, the *cya* and *crp* mutants differ in that the *crp* mutant initiates aerial hyphae differentiation normally, regardless of the medium pH and, moreover, displays enhanced sporulation (Derouaux *et al*, 2004b). By contrast, overexpressing Crp from a high-copy-number plasmid confers a "bald" phenotype and precocious Act production (Piette *et al*, 2005). Altogether, these data suggest that Crp is important for regulating morphological development and also affects actinorhodin production. To better understand how Crp functions in the cell, a comparative proteomic study evaluated the changes in protein levels between dormant and germinating spores in the wild-type and *crp* mutant strains using 2-D gel electrophoresis (Piette *et al*, 2005). The results generally show lower reactivation of essential metabolic activities in the *crp* mutant in response to spore germination, since a number of proteins involved in metabolism, protein synthesis, transcription, and respiration are expressed at considerably lower levels compared to the wild-type (Piette *et al*, 2005).

At a molecular level, a partial palindromic sequence that matches the consensus binding motif of the *E. coli* Crp is present in the promoter region of *SCO3571*, and is found eight times in the upstream regions of *S. coelicolor* genes (Derouaux *et al*, 2004b). However, although cAMP binding has been shown for *S. coelicolor* Crp using cAMP affinity chromatography, gel mobility shift assays using purified protein have failed to demonstrate interactions between Crp and this consensus sequence, even when Crp is purified in a cAMP-bound form from *Streptomyces* cells (Derouaux *et al*, 2004a; Derouaux *et al*, 2004b). It is therefore proposed that *S. coelicolor* Crp may function in a mechanistically different manner than the *E. coli* Crp. It may recognize a different binding site, or it may require another co-factor to perform its DNA-binding function.

1.4.4 Functions of Crp in Other Actinomycetes and Some Gram Negative Pathogens

Two other well characterized Crp orthologs in the actinobacteria are Rv3676 from *M. tuberculosis* and GlxR from *C. glutamicum* (Fig.1.9). Rv3676 shares 54.9% overall sequence identity with *S. coelicolor* Crp, and their DNA binding domains are 95.6% identical. Previous studies have shown that cAMP levels within *Mycobacteria* increase dramatically upon entry into macrophages (Bai *et al*, 2009) and the *Rv3676* null mutant is attenuated for survival and replication in mice and macrophages (Rickman *et al*, 2005). The mutant strain also shows altered transcription of many genes, among which *rpfA* and *whiB1* exhibit the greatest reduction in expression (Rickman *et al*, 2005). *rpfA* encodes a resuscitation promoting factor involved in regulating entry into, and exit from, bacterial dormancy, while *whiB1* encodes a Wbl family protein that binds a redox-sensitive [4Fe-4S] cluster and potentially controls developmental processes (Haiser *et al*, 2009; Mukamolova et al, 2002; Soliveri *et al*, 2000). Gel mobility shift assays showed that Rv3676 activates expression of these two genes directly by binding to a consensus motif

similar to that of the *E. coli* Crp, GTGNNA$(N)_5$CACA, in their promoter regions (Agarwal *et al*, 2006; Rickman *et al*, 2005). However, it has not been determined whether cAMP is absolutely required for the DNA binding of Rv3676 *in vitro* (Agarwal *et al*, 2006; Rickman *et al*, 2005). In contrast to the *E. coli* and *S. coelicolor* Crp, which are not necessary for bacterial survival, deletion of the *glxR* gene in *C. glutamicum* causes very severe growth defects and this has greatly hindered assessment of the physiological role of GlxR (Kim *et al*, 2004). Recently, chromatin immunoprecipitation-microarray (ChIP-chip) analyses revealed around 200 binding targets of GlxR across the genome of *C. glutamicum* (Toyoda *et al*, 2011). Gel mobility shift assays and promoter-reporter fusion assays verified that GlxR activates genes for aerobic respiration, ATP synthesis and glycolysis and represses a citrate uptake gene through direct interaction with the consensus Crp binding motif in the promoter sequences of these genes (Toyoda *et al*, 2011). Although the binding affinity is significantly decreased in an *cya* deficient mutant, cAMP is not strictly required for the function of GlxR (Toyoda *et al*, 2011).

Crp family proteins also have immense medical relevance since their orthologs encoded by pathogenic bacteria are often important for virulence. For example, Vfr (virulence factor regulator) of *P. aeruginosa* acts as a global virulence regulator, activating the production of exotoxin A, type IV pili, and a type III secretion system involved in virulence factor secretion. It also indirectly turns on many other virulence genes by regulating the *las* quorum sensing system (Fuchs *et al*, 2010; West *et al*, 1994). Similarly, the *Yersinia enterocolitica* Crp is necessary for normal expression of the virulence proteins Ysc and Ysa, and a type III secretion system (Petersen & Young, 2002).

1.5 AIMS OF THIS STUDY

For a long time, the cAMP-Crp signal transduction system in *E. coli* has been a model system for studying transcription regulation. However, more and more research is revealing that Crp regulation is far from being thoroughly understood, and that what is true in *E. coli* can not necessarily be directly applied to other bacteria.

Previous studies on *S. coelicolor* Crp had focused on its role in germination, but efforts to characterize its activity at a more molecular level were hampered by an inability to establish Crp-DNA interactions *in vitro*. Here, I started off by performing more in-depth phenotypic analyses of the *S. coelicolor crp* mutant including germination assays, stress tests, antibiotic production assays and scanning/transmission electron microscopy. This was followed by elucidating the downstream Crp targets using chromatin-immunoprecipitation and transcriptome profiling assays. Both phenotypic assays and target identification pointed towards a role for Crp as a master regulator controlling multiple aspects of the *S. coelicolor* life cycle. Unexpectedly, I found that multiple antibiotic biosynthesis clusters were upregulated upon Crp expression, prompting us to redirect our research focus to investigating the regulation of secondary metabolism by the *S. coelicolor* Crp. Bioassays and comparative metabolic profiling were subsequently

performed to evaluate the impact of Crp overexpression in its native and heterologous hosts. Although it is challenging to delineate a consensus Crp binding motif due to the difficulty in showing Crp-DNA interaction *in vitro*, the work presented here will still greatly help decipher the precise function of *S. coelicolor* Crp.

1.6 FIGURES AND TABLES:

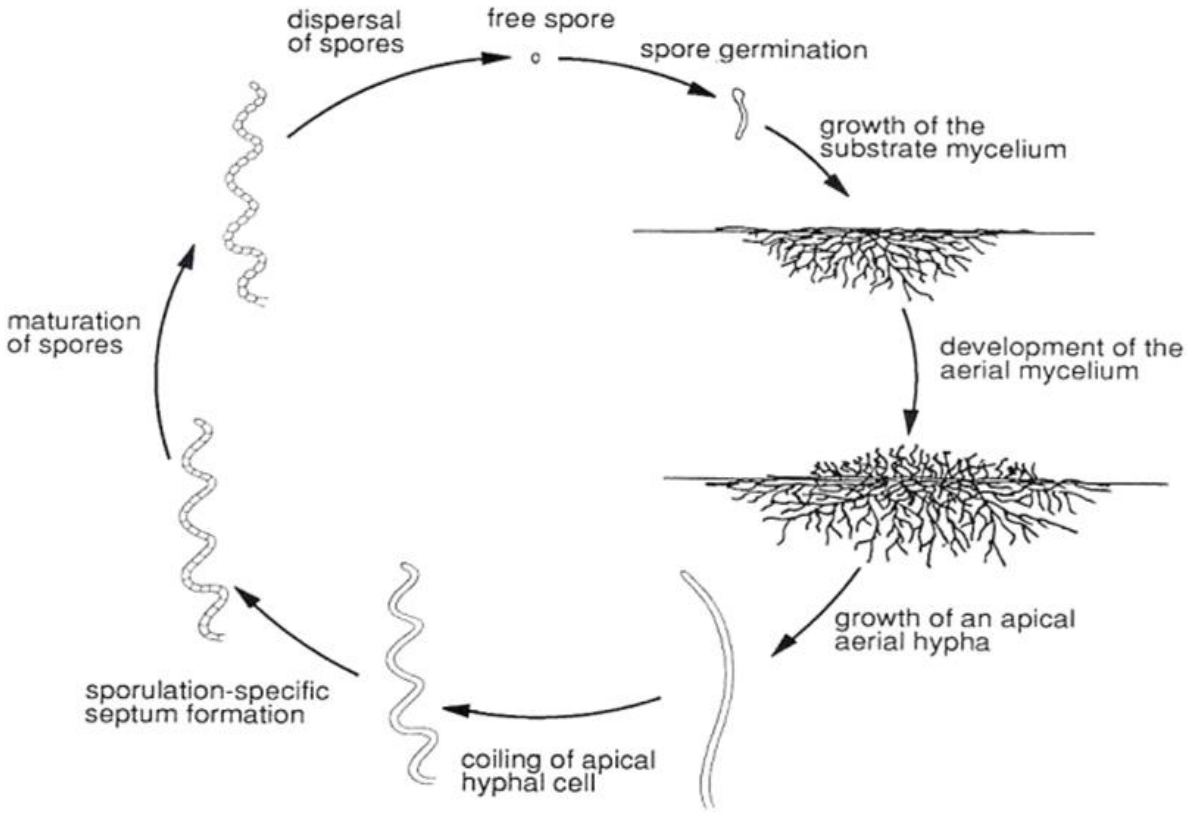

Fig.1.1 Life cycle of *Streptomyces coelicolor*. A spore germinates in response to as yet unknown environmental signals. Germ tubes elongate by tip extension. The resulting vegetative hyphal filaments branch frequently, producing a thick network of hyphae termed the substrate mycelium. As nutrients are depleted, the newly emerging filaments are coated in a hydrophobic sheath, allowing them to grow away from the substrate hyphae. Each of these aerial filaments is fated to undergo a single synchronous cell division event along its length, forming ~50-100 uninucleoid cells that go on to develop into spores.

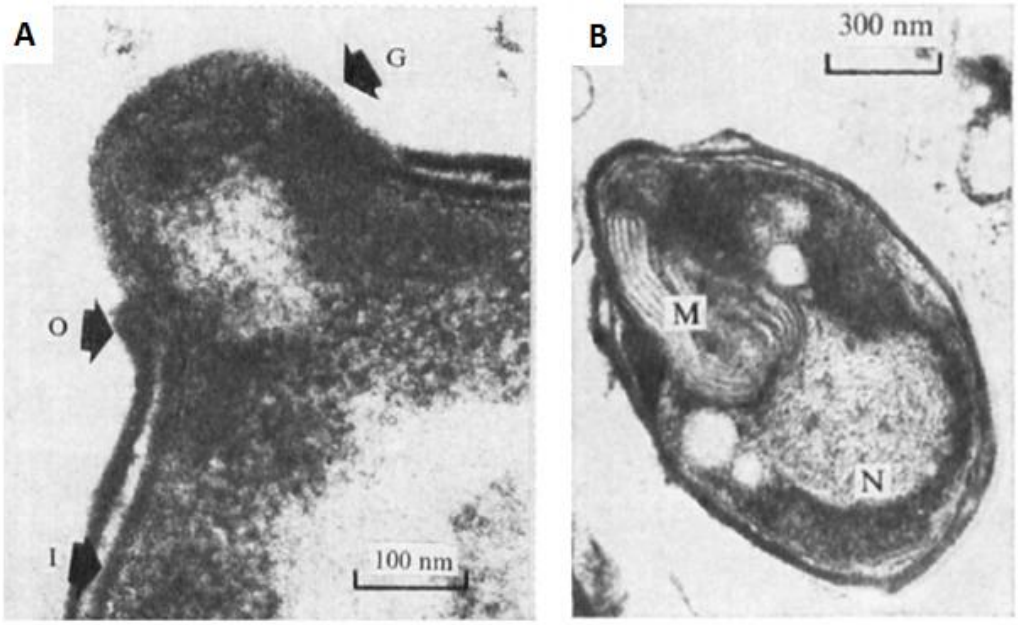

Fig.1.2 Transmission electron micrographs of germinating *Streptomyces ostreogriseus* spores. (A) Germ tube emergence: (G) a germ tube; (O) ruptured spore wall outer layer; (I) continuous spore wall inner layer. (B) Visible nuclear material in germinating spores: (M) mesosome; (N) distinct nuclear material. Image modified from: (Gerhardt *et al*, 1976)

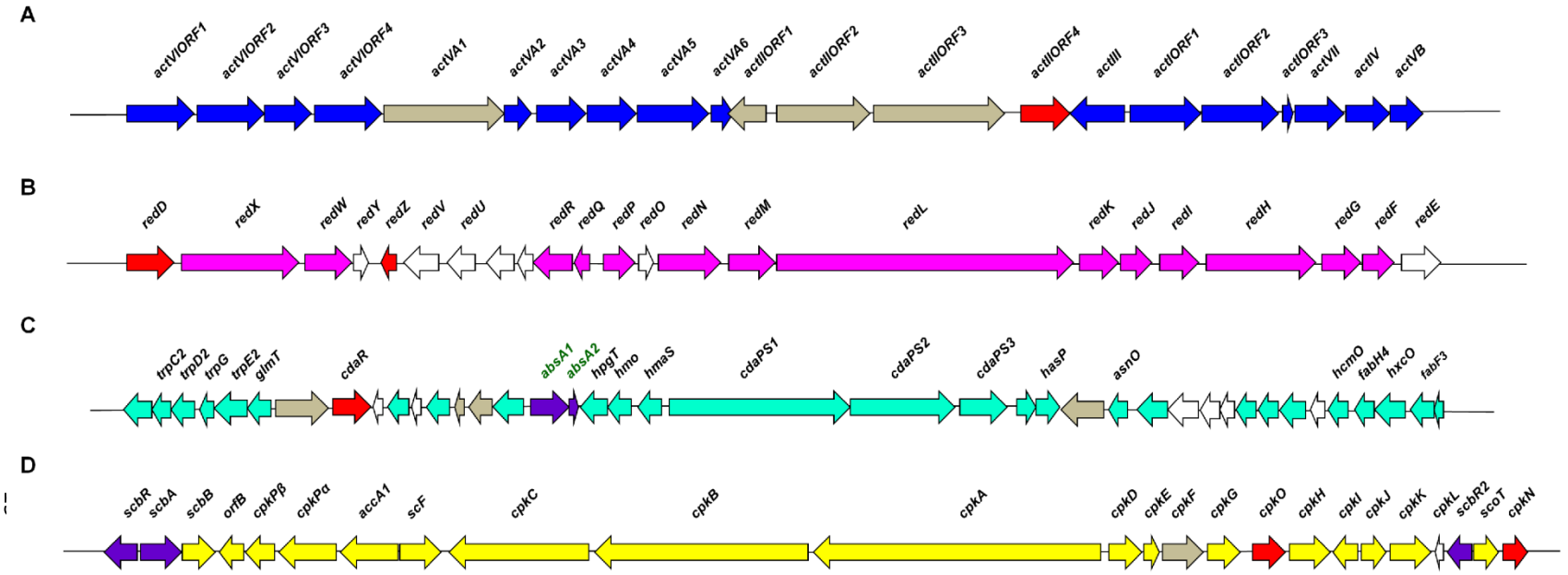

Fig. 1.3 Biosynthesis clusters for (A) actinorhodin (*act*), (B) undecylprodigiosin (*red*), (C) calcium-dependent antibiotic (*cda*) and (D) yellow cryptic polyketide (*cpk*). The genes encoding the biosynthesis enzymes are shaded in blue (*act*), pink (*red*), cyan (*cda*) and yellow (*cpk*) respectively. The resistance genes are shaded in grey. The pathway-specific and global regulatory genes are highlighted in red and purple respectively.

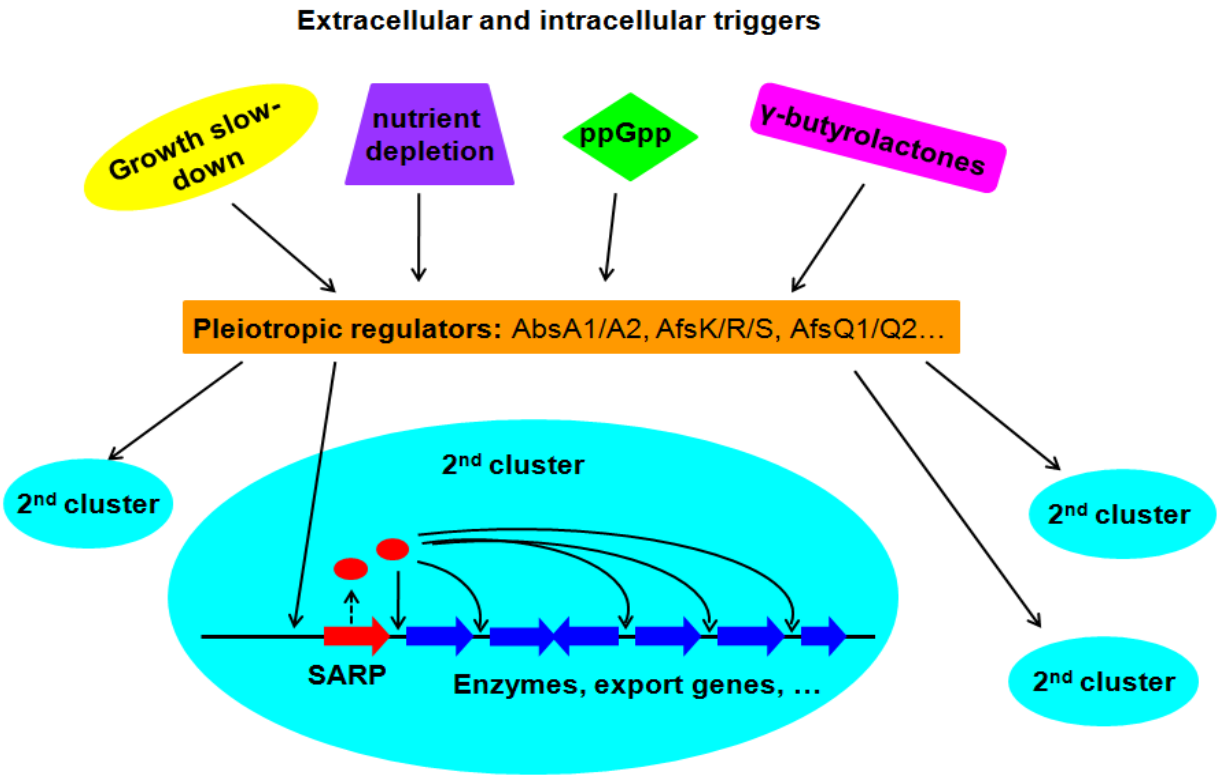

Fig. 1.4 Simplified secondary metabolism regulation cascade in *S. coelicolor*. In response to triggers such as growth slow-down, nutrient depletion, and signaling molecules, pleiotropic regulators are activated and then transmit the signals to individual secondary metabolic clusters by stimulating the transcription of the pathway-specific regulatory genes (most of which encode SARP family proteins). The pathway-specific regulators in turn promote the expression of their cognate clusters, resulting in secondary metabolite production. 2nd cluster: secondary metabolic clusters.

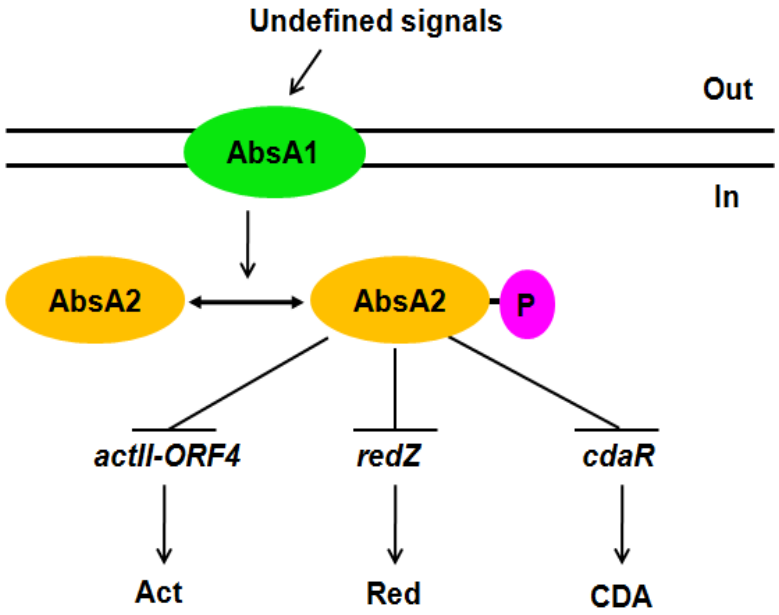

Fig. 1.5 Regulatory mechanism of the pleiotropic regulator AbsA1/A2. Depending on the nature of the signals, the sensor kinase AbsA1 can phosphorylate (and activate) the response regulator AbsA2 so that the latter can directly interact with the promoter regions of *actII-ORF4*, *redZ*, and *cdaR* to reduce their transcription. Alternatively, AbsA1 can also relieve this repression by dephosphorylating AbsA2-P.

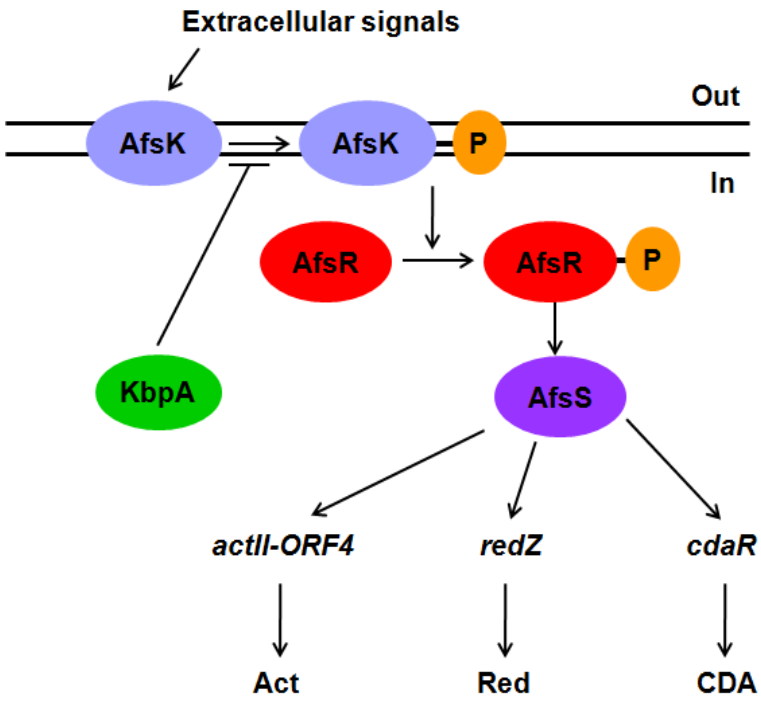

Fig. 1.6 Mechanism of the pleiotropic regulatory system AfsR/S/K. Once triggered by environmental signals, the membrane bound kinase AfsK autophosphorylates and subsequently phosphorylates the pleiotropic regulator AfsR. Once phosphorylated, AfsR is capable of stimulating the *afsS* gene, whose gene product enhances antibiotic production through an unknown mechanism. The pathway can be blocked by KbpA, which binds unphosphorylated AfsK and inhibits its autophosphorylation.

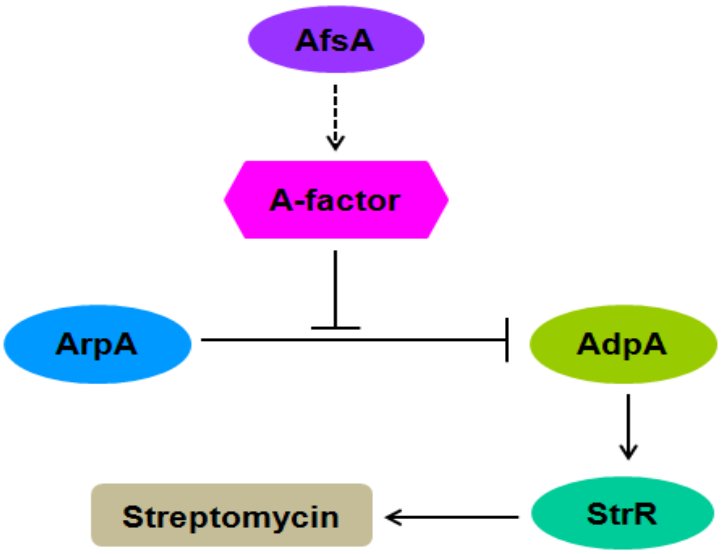

Fig. 1.7 Regulation of streptomycin production by A-factor in *S. griseus*. Upon synthesis by the gene product of *afsA*, A-factor binds its receptor protein ArpA, releasing its repression of the *adpA* gene. AdpA, when expressed, passes on the signal by stimulating the transcription of the pathway-specific regulator of the streptomycin synthesis cluster (*strR*), leading to the synthesis of streptomycin.

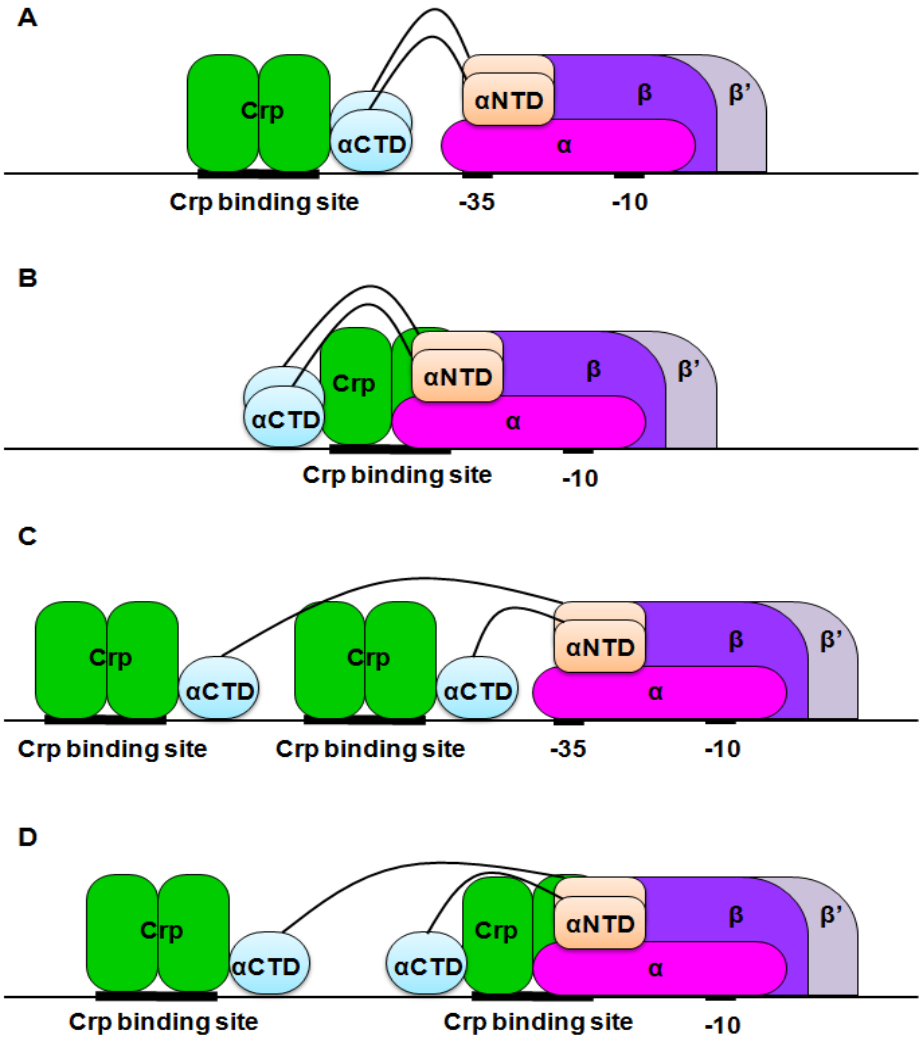

Fig. 1.8 Transcription activation of different promoters by *E. coli* Crp. (A) Class I promoter: the Crp homodimer binds to a single site upstream of the -35 promoter sequence (usually centered around -93, -83, -72, or -62) and interacts with the C-terminal domain of the RNAP α subunit through the downstream subunit of the Crp homodimer. (B) Class II promoter: the Crp binding site overlaps the -35 promoter sequence. The upstream subunit of the Crp homodimer interacts with the C-terminal domain of the RNAP α subunit. (C) and (D) Class III promoters: the promoter has multiple Crp binding sites (only two sites are shown in the diagram), with varying distances between the Crp binding sites, and between the RNP binding site and Crp binding sites. Activation at such promoters is accomplished through additive combinations of class I and class II mechanisms. Image is recreated based on (Busby & Ebright, 1999).

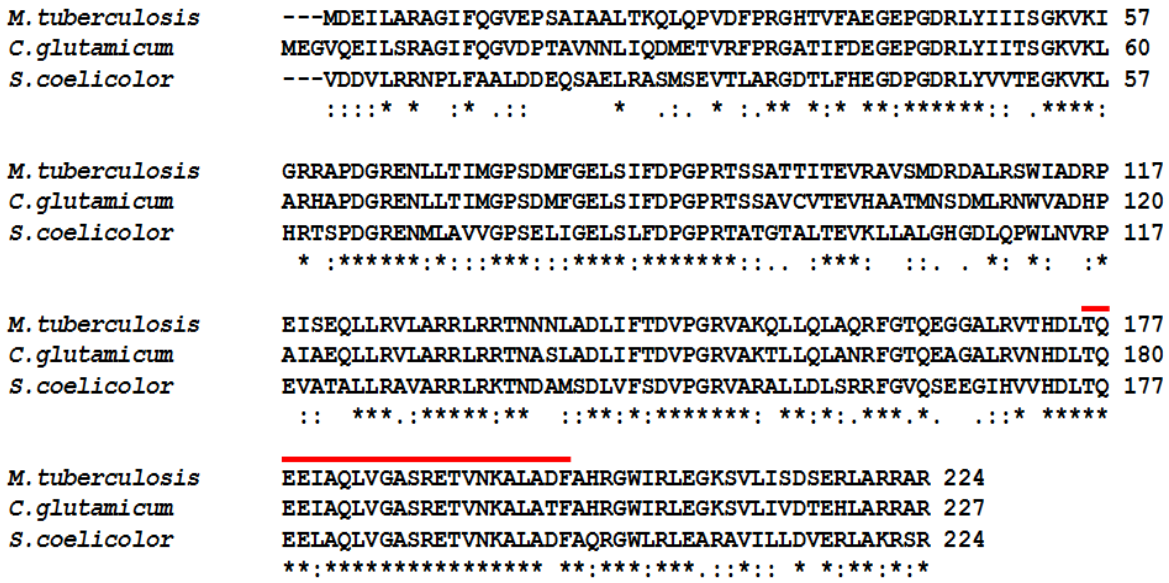

Fig. 1.9 Multiple sequence alignment of Crp from three actinomycetes: *M. tuberculosis*, *S. coelicolor*, and *C. glutamicum*. The helix-turn-helix motif for DNA binding is marked by a red bar on top. (*) indicates a fully conserved residue; (:) indicates a strongly conserved residue; (.) indicates a weakly conserved residue.

Table 1.1 Summary of genes involved in morphogenesis

	Product	Reference
divIVA	Coiled-coil protein	(Flardh, 2003; Hempel *et al*, 2008)
chpA-H	Hydrophobic cell wall-associated proteins	(Claessen *et al*, 2003; Elliot *et al*, 2003)
rdlA/B	Secreted cell wall-associated protein	(Claessen *et al*, 2004; Claessen *et al*, 2002)
ramCSAB	Biosynthetic cluster for lantibiotic-like peptide surfactant SapB	(Kodani *et al*, 2004; Ma & Kendall, 1994)
bldA	tRNA for the rare Leu codon UUA	(Lawlor *et al*, 1987)
bldC	MerR-like DNA-binding protein	(Hunt *et al*, 2005)
bldD	Small DNA-binding transcription factor	(Elliot *et al*, 1998)
bldG	Anti-anti-sigma factor	(Bignell *et al*, 2000)
bldH	DNA-binding transcription factor;	(Nguyen *et al*, 2003; Takano *et al*, 2003)
bldK	ATP-binding cassette oligopeptide transporter	(Nodwell *et al*, 1996; Nodwell *et al*, 1999)
bldM	"Orphaned" response regulator	(Molle & Buttner, 2000)
bldN	Sigma factor	(Bibb *et al*, 2000)
whiA	Regulatory protein	(Ainsa *et al*, 2001)
whiB	Wbl family protein	(Jakimowicz & Cuschieri, 2005; Soliveri *et al*, 2000)
whiD	Wbl family protein	(Jakimowicz & Cuschieri, 2005; Molle *et al*, 2000)
whiE	Locus specifying the synthesis of the grey type II polyketide spore pigment	(Kelemen *et al*, 1998)
whiG	Sigma factor	(Chater *et al*, 1989)
whiH	GntR-like transcription factor	(Ryding *et al*, 1998)
whiI	Two component system response regulator	(Ainsa *et al*, 2000)
whiJ	Transcription factor	(Ainsa *et al,* 2010)
sigF	Stress response sigma factor	(Potuckova *et al*, 1995)
ftsZ	Tubulin-like cell division protein	(McCormick *et al*, 1994)
ssgA	Small protein	(Willemse *et al*, 2011)
ssgB	Small protein	(Willemse *et al*, 2011)
mreB	Actin-like ATPase	(Mazza *et al*, 2006)
parAB	ATPase (ParA) and DNA binding protein (ParB)	(Jakimowicz *et al*, 2005; Kim *et al*, 2000)
ftsK	Membrane-associated DNA translocase	(Wang *et al*, 2007)

Table 1.2 Summary of genes involved in secondary metabolism

Gene name	Product	Reference
actIIORF-4	SARP family pathway-specific activator of the *act* cluster	(Arias *et al*, 1999)
redD	SARP family pathway-specific activator of the *red* cluster	(Takano *et al*, 1992)
redZ	"Orphaned" response regulator	(Huang *et al*, 2001; White & Bibb, 1997)
cdaR	SARP family pathway-specific activator of the *cda* cluster	(Ryding *et al*, 2002)
cpkO	SARP family pathway-specific activator of the *cpk* cluster	(Pawlik *et al*, 2007)
scbR	γ-butyrolactone receptor	(Takano *et al*, 2001)
scbR2	γ-butyrolactone receptor	(*Gottelt et al*, 2010)
absA1/A2	Two-component system	(Brian *et al*, 1996; McKenzie & Nodwell, 2007)
afsK	Membrane-associated kinase	(Petrickova & Petricek, 2003)
afsR	SARP family positive pleiotropic antibiotic regulator	(Floriano & Bibb, 1996)
afsS	Sigma factor like protein	(Vogtli *et al*, 1994)
atrA	TetR-like regulator	(Uguru *et al*, 2005)
cprA/B	A-factor receptor protein homologue	(Onaka *et al*, 1998)
cutR/S	Two-component system	(Chang *et al*, 1996)
ecrA1/A2	Two-component system	(Li *et al*, 2004)
afsQ1/Q2	Two-component system	(Ishizuka *et al,* 1992)
abrA1/A2	Two-component system	(Yepes *et al*, 2011)
abrC1/C2/C3	Two-component system	(Yepes *et al*, 2011)
absB	RNase III	(Price *et al*, 1999)
dasR	GntR-like transcription factor	(Rigali *et al*, 2008)
relA	GTP pyrophosphokinase	(Chakraburtty & Bibb, 1997)
ppk	Polyphosphate kinase	(Kornberg *et al*, 1999)
phoR/P	Two-component system	(Santos-Beneit *et al*, 2011; Sola-Landa *et al*, 2003)

CHAPTER 2:

CRP: A MASTER REGULATOR OF DEVELOPMENT AND ANTIBIOTIC PRODUCTION IN *STREPTOMYCES COELICOLOR*

2.1 INTRODUCTION

Streptomyces are multicellular Gram positive bacteria with a complex life cycle consisting of germination, vegetative mycelia growth, aerial hyphae development, and spore morphogenesis (Elliot *et al*, 2007; Flardh & Buttner, 2009) The formation of aerial hyphae coincides with the production of a plethora of secondary metabolites, including anti-cancer agents, immunosuppressants and two-thirds of clinically useful antibiotics. *Streptomyces coelicolor* produces four well characterized antibiotics: actinorhodin (Act) (Wright & Hopwood, 1976a), undecylprodigiosin (Red) (Malpartida *et al*, 1990; Rudd & Hopwood, 1980), methylenomycin (Mmy) (Chater & Bruton, 1985; Wright & Hopwood, 1976b), and calcium-dependent antibiotic (CDA) (Chong *et al*, 1998). They are all synthesized through the cooperative activity of the enzymes encoded by their respective secondary metabolic gene clusters (Bibb, 2005). In addition to these structural genes, the clusters also contain pathway-specific regulatory genes whose products are responsible for activating the transcription of the entire clusters (Bibb, 2005). Apart from the well characterized antimicrobial secondary metabolites, there also exist 'orphan' gene clusters for which no product has been assigned (Bentley *et al*, 2002). Recently, a yellow-pigmented polyketide (yCPK) was discovered following deletion of the regulatory *scbR2* gene from the yCPK biosynthetic cluster (Gottelt *et al*, 2010). ScbR2 is a γ-butyrolactone receptor protein that represses yCPK production in the absence of ligand binding (Gottelt *et al*, 2010). The pathway-specific regulators ActII-ORF4 (encoded by *SCO5085* in the *act* cluster) (Arias *et al*, 1999), RedD (encoded by *SCO5877* in the *red* cluster) (Takano *et al*, 1992), CdaR (encoded by *SCO3217* in the *cda* cluster) (Ryding *et al*, 1998), CpkO (encoded by *SCO6080* in the *cpk* cluster) (Takano *et al*, 2005), and MmyB (encoded by *SCP1.230* in the *mmy* cluster) (O'Rourke et al, 2009) were proven or presumed to positively regulate the production of Act, Red, CDA, yCPK and methylenomycin, respectively, through interactions with promoter regions within the individual clusters. Apart from the pathway-specific regulators, these biosynthetic clusters are also governed by pleiotropic antibiotic regulators that control multiple gene clusters. These include the AbsA two-component system, RelA, AfsQ1/Q2, AfsK/R/S, and AbsB (Chakraburtty & Bibb, 1997; Floriano & Bibb, 1996; Ishizuka *et al*, 1992; Price *et al*, 1999). They all function in a global manner to control the transcription of antibiotic biosynthesis genes. Secondary metabolism is also usually coordinated with bacterial growth. In liquid culture, *S. coelicolor* does not differentiate or sporulate and secondary metabolism initiates during stationary phase, whereas in surface-grown cultures, secondary metabolism coincides with the onset of morphological differentiation, indicating that the two processes share some common elements in their genetic control. This adds additional complexity to the system (Bibb, 2005).

Cyclic AMP (cAMP) is a low molecular weight signaling molecule used by both eukaryotes and prokaryotes for intracellular and intercellular communication (Susstrunk *et al*, 1998). In *S. coelicolor*, disruption of the gene encoding adenylate cyclase (*cya*), the enzyme that catalyzes the synthesis of cAMP, results in a significant delay in germination and aerial hyphae formation, and leads to defective Act production. These defects can be rescued by the exogenous addition of cAMP (Susstrunk *et al*, 1998). In *Escherichia coli*, cAMP is the effector molecule for the cAMP receptor protein (Crp), which is a global regulator of over one hundred genes, most notably those involved in utilizing alternative carbon sources other than glucose (Fic *et al*, 2009). cAMP binds to the N-terminus of Crp, leading to a conformational change in the C-terminal domain that promotes DNA interaction with an imperfect palindrome sequence $TGTGA(N)_6TCACA$ in or near the promoter sequences of target genes. The open reading frame *SCO3571* in the *S. coelicolor* genome has been annotated as a potential *crp*-like gene (Bentley *et al*, 2002). Similar to the *cya* mutant, deletion of this gene significantly delays spore germination and reduces Act production. However, this mutation also causes accelerated sporulation, in contrast to the *cya* strain which failed to form fluffy aerial hyphae on unbuffered media (Derouaux *et al*, 2004b; Susstrunk *et al*, 1998). Conversely, overexpression of *crp* has no effect on germination and vegetative growth, but the strain exhibits precocious antibiotic production and an inability to form aerial hyphae and spores (Piette *et al*, 2005).

A partial palindrome [$GTG(N)_8CAC$] that matches the consensus binding motif of *E. coli* Crp is present upstream of *SCO3571*, and is found eight times in the upstream regions of *S. coelicolor* genes (Derouaux *et al*, 2004b). Although the *S. coelicolor* Crp protein binds cAMP, as shown using cAMP affinity chromatography, gel shift assays using purified protein have failed to demonstrate interaction with this consensus sequence, even when the protein is purified in a cAMP-bound form from *Streptomyces* cells (Derouaux *et al*, 2004a; Derouaux *et al*, 2004b). Moreover, *S. coelicolor* Crp does not complement the maltose- and lactose-negative phenotype of *E. coli crp*- strain (when immunoblots confirmed the expression of *S. coelicolor* Crp in *E. coli*), indicating that *S. coelicolor* Crp may recognize a different binding sequence than its *E. coli* counterpart, or that it may function in a mechanistically different manner (Derouaux *et al*, 2004a). Proteomic studies identified an array of proteins having different expression profiles in the wild-type and *crp* strains. These proteins play important roles in diverse cellular activities such as protein translation and folding, carbohydrate metabolism, nucleotide metabolism, and small molecule transport (Piette *et al*, 2005). Based on these results, Crp is hypothesized to be a global regulator engaged in multiple stages during the development of *S. coelicolor*. In this study, I investigated the impact of Crp on the cellular development and metabolism of *S. coelicolor*. For the first time, I uncover the regulatory network controlled by Crp and report a direct role for Crp as a global antibiotic regulator.

2.2 MATERIALS AND METHODS

2.2.1 Bacterial strains, plasmids, and culture conditions

Streptomyces strains, *Escherichia coli* strains, and all plasmids used in this study are summarized in Table 2.1. *Streptomyces* strains were grown at 30°C on solid MS (mannitol, soy flour), R2YE (rich), Difco nutrient agar (Difco) and R5 (rich) agar media or in liquid R5, as described previously (Kieser, 2000). *E. coli* strains were grown at 37°C on solid or in liquid LB (Luria Bertani) medium or in liquid 2 × YT (yeast-tryptone) broth unless otherwise indicated (Kieser, 2000). Antibiotics were added to maintain plasmids when necessary. For *E. coli* cultures, antibiotics were added to the following final concentrations (μg/ml): ampicillin, 100; apramycin, 50; chloramphenicol, 25; hygromycin B, 50; kanamycin sulphate, 50. For *Streptomyces* cultures, antibiotics were used at the following final concentrations (μg/ml): apramycin, 50; hygromycin B, 50; nalidixic acid, 25; thiostrepton, 50. To obtain *Streptomyces* spore stocks, cells were cultured on agar medium overlaid with cellophane disks until sporulation was obvious. Spores were scraped off, pelleted, and stored in 20% glycerol at -20°C (Kieser, 2000). *E. coli* stocks were made by mixing overnight liquid cultures and 40% glycerol in a 1:1 ratio, followed by storage at -80°C.

2.2.2 Oligonucleotides

All DNA oligonucleotides used in this study are summarized in Table 2.2.

2.2.3 Mutant strain construction

The *crp* deletion mutant was created using a standard PCR-targeted gene replacement method (Gust *et al*, 2003). The *crp* gene in the cosmid StH17 was replaced by the *oriT*-containing *aac(3)IV* (apramycin) resistance cassette of pIJ773 (Redenbach *et al*, 1996). The mutant cosmid carrying the *crp* deletion was confirmed by digestion and PCR amplification using different combinations of primers annealing upstream, within, and downstream of the deleted gene. The cosmid was subsequently conjugated into *S. coelicolor*, screened for double-crossover recombination, and confirmed using PCR.

2.2.4 Plasmid construction and conjugation into *Streptomyces*

To generate a complementation construct, the *crp* gene was PCR amplified from the cosmid StH17 along with 100 nt upstream and 180 nt downstream sequences so as to include all regulatory elements. The fragment was digested with *Xba*I/*Bam*HI and cloned into the integrating *Streptomyces* vector pIJ82 digested with the same enzymes (Table 2.1). To create an inducible *tipA-crp* construct, the *crp* gene and extended downstream sequence (180 nt) were PCR amplified and cloned into pCR2.1-TOPO vector (Invitrogen). The clone containing the correct insert was digested with *Nde*I/*Eco*RI and the fragment was subcloned downstream of the *tipA* promoter in the integrating *Streptomyces* vector pIJ6902 (Table 2.1). A constitutive *crp*-overexpression plasmid was made by first cloning the *crp* gene and its downstream sequence (180 nt) into the *Eco*RI/*Bam*HI sites downstream of the *ermE** promoter in the pMC500 vector (Table 2.1)

before *ermE*-crp* was excised using *Bgl*II, and inserted into the *Bam*HI site of pIJ82. Plasmids were introduced into *Streptomyces* strains via conjugation from the non-methylating *E. coli* strain ET12567 containing the conjugation 'helper' plasmid pUZ8002 (Kieser, 2000). The *E. coli* strain ET12567/pUZ8002 carrying the plasmid to be transferred was grown in liquid LB until OD_{600} reached 0.4-0.6 before being washed with 2 × YT and mixed together with heat shocked *S. coelicolor* spores (heat shocked at 50°C for 10 min). The mixture was then spread onto MS agar containing 10 mM $MgCl_2$ and incubated for 16 h. The plates were overlaid with 1 ml sterile dH_20 supplemented with 25 µl apramycin (50 mg/ml), 6 µl thiostrepton (50 mg/ml) or 25 µl hygromycin B (50 mg/ml) to select for plasmids, and 20 µl naladixic acid (25 mg/ml) to kill the *E. coli* donor strain.

2.2.5 Germination assay

Spores of the wild-type, *crp* mutant and *crp* (pIJ6902*crp*) strains were spread on MS agar overlaid with cellophane disks and incubated at 30°C for 12 h. A cellophane disk square was excised every 2 h and viewed under a light microscope (10 × times 100 ×) to score germinated and non-germinated spores. Germination rates were calculated as percentages of germinated spores versus the total population under the scope. Three independent experiments were performed, and in each experiment, approximately 200 spores were counted for each strain.

2.2.6 Heat shock assay

Spores of the wild-type and *crp* mutant strains were diluted in distilled water (dH_2O) to a final concentration of ~300 spores/50 µl. Aliquots of the diluted spores were incubated at 60°C for 10, 20, 40, 60, and 90 min, followed by cultivation on DNA agar at 30°C. The number of surviving spores was quantified after 3 days for the wild-type strain and 5 days for the *crp* mutant strain. Survival rates were determined by dividing the number of colonies on plates after heat treatment by the number of colonies on plates without heat exposure, and were expressed as percentages. Results were obtained from three independent experiments.

2.2.7 Scanning and transmission electron microscopy

The wild-type, *crp* deletion mutant and *crp* (pIJ6902*crp*) strains were streaked on MS agar and incubated for 6 days to obtain single colonies. Individual colonies were excised from the plates and fixed overnight in 2% glutaraldehyde. The colonies were then rinsed twice in dH_2O and post-fixed in 1% osmium tetroxide for 1 h. Subsequently, samples were dehydrated using an ethanol gradient (50%, 70%, 95%, and 100% ethanol). For scanning electron microscopy (SEM), samples were dried on a critical point dryer and then mounted onto SEM stubs, where they were coated with gold before being viewed on a JEOL JSM840 SEM. For transmission electron microscopy (TEM), the samples were dehydrated in 100% propylene oxide, before being infiltrated with Spurr's resin and transferred to an embedding mold. Once on the mold, they were polymerized in Spurr's

resin at 60°C overnight. Sections were then sliced using a Leica UCT ultramicrotome and viewed using a JEOL JEM 1200 TEM. All images were analyzed using ImageJ 1.41a (Abramoff *et al*, 2004) and Adobe Photoshop CS version 8.0 software. The spore wall thickness was determined for 100 spores for each strain.

2.2.8 Cell extract preparation, SDS-PAGE, and immunoblotting

Streptomyces cells grown in liquid R5 were pelleted by centrifugation at 16,000 × g at 4°C for 5 min, washed once with 10% sucrose, and then resuspended in appropriate volumes (4 µl/mg biomass) of lysis buffer [10 mM Tris-HCl pH 8.0, 50 mM NaCl, 3 mg/ml lysozyme, 1 × protease inhibitor (Roche)] . The mixtures were incubated at 30°C for 30 min and sonicated at 50% power for 4 cycles of 15 s each (Branson, Sonifier Cell Disruptor 350). Cell debris was removed by twice centrifuging at 16,000 × g at 4°C for 25 min. Total protein concentration of the supernatant was measured using a Bradford assay (Bradford, 1976). To standardize total protein concentrations of all samples, lysis buffer was added where necessary. Cell extracts were then mixed with equal volumes of 2 × Laemmli buffer (Laemmli, 1970) and boiled at 100°C for 5 min. Ten microliters of each sample were separated by electrophoresis in a 12.5% (w/v) denaturing SDS polyacrylamide gel at 180 V for 1 h (Laemmli, 1970). To ensure equal sample loading in each lane, the gels were immersed in Coomassie blue staining solution (0.1% Coomassie Brilliant Blue R-250, 50% methanol, 10% acetic acid) and microwaved for 30 s, followed by staining on a shaker for 30 min, after which the gels were shaken in destaining buffer (40% methanol and 10% glacial acetic acid) overnight. For immunoblotting, gels and 0.45 µm PVDF-Plus transfer membranes (GE) were equilibrated in transfer buffer (25 mM Tris, 192 mM Gly, 10% methanol) for 20 min. Proteins were transferred to membranes in a Trans-Blot semi-dry transfer cell (Bio-Rad) at 15 V for 1 h. At room temperature, membranes were blocked in TBS-T (Tris-buffered saline, 0.05% Tween 20) containing 5% (w/v) low fat milk for 1 h, and were then blotted using an anti-Crp polyclonal antibody (1:2000) for 1 h. Membranes were washed with TBS-T for 3 × 10 min, before being incubated with an anti-Rabbit IgG horseradish peroxidase-linked secondary antibody (1:3000 dilution; Cell Signaling) for 40 min. After membranes were washed (3 × 15 min), Amersham ECL™ Western Blotting Detection Reagents (GE Healthcare) and Biomax XAR film (Kodak) were used for detection.

2.2.9 Protein overexpression and purification

To make a Crp overexpression construct, the *crp* gene was PCR amplified from the cosmid StH17 and was digested with *Bam*HI and *Nde*I. PCR was performed using the iProof polymerase system (Bio-Rad). The fragment was ligated into the pET15b vector (Novagen) digested with the same enzymes. Construct pET15b*crp* was transformed into *E. coli* BL21 (DE3) (Novagen) (Table 2.1). The plasmid-containing strain was subcultured to exponential growth phase (OD_{600}= ~0.5) before IPTG was added to a final concentration of 0.5 mM to induce overexpression. Induction was performed at 26°C overnight, after which cells were collected by centrifugation at 1,000 × g at 4°C for 5

min. For every 200 ml of culture, the pelleted cells were resuspended in 10 ml lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole pH 8.0) containing 1× protease inhibitor (Complete Mini, Protease Inhibitor Cocktail, Roche) and were then disrupted by four passages through a French Press cell (Thermo Fisher Scientific) at 10,000 psi. Cell debris was removed by twice centrifuging at 12,000 × g at 4°C for 30 min. The resulting supernatant was transferred to a 15 ml Falcon tube and incubated with Ni-nitrilotriacetic acid (Ni-NTA) agarose resin (Qiagen) for 1 h at 4°C, allowing the histidine (His_6)-tagged Crp to associate with the resin. The slurry was loaded onto a PolyPrep chromatography column (Bio-Rad). Flow-through was discarded and the column was washed with five column volumes of wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole pH 8.0). Proteins were eluted using increasing concentrations of imidazole (100 mM to 500 mM). All eluted fractions were examined using SDS-PAGE and Coomassie blue staining to check for the presence of Crp. The fraction containing the majority of Crp was subjected to buffer exchange with phosphate buffered saline (PBS) (pH 7.4) using Macrosep Advance Centrifugal Devices (PALL) with a cut-off value of 10 kDa. The protein was stored at -20°C in the presence of 50% glycerol and the concentration was determined using a Bradford assay (Bradford, 1976).

2.2.10 Antibody generation and purification

Purified His_6-tagged Crp was used to generate polyclonal antibodies (Cederlane Labs; Burlington, ON). To remove His_6-tag-reactive species from the crude antiserum, an unrelated His_6-tagged protein (VirB8 protein from *Brucella suis*) was overexpressed and immobilized on the PolyPrep column loaded with Ni-NTA agarose as described in section 2.2.9. Following five washes (one column volume each), anti-Crp antiserum was passed through the column and the flow-through was collected as the "pre-cleared antiserum". At the same time, His_6-Crp was bound to the Ni-NTA agarose and the beads were then washed with five column volumes of wash buffer and five column volumes of equilibration buffer (150 mM NaCl, 50 mM Tris-Cl, pH 7.4), before being incubated with the pre-cleared antiserum overnight at 4°C. The mixture was loaded onto a PolyPrep column and the flow-through was discarded. The resin was washed with 5 ml equilibration buffer and then 5 ml equilibration buffer containing 2 M NaCl. The column bottom was sealed with the stopper provided and 2 ml of 4 M $MgCl_2$ was added before incubation at room temperature for 10 min. The column was transferred back to 4°C and the stopper was removed. The liquid eluted from the column was collected in four fractions of 0.5 ml each. Each fraction was subjected to buffer exchange separately in dH_2O and then in PBS as described above. SDS-PAGE and Coomassie blue staining were performed to detect the antibody. The fractions containing the antibody were pooled and stored at -20°C in 50% glycerol.

2.2.11 Chromatin immunoprecipitation–microarray (ChIP-chip) assay

The wild-type strain M145 was grown in liquid R5 medium for 16 h before adding formaldehyde to a final concentration of 1% (v/v). The *crp* mutant strain was grown for 2

days and 16 h to get an equivalent amount of biomass. Cultures were grown in the presence of formaldehyde at 30°C for 25 min before glycine was added to a concentration of 125 mM to terminate the cross-linking reaction. The mycelia were washed twice with PBS (pH 7.4), resuspended in lysis buffer (10 mM Tris-HCl pH 8.0, 50 mM NaCl, 10 mg/ml lysozyme, 1 × protease inhibitor) and incubated at 25°C for 25 min, after which an equal volume of immunoprecipitation (IP) buffer (100 mM Tris-HCl pH 8.0, 250 mM NaCl, 0.5% Triton X-100, 0.1% SDS, 1 × protease inhibitor) was added. The mixture was sonicated at 50% power for 5 cycles of 15 s each (Branson, Sonifier Cell Disruptor 350) to shear the chromosomal DNA into fragments of 0.3-1 kb, centered around 0.5 kb. The sample was centrifuged twice at 16,000 × g at 4°C for 20 min to remove cell debris. Ten microlitres of the cell extract were saved as the "total DNA" control. The remainder of the sample was pre-cleared by incubating with 1/10 volume of 50% protein A-Sepharose slurry for 1 h at 4°C with mild agitation. The beads were subsequently removed by centrifugation at 16,000 × g at 4°C for 15 min. The supernatant was incubated with 1/10 volume purified anti-Crp antibody overnight at 4°C with mild shaking. One tenth volume of 50% protein A-Sepharose slurry was then added to pull down the anti-Crp-Crp-DNA complexes and the sample was incubated for 4 h at 4°C with mild shaking. The beads were then pelleted by centrifugation at 1,000 × g at 4°C for 5 min and washed twice with 0.5 × IP buffer and then twice with 1 × IP buffer. The pellet and 10 µl of the total cell extract were both incubated in 150 µl IP elution buffer (50 mM Tris-HCl pH 7.6, 10 mM EDTA, 1% SDS) overnight at 65°C to reverse the cross-links. The sample was centrifuged at 16,000 × g for 5 min to remove the beads and the pellet was re-extracted with 50 µl IP elution buffer. The combined eluted DNA was treated with 0.1 mg/ml Proteinase K (Roche) at 55°C for 2 h. The samples were subjected to phenol-chloroform extraction and precipitated with 1/10 volume of 3 M NaOAc (pH 5.2) and 2.5 volumes of 95% ethanol. Immunoprecipitated DNA was resuspended in 10 µl dH_2O and total DNA was resuspended in 50 µl dH_2O. The DNA samples were quantified using a NanoDrop 1000 spectrophotometer (Thermo Scientific).

Total and immunoprecipitated DNA were labelled with Cy5 and Cy3 (fluorescent cyanine dyes) respectively, and hybridized to a microarray consisting of 44,000 60-mer oligonucleotide probes covering the entire genome of *S. coelicolor* (Oxford Gene Technology, Oxford UK). Duplicate, independent DNA samples for each strain were examined. DNA labeling, hybridization and microarray scanning were performed by Oxford Gene Technology (OGT) according to their standard protocols (http://www.ogt.co.uk). For each array, the signals of all probes were normalized to the median channel signal for the respective array to correct for any systematic errors. Signal ratios between the immunoprecipitated DNA and the total reference DNA were obtained for both the wild-type and the mutant experiments. A final interaction score was calculated by taking the $\log_2$ value of the ratio between the wild-type and the mutant values (MUT/WT) for each probe. A probe was defined as a Crp association site when its interaction score is three times the standard deviation above the median interaction score (0.104 + 0.45 × 3) and at the same time, at least one adjacent probe shows a score higher

than the cut-off value; however the scores of the adjacent probes have to be lower than that of the peak probe.

A map of the *S. coelicolor* genome with data from the ChIP-chip and the microarray assays was generated using *CGView* (Stothard & Wishart, 2005). Hits from both datasets were positioned using probe start and stop positions, and scored using the $\log_2$(MUT/WT) and fold change values for the ChIP-chip and microarray assays respectively. Secondary metabolic operons were added and positioned accordingly (Bentley *et al*, 2002).

2.2.12 RNA isolation

Cultures of *crp* (pIJ6902*crp*) (spore stock harvested in the presence of the inducer thiostrepton in order to avoid suppressor mutation that always appeared when un-induced stock was used) and negative control *crp* (pIJ6902) were grown in liquid R5 medium for 16 h and 2 days plus 16 h, respectively, before addition of thiostrepton to induce *crp* expression. RNA was harvested from cell aliquots before induction, and at 15, 30, 45, and 60 min after thiostrepton induction. Total RNA was harvested as described previously (Hopwood *et al*, 1985) followed by clean-up using the RNeasy mini kit (Qiagen). All RNA samples were quantified using UV spectroscopy on an Ultrospec 3100 pro (Biochrom). RNA quality was examined using agarose gel electrophoresis.

2.2.13 RT and qPCR

Reverse transcription (RT) reactions were performed using SuperScript III reverse transcriptase (Invitrogen). Two picomoles of gene-specific primer, 2 µg of total RNA, and 2 µl of 10 mM dNTPs were mixed together, incubated at 95°C for 5 min and cooled on ice. Four microlitres of 5 × First strand buffer, 1 µl 0.1 M DTT, and 40 U RNaseOUT (Invitrogen) were added to each reaction, followed by incubation at 42°C for 2 min. Two hundred units of SuperScript reverse transcriptase were then added and each mixture was incubated at 42°C for 50 min. Following this, the enzyme was inactivated at 70°C for 10 min. One microlitre of cDNA was used for each 25 µl real time PCR (qPCR) reaction in the presence of 1× PCR buffer, 2 mM $MgSO_4$, 0.2 mM dNTP, 1 mM of each gene specific primer, 7.5% DMSO, 0.5 µl SYBR Green I dye (50 × in DMSO) (Invitrogen) and 1.25 U Taq DNA polymerase (Norgen). The reactions were conducted in a CFX96 qPCR detection system (Bio-Rad)**.** The cycling conditions used were: 95°C for 5 min; 95°C for 30 s; 58 or 60°C for 1 min (annealing); 72°C for 30 s (extension); 72°C for 10 min. Steps 2 to 4 were repeated for a total of 40 cycles. All reactions were performed in biological and experimental triplicate. For each pair of primers, the annealing temperature was optimized to obtain a specific product. Standard curves for each primer pair was constructed by amplifying the same fragment using 10-fold dilutions of chromosomal DNA as template. Amplification efficiency obtained from the standard curve for each set of primers fell in the range of 95%-105%. Therefore, the comparative Ct (threshold cycle)

method was employed to calculate the fold-change of the target DNA levels in different samples (Schmittgen & Livak, 2008)

2.2.14 Transcriptome profiling

RNA samples were prepared as described above in duplicate and were sent to the London Regional Genomics Center for microarray analysis. cDNA samples were created by reverse transcription, before being biotinylated, fragmented and hybridized to a custom-designed Affymetrix GeneChip array according to Affymetrix published protocols. The array contained 226,576 perfect match oligonucleotide probes (25 bp), with 8,205 probe sets (each set contains multiple probes annealing to the same coding sequence) targeting protein-coding regions, 10,834 probes sets (each set contains multiple probes annealing to the same intergenic region) tiling the intergenic regions, and 3,672 probe sets (each set contains multiple probes annealing to the same non-coding RNA region) targeting non-coding RNAs in the *S coelicolor* genome (Nieselt *et al*, 2010). The hybridized GeneChips were stained and washed using an Affymetrix Fluidics Station 450 and scanned with an Affymetrix® Scanner 3000 7G. The $\log_2$ values of the signals were normalized to the median value of the respective arrays. Transcriptional fold-change of each gene was calculated as the ratio between the induced and the uninduced sample. ANOVA tests were performed to evaluate the statistical significance of the results. Selected targets were validated with qPCR. In addition to the genes of interest, 16S rRNA was included as a reference for qPCR. For each time point, the Ct of a target gene was normalized to the Ct of 16S rRNA, which was obtained using the same cDNA. The uninduced (time 0) sample was used to determine the baseline expression level, and to determine the fold-change in transcript levels at each subsequent point in the time course. ANOVA tests were performed using SPSS v17.0 to determine whether the results were statistically significant.

2.2.15 'Crp induction'-ChIP assay

Induction of *crp* (pIJ6902*crp*) (using spore stocks harvested in the presence of thiostrepton) was conducted as described for the microarray assays. Fifty milliliters of culture were taken from a starting culture of 250 ml at 0 (uninduced), 15 and 45 min, and the cells were washed twice with equal volumes of fresh R5 medium. The mycelia were resuspended in 50 ml R5 medium and subjected to the ChIP procedure described in 2.2.11. Both the immunoprecipitated DNA and the total DNA were used as templates to amplify select target sequences. For each time point, the immunoprecipitated DNA Ct value was normalized to the total DNA Ct value. The uninduced sample was used to assess the fold-change of the DNA levels in the 15 and 45 min samples. Three independent cultures were set up for the ChIP experiments and each qPCR was done in triplicate for each biological replicate. ANOVA tests were performed using SPSS v17.0 to test the statistical significance of the results.

2.2.16 Antibiotic production assay

Antibiotic production was quantified as described before (Kang *et al*, 1998; McKenzie & Nodwell, 2007). M145 (pIJ82), *crp* (pIJ82), *crp* (pIJ82*crp*), and M145 (pIJ82*ermE*crp*) were grown in liquid R5 medium for 7 days. To determine the level of total Act, 0.5 ml of culture was taken from each culture every 24 h and mixed with an equal volume of 2 M KOH. The mixture was vortexed and spun at 16,000 × g for 5 min. The absorbance of the supernatant was measured at 633 nm using an Ultrospec 3100 pro (Biochrom). For analysis of Red levels, 1 ml of culture was pelleted, the culture supernatant was removed, and the cells were resuspended in an equal volume of acidified methanol (0.5 M HCl). The cells were extracted overnight, after which cell debris was removed by centrifugation (16,000 × g, 5 min) and the absorbance of the supernatant at 530 nm was measured. Three independent cultures were set up for each strain, and two aliquots from each culture were tested. To assess CDA production, spores of all strains were grown in liquid R5 for 48 h. Cells were removed by centrifugation at 16,000 × g for 5 min and the supernatant was filtered through 0.2 µm pore size Acrodisc syringe filters (PALL). ONA (Oxoid nutrient agar) plates with or without 12 mM $CaCl_2$ were overlaid with ONA soft agar (1:1 ONA plus Oxoid nutrient broth) containing 1/100 volume of overnight culture of *Staphylococcus aureus*. Sterilized 6 mm antibiotic assay disks (Whatman) were placed on top of the soft agar. Fifteen microliters of supernatant were loaded on each disk and the plates were incubated overnight at 37°C. Four replicates were conducted for each strain. CDA production was quantified by measuring the diameters of the inhibition zones. The levels of all antibiotics were normalized relative to the biomass of the mycelia from which the antibiotics were extracted.

2.2.17 Antibiotic bioassay

S. coelicolor A3(2) M145, *S. venezuelae* ATCC 10712, *S. pristinaespiralis* ATCC25486, *S. sp.* SPB74, *S. sp.* Ja2b, and *S. sp.* WAT1 (Table 2.1) were tested against the following indicator strains: Gram-negative bacteria *E. coli, Pseudomonas aeruginosa, Klebsiella pneumoniae*; Gram-positive bacteria *S. aureus,* methicillin-resistant *Staphylococcus aureus* (MRSA), *Bacillus subtilis,* vancomycin-resistant *Enterococcus* (VRE) and *Micrococcus luteus* (Table 2.1). Approximately 10^6 spores (in 5 µl of sterile dH_2O) of the *Streptomyces* strains were spotted on DNA agar plates and incubated at 30°C for 48, 72, 96, 120 and 144 h before being overlaid with soft agar (1:1 DNA plus Difco nutrient broth) containing a 100-fold dilution of overnight culture of indicator strains in liquid LB. The plates were incubated overnight at 37°C before measuring the size of the inhibition zone (the median distance from the outer edge of each *Streptomyces* circular patch to the edge of the zone of clearing). This is different from the method used in CDA assay since the shape of inhibition zone in this experiment is not a perfect circle in most cases. Each experiment included four replicates for each strain and was performed three times.

2.2.18 Secondary metabolite extraction

S. coelicolor A3(2) M145, *S. venezuelae* ATCC 10712, *S. pristinaespiralis* ATCC25486, *S. sp.* SPB74, *S. sp.* Ja2b, and *S. sp.* WAT1 were engineered to carry either pIJ82 or pIJ82*ermE*crp* (overexpression construct). Equivalent amounts of spores of both strains (containing either plasmid alone or the *crp* overexpression construct) were spread on R5 agar medium and incubated for 3 and 7 days. The cultures, along with the agar, were cut into small pieces, soaked in 25 ml n-butanol and sonicated in a Branson 2520 Tabletop Ultrasonic Cleaner for 3 min before being macerated in butanol at room temperature overnight. The mixture was filtered through Whatman filter paper and dried down in a HT-4X centrifugal vacuum evaporator (Genevac), followed by reconstitution in 500 µl of acetonitrile:dH_2O (1:1, HPLC grade). R5 medium alone was processed in parallel as a negative control. Each sample was prepared in quadruplicate.

2.2.19 High performance liquid chromatography and mass spectrometry (HPLC-MS)

LC-MS analysis was performed using a Agilent 1200 series analytical HPLC system equipped with a reverse phase C18 column (2.1×100 mm, 2.6 µm, 100 Å) (Kinetex) coupled to a benchtop time-of-flight spectrometer Bruker MicroTOF II (Bruker Daltonics). The samples were separated at 50^{o}C using a gradient of 5% - 95% acetonitrile (0.1% formic acid, v/v) over 22 min. The flow rate was 0.2 ml/min. Positive electrospray ionization was performed at a voltage of 4.5 kV and the ions were scanned over the mass range of 200-1,700 m/z. Data were analysed using MZmine 2 software (Pluskal *et al*, 2010).

2.2.20 cAMP concentration measurement

Spores of the wild-type, *crp* mutant, *crp* (pIJ6902*crp*) and *crp* (pIJ6902) strains were pregerminated and cultured in liquid R5 medium. For the wild-type strain, samples were harvested at 12, 16 and 20 h, while the *crp* (pIJ6902*crp*) cultures were induced at 16 h and samples were harvested at 16 (pre-induction), 18 and 20 h. The cultures of the negative controls *crp* and *crp* (pIJ6902) were set up 48 hours ahead of the wild-type and *crp* (pIJ6902*crp*) respectively and then followed the same time course in each case. At each time point, 7 ml of culture was removed and the cells were pelleted at 8,000 × g for 5 min at 4°C. For determining extracellular cAMP levels, the supernatant was heated at 95°C for 5 min and then diluted 2- (for M145) or 50- (for all other strains) fold in work buffer (BTI, Biomedical Technologies). Samples were assayed using a cAMP EIA (enzyme immunoassay) kit (BTI) that allows cAMP quantification in the range of 0.5–100 pmol/ml. For quantifying intracellular cAMP, the cell pellets were washed in an equal volume of PBS (0.8% NaCl; 0.02% KCl; 0.15% Na_2HPO_4; 0.024% KH_2PO_4; pH 7.4) and resuspended in 1 ml work buffer (BTI). The mycelia were sonicated on ice at 50% duty cycle for 4 × 15 s (Branson Sonifier Cell Disruptor 350), followed by centrifugation at 16,000 × g for 30 min. The cell extract supernatant was heated at 95°C for 5 min before being assayed using the kit. Each strain was examined in biological

triplicate and the concentrations were normalized relative to the biomass of the mycelial pellets.

2.3 Results

2.3.1 Characterization of the *crp* mutant phenotype

Under laboratory conditions, wild-type strain M145 usually initiates aerial growth between 36 and 48 h, and this is followed shortly thereafter by sporulation. However, on both MS (poor carbon source) and R2YE (rich) agar media, the *crp* deletion strain displayed a growth delay of more than 24 h. Furthermore, in contrast to fully grown wild-type colonies having diameters of around 3.5 mm, the mutant colonies had diameters of only ~1 mm and they were more grey-pigmented at the late-sporulation stage (Fig. 2.1). These defects could be restored by complementing the strain with pIJ82 (an integrating plasmid vector) carrying the *crp* gene along with its native promoter (Fig. 2.1). Stress tests including lysozyme, SDS, and heat shock treatments were conducted to evaluate mutant spore viability, as spore defects have been associated with increased sensitivity to heat and chemical stress (Margot *et al*, 1998; Mazza *et al*, 2006). Although no change in resistance to SDS and lysozyme was detected, the survival rate following exposure to 60 ^{o}C for 90 min was increased 4-fold for the mutant, indicating the mutant spores were more resistant to heat treatment (Fig. 2.2C). This was further supported by TEM micrographs showing that the spore wall thickness was doubled in the *crp* strain relative to that of the wild-type (Fig. 2.2A, B). A close-up examination of the colony surface revealed that after 6 days' growth, all of the spores of the wild-type strain were organized into long filaments, while in the *crp* mutant, the majority of the spores were present as free spores and they lacked the organized structure of the wild-type spore chains (Fig. 2.4). Taken together, the initial strain characterization suggested an important role for Crp during vegetative growth and spore formation of *S. coelicolor*.

2.3.2 Dissecting the role of Crp at different developmental stages

With the goal of clarifying the precise role of Crp, we set out to dissect its function during spore formation and germination. *crp* was cloned downstream of the thiostrepton-inducible promoter *tipA* in the integrating vector pIJ6902 and was introduced into the *crp* mutant strain. Spores of the complemented strain were harvested either in the presence or absence of thiostrepton, and each spore stock was subsequently cultured on media either containing or lacking the thiostrepton inducer. This gave rise to four different scenarios: (i) a strain where Crp is continually induced, which should mimic the wild-type strain; (ii) a strain completely lacking Crp, which should resemble the *crp* mutant; (iii) a strain where Crp is expressed during sporulation but not during subsequent re-growth; and (iv) a strain where Crp is expressed exclusively during re-growth. The latter two cases allow us to distinguish whether the defects observed in the phenotypic assays stem from defective spore formation or from an important role for Crp during spore germination and out-growth.

As expected, on both MS and R2YE media, the uninduced strain (designated '-/-') had a morphology similar to that of the deletion mutant, and when *crp* was induced at both stages (denoted '+/+'), the defects were fully complemented, suggesting that our induction system was effective (Fig. 2.3A). When *crp* was induced exclusively during spore harvest but not during re-growth (termed '+/-'), the colonies resembled those of the deletion mutant, while when *crp* was induced exclusively during regrowth (dubbed '-/+'), wild-type colony morphology was observed (Fig. 2.3A). Given that previous work has shown *crp* to be essential for germination, the wild-type, deletion mutant, and *tipA-crp* containing strains were examined using germination assays. Germ tube development was followed by incubating spores on MS agar at 30°C for up to 12 h. The wild-type strain exhibited little germination before 6 h, but germination progressed rapidly after that and by 12 h, ~80% of the spores had germinated (Fig. 2.3B). In accordance with the morphological studies, germination of the *crp* strain was delayed by at least 6 h, with fewer than 5% of the spores having germinated by 12 h. As expected, no and full Crp induction (-/- and +/+) showed similar profiles to the deletion mutant and wild-type strains, respectively. Surprisingly, however, exposure to thiostrepton during spore harvest, but not afterwards (+/-), resulted in wild-type levels of germination, while Crp induction exclusively during subsequent replating (-/+) only partially restored germination, despite the fact that the colony morphology of the previous resembled that of the mutant strain, and the latter strain looked wild-type when grown on agar plates (Fig. 2.3A, B). Scrutinizing these colonies using SEM revealed that the absence of Crp during regrowth (-/- and +/- strains) led to greater numbers of free spores, whereas if Crp was induced during replating (-/+ and +/+ strains), wild-type-like spore chain integrity was restored to both strains. Moreover, it appeared that continual induction of Crp (in the +/+ strain) resulted in less septated filaments than observed in the wild-type strain (Fig. 2.4).

These results altogether suggested that the lack of Crp in the previous life cycle (during sporulation/spore harvest) resulted in spores defects, making them less amenable to germination. Conversely, although the spores harvested in the presence of Crp are capable of initiating normal germination, they can not sustain wild-type level of post-germination vegetative growth, as manifested by the drastic difference in colony appearance on agar plates.

2.3.3 Crp negatively regulates cAMP synthesis

Since cAMP is the effector molecule for the Crp protein (at least in *E. coli*), we were interested in monitoring its concentration in *S. coelicolor*. This investigation would also provide guidance for us in selecting conditions under which to conduct the ChIP-chip and transcriptome profiling experiments described later. Liquid cultures of wild-type and *crp* mutant strains were checked at 12, 16 and 20 h, whereas for the *crp* mutant containing either the *tipA-crp* construct or the empty vector, the first sample was taken at 16 h, after which thiostrepton was added and cAMP levels were examined at 18 and 20 h (cultures

of *crp* mutant and *crp* mutant containing the empty vector were set up 48 hs ahead of time and subsequently subjected to the same time course in order to get enough biomass). First of all, the results confirmed previous reports that cAMP is present at very low levels (< 10 pmol/mg) in wild-type *Streptomyces* cultures and most of it is found in the culture medium (Ruiz *et al*, 2010; Ullmann & Danchin, 1983). Over 20 h, cAMP concentrations in the wild-type strain were maintained at very low levels; however, they were around 20-fold higher in the *crp* mutant strain throughout the time course (Fig. 2.5A). Before addition of thiostrepton to the cultures of *crp* (pIJ6902*crp*) and *crp* (pIJ6902), the cAMP levels in both stains were approximately 150 pmol/mg. In the control carrying the empty vector, the level decreased slightly over the time course, while in the Crp induced samples, the decrease was much more prominent, with the cAMP concentration dropping to only ~50 pmol/mg (Fig. 2.5B). This drastic change upon Crp expression, along with the difference between the wild-type and the *crp* mutant strains, suggest negative feedback regulation of cAMP production by Crp. This phenomenon prompted us to test whether the phenotype of *crp* mutant stemmed from the lack of *crp* gene itself or from the immense elevation in cAMP levels. Upon cultivation on both MS and R2YE agar supplemented with 2 mM cAMP, the wild-type strain did not exhibit any morphological change compared to those grown on regular medium, confirming the importance of Crp *per se* as a regulator during development.

2.3.4 Deletion of *crp* affects antibiotic production in *S. coelicolor*

In addition to aberrant colony and spore morphology, we also observed a pronounced reduction in levels of the blue-pigmented antibiotic Act during growth of the *crp* mutant on both MS and R2YE agar (Fig. 2.3A). This has been reported previously, but has never been investigated in detail (Derouaux *et al*, 2004b; Piette *et al*, 2005). In order to quantitatively assess the levels of antibiotics produced by the wild-type and the *crp* mutant strains, total Act (actinorhodin and γ-actinorhodin) was extracted by treating the culture with KOH, while Red was extracted by treating the mycelia with acidified methanol. Relative absorbances were measured at 633 (Act) and 530 (Red) nm, followed by normalization relative to the biomass of the cell pellets (Kang *et al*, 1998). In R5 liquid medium, Act levels in the wild-type strain increased rapidly between day 2 and 4 and then decreased slightly on day 5 and day 6, before increasing again on day 7. By contrast, Act production was drastically diminished in the *crp* mutant at all times (Fig. 2.6A). Red synthesis profiles, however, were very similar for both strains, where a sharp increase was observed early during the growth, followed by a significant drop, and a minor peak again later in the time course. There was, however, a reproducible delay of ~ 24 h in Red production by the *crp* mutant (Fig. 2.6B). We also assessed CDA production using a bioassay against the indicator strain of *Staphylococcus aureus*. The sizes of the zones of inhibition were measured and normalized to mycelia biomass (McKenzie & Nodwell, 2007). We found loss of *crp* completely inhibited CDA production compared to the wild-type strain (Fig. 2.6C). The synthesis of all three antibiotics could be restored by complementing the deletion mutant with a construct carrying *crp* expressed from its native promoter, verifying that the phenotypes were due to the deletion of *crp* (Fig. 2.6).

Taken together, our data suggest that Crp plays an important role in secondary metabolite production.

2.3.5 Identification of Crp association sites in the genome of *S. coelicolor*

To gain further insight into the defects in cell growth and antibiotic production associated with a *crp* mutation, and to determine the regulatory targets of Crp in *S. coelicolor*, we pursued a ChIP-chip approach for Crp target identification. Unlike *in vitro* gel shift mobility assays, which had previously failed to produce any positive results (Derouaux *et al*, 2004a; Derouaux *et al*, 2004b), ChIP-chip assays have the advantage of being able to capture transient and weak interactions *in vivo* (Buck & Lieb, 2004). Liquid cultures were used due to technical limitations (the formaldehyde cross-linking does not work well for solid-grown cultures) and since *S. coelicolor* does not differentiate in liquid medium, the experiment would mainly provide information on the targets of Crp during vegetative growth and secondary metabolism. Any potential consensus binding motif derived will serve as the basis for further genome mining to discover additional Crp target genes involved in aerial hyphae formation and sporulation.

Initially, Crp expression was tested at both transcriptional and translational levels throughout a 48 h time course. Reverse transcription-PCR showed the Crp transcript to be most abundant at 12 h, and decreased thereafter (Fig. 2.7A), whereas the Crp protein was present at a relatively constant level throughout growth (Fig. 2.7B). The cellular abundance of the effector molecule cAMP also gradually dropped over a similar time course (Fig. 2.5A). In order to get enough biomass and at the same time ensure abundant cellular levels of Crp and cAMP, 16 h was picked as the time point to start cross-linking and harvest samples. Crp was crosslinked to target DNA *in vivo* and the DNA was fragmented to the size of 0.3-1 kb, centered around 0.5 kb. Crp-DNA complexes were immunoprecipitated using an affinity purified Crp-specific polyclonal antibody. The DNA samples were eluted, fluorescently labelled, and hybridized with microarrays containing oligonucleotide probes distributed across the entire *S. coelicolor* genome (on both strands). A probe is considered to represent a binding (or association) site only when the wild-type / mutant signal ratios of itself and at least one adjacent probe were three times the standard deviation above the median ratio of the entire array. At the same time, the score of this probe has to be higher than its two adjacent probes, making it a local peak. A total of 393 binding sites were obtained and mapped to the genome (Fig. 2.8). Crp binding sites were distributed fairly evenly throughout the genome, with 234 sites found within coding regions, 151 located in intergenic regions (IGRs), and eight spanning both coding and non-coding regions.

Each Crp binding/association site (indicated as Crp sites hereafter) was examined manually in relation to its chromosome context. If a site was identified within an intergenic region, adjacent, divergently-encoded genes were considered to be potentially regulated by Crp, although any flanking genes for which the start codon was more than 500 bp away from the center of the Crp sites were excluded from consideration. Five

hundred base pairs was selected as a cut-off value on since it is the longest distance between a Crp ortholog binding site and the start codon of the regulated gene in the published literature (Rickman *et al*, 2005). For Crp sites within a coding region, the associated gene, along with divergent flanking genes (within 500 bp), were extracted as potentially regulated candidates. Each candidate target gene was classified according to its function, as published in the literature or as annotated in the *Streptomyces* database StrepDB (Table 2.3). Among the genes with assigned functions, the most abundant functional groups were transcription regulators (9.9%) and proteins involved in metabolism (17.6%), one third of which participate in secondary metabolism (5.1%). This is in accordance with the canonical role of Crp in controlling metabolism and numerous downstream transcription regulators (Fic *et al*, 2009). Compared to the overall proportion (2.4%) of the genome dedicated to secondary metabolism, genes in this category are obviously overrepresented in the list of Crp associated targets. Eight out of 22 characterized or predicted secondary metabolic clusters were associated with Crp sites (Fig. 2.8, Table 2.4) (Bentley *et al*, 2002). *SCO5071-5092*, *SCO5877-5898*, *SCO3210-3249*, and *SCO6273-6288* encode the biosynthetic pathways of Act, Red, CDA, and yCPK respectively. Within these clusters, Crp sites were associated with biosynthetic genes, and Crp was also found to co-immunopreciptate with the coding regions of the pathway specific regulatory genes of *act* (*SCO5085*, *actII-ORF4*), *red* (*SCO5881*; *redZ*), and *cda* (*SCO3217*; *cdaR*) clusters. The other four Crp-associated metabolic clusters are predicted to code for nonribosomal peptide synthetases (*SCO6429-6438*), a sesquiterpene cyclase (*SCO5222-5223*), a type II fatty acid synthase (*SCO1265-1273*), and deoxysugar synthases/glycosyl transferases (*SCO0381-0401*). Each of these four uncharacterized clusters contained Crp sites in either intergenic or coding regions.

Additionally, substantial fractions of the regulated genes belong to the categories of DNA replication, recombination and repair (7.9%), and morphological differentiation and sporulation (5.1%), despite the fact that experiments were conducted using liquid cultures. Similar results were seen before when liquid cultures were subjected to ChIP-chip assays to identify BldD binding sites (den Hengst *et al*, 2010). Important developmental genes associated with Crp sites include (Table 2.4): *bldG*, which encodes an anti-anti-sigma factor involved in aerial development and stress responses in *S. coelicolor* (Parashar *et al*, 2009; Sevcikova *et al*, 2010), *whiB*, which encodes a small protein with a [4Fe-4S] cluster that may respond to redox changes during sporulation (Davis & Chater, 1992; Jakimowicz *et al*, 2005), and *ssgB*, which encodes a protein that directs FtsZ localization during aerial hyphae septation (Willemse *et al*, 2011). Crp sites also include sequences associated with several chaplin genes (*SCO1674*, *SCO1675*, and *SCO2717*), where the chaplins are secreted hydrophobic proteins that polymerize on the surface of *Streptomyces* cells so as to facilitate aerial hyphae growth (Claessen *et al*, 2003; Elliot *et al*, 2003)

As mentioned above, a previous proteomic study had identified a number of candidate Crp targets, based on the fact that their protein products were far less abundant in a *crp* mutant than in a wild-type strain during the first 5 h of germination (Piette *et al*, 2005).

Crp was found to associate with the upstream sequences of several of these genes, including *SCO3878* and *SCO4296*. *SCO3878* encodes the β chain of DNA polymerase III (DnaN) while *SCO4296* encodes the protein chaperone GroEL. DNA polymerase III is a primary enzyme complex for prokaryotic DNA replication and GroEL is an important chaperone protein. Both of these genes were downregulated in the *crp* deletion strain during the initial growth stage according to 2D gel electrophoresis analysis, indicating a possible role for Crp in the reactivation of DNA replication machinery and protein folding (Piette *et al*, 2005).

Selected Crp sites sequences were subjected to gel mobility shift assays to validate their interactions *in vitro*. However, reproducible shifts were not detected even after varying the buffer composition, pH, and the Crp samples used (either purified protein or crude cell extracts). Adding cAMP in the binding reaction, the gel and the running buffer had no effect on binding. This adds further weight to the notion that *S. coelicolor* Crp may behave differently from the *E. coli* Crp.

2.3.6 Validation of ChIP-chip association sites using 'Crp induction'-ChIP assays

Since all efforts to demonstrate Crp-DNA interactions *in vitro* failed, a 'Crp induction-ChIP' assay was adopted as a complementary approach. Cultures of the *crp* (pIJ6902*crp*) strain were induced with thiostrepton for 0 (uninduced), 15, and 45 min respectively before crosslinking and chromatin immunoprecipitation as described. Instead of being loaded onto microarrays, the immunoprecipitated and total DNA samples were used as templates for PCR amplification of fragments harboring selected Crp target sites discovered in ChIP-chip experiments. qPCR was employed to quantitatively assess the levels of the Crp targets in the immunoprecipitated DNA samples. If Crp associates with the target DNA, we anticipated enrichment of the targets upon Crp induction following normalization with respect to total DNA signals. The sequences upstream of or within the secondary metabolic pathway-specific regulatory genes *SCO5085* (*actII-ORF4*), *SCO3217* (*cdaR*), and *SCO5881* (*redZ*) were selected for investigation. Primers annealing to a non-ChIP-chip target gene *SCO4662* (*tuf-1* gene) were used as a negative control. All fragments except for the negative control were enriched in the immunoprecipitated DNA later in the time course, with maximal levels observed at 45 min (Fig. 2.9). This experiment indirectly confirmed *SCO5085* (*actII-ORF4*), *SCO3217* (*cdaR*), and *SCO5881* (*redZ*) as Crp targets, although without *in vitro* evidence, direct regulation can not be definitively concluded.

2.3.7 Transcriptome profiling

In addition to ChIP-chip assays, transcriptome profiling was also performed to get more in-depth information about Crp regulation. RNA samples were prepared from strains carrying the inducible *tipA-crp* construct and the empty vector respectively before and after thiostrepton induction. In this experiment and the 'Crp induction-ChIP' assays, the spore stock of *crp* (pIJ6902*crp)* harvested in the presence of thiostrepton was used in

order to circumvent the suppressor mutation problem that always arose when the spore stock harvested in the absence of thiostrepton was used. To address the concern that the presence of Crp in the spores may compromise the induction efficiency, cellular levels of Crp throughout the time course were examined. Immunoblots showed that although Crp was detected at a low level before induction, its cellular abundance was dramatically increased following induction (Fig. 2.10), justifying our choice of the spore stock. The RNA samples were assayed using Affymetrix-based microarrays in order to detect changes in the *S. coelicolor* transcriptome upon Crp induction. Genes showing at least 2-fold changes in their transcription levels after induction in the *tipA-crp* samples, but not in the negative control, were considered for follow-up investigation. Overall, 360 genes were stimulated and 91 genes were repressed (Fig. 2.8).

Consistent with the ChIP-chip assay results, functional classification of the Crp-affected genes in our microarrays supported a versatile role for Crp with an emphasis on metabolism (48.6%), particularly secondary metabolism (17.5%). Crp was also found to affect the expression of diverse transcription regulators (9.5%), which would further account for the pleiotropic effects associated with its loss (Table 2.5). The microarray results underlined an even stronger correlation between Crp and secondary metabolism than was observed with the ChIP-chip experiments, with secondary metabolic genes being significantly overrepresented (17.5%) as Crp targets compared to their overall genomic percentage of 2.4%. A number of affected genes known or predicted to be involved in the same cellular activity were clustered together, and all clustered genes were affected in the same way (all activated or downregulated). Among these, the most striking were four antibiotic biosynthetic gene clusters: *act* (*SCO5072-5092*), *red* (*SCO5877-5898*), *cda* (*SCO3210-3249*), and *cpk* (*SCO6266-6288*) (Fig. 2.11). The expression of the pathway-specific activator genes *actII-ORF4*, *redD*, *cdaR*, and *cpkN* (an uncharacterized pathway-specific activator of the *cpk* pathway) were all upregulated following Crp induction, and they all contained Crp-associated sequences except for *cpkN*. Other activated genes included: *SCO5029* and *rpfA* (*SCO3097*), two cell wall hydrolase genes implicated in cell wall remodelling during morphological differentiation (Haiser *et al*, 2009); *hspR*, a regulator controlling the *dnaK* operon in response to heat shock stress (Kallifidas *et al*, 2010); *bldK* and *ramCS*, genes essential for aerial hyphae erection (Kodani *et al*, 2004; Nodwell *et al*, 1996); and *mreC* and *ssgD*, which encode proteins important for cell division and sporulation (Burger *et al*, 2000; Noens *et al*, 2005).

Unlike the stimulated genes, the majority of the repressed genes encode proteins of unknown function. When examined in their genomic context, there were two extensive regions of downregulation, *SCO0162-0181* and *SCO0197-0220*. Despite the lack of validated annotation for these genes, it was noticeable that six out of twelve *S. coelicolor* "universal stress protein"-encoding genes are distributed in these regions (Kvint *et al*, 2003). This family of proteins is common in diverse organisms. Although their precise role(s) in *S. coelicolor* has not been determined, in many other bacteria, they are essential

for coping with wide-ranging stresses such as nutrient depletion, oxygen depletion, heat shock, and antibiotic treatment (O'Toole & Williams, 2003).

When examined side by side, ChIP-chip and transcriptome profiling assays revealed 33 common target genes, almost half of which have not been characterized (Table 2.6). Over 60% of the characterized ones encode either secondary metabolic biosynthesis enzymes or regulators. Of the remaining overlapping targets, *SCO4878* is clustered with the phosphorus-free teichuronic acid biosynthesis genes (*SCO4880* and *SCO4881*) to form a cell envelope biosynthesis cluster that is activated by a response regulator PhoP upon phosphate starvation (Rodriguez-Garcia *et al*, 2007). Genes in this cluster are also under the control of AfsS (Lian *et al*, 2008), a pleiotropic antibiotic regulator, suggesting that there may be cross regulation between Crp, PhoP and AfsS. Transcription of *SCO4562* was downregulated by Crp and its upstream sequence contains two Crp association sites. Its gene product NuoA is a subunit of the respiratory chain NADH dehydrogenase (Bentley *et al*, 2002). *SCO4921*, encoding AccA2 (acyl-CoA carboxylase complex A subunit) and *SCO4979*, encoding PckA (phosphoenolpyruvate carboxykinase) were both stimulated by Crp and they both have a Crp association site in their upstream region. AccA and PckA are important enzymes in carbohydrate metabolism. AccA is essential for converting acetyl-CoA to malonyl-CoA, which is a precursor used in antibiotic biosynthesis. Overexpressing this enzyme results in a marked increase in Act production in wild-type *S. coelicolor* (Ryu *et al*, 2006), suggesting that Crp may also indirectly affect secondary metabolism by modulating primary metabolic pathways.

Combined with the ChIP-chip assay results, the transcriptome profiling data further support a role for Crp as a master regulator, particularly in the control of secondary metabolism.

2.3.8 The impact of Crp overexpression on secondary metabolism in diverse *Streptomyces* species

Multiple sequence alignment revealed 90.6% identity shared between Crp protein sequences across different *Streptomyces* species (Fig. 2.12). Given the importance of Crp in secondary metabolism in *S. coelicolor* as mentioned above, the *crp* gene was cloned downstream of a strong constitutive promoter *ermE** on a ΦC31-based integrating plasmid vector (whose target sequence is found in all sequenced *Streptomyces* species to date). It was then transformed into several *Streptomyces* species (Table 2.1), followed by assessment of its impact on secondary metabolism (Schmitt-John & Engels, 1992). SDS-PAGE and western blots showed that Crp levels were higher in all strains carrying *ermE*-crp* compared to controls carrying the empty plasmid vector; as a control for protein loading, Coomassie blue staining revealed similar total protein levels in overexpression and control strains (Fig. 2.13). Quantitative antibiotic production assays using *S. coelicolor* as the host strain showed that in liquid R5 medium, overproduction of Crp brought about a moderate increase in Act synthesis but had very subtle effects on Red and CDA production (Fig. 2.6). Upregulation of Act production was much more

prominent in surface-grown cultures, as indicated by the more heavily blue-pigmented bacterial lawn of the overexpression strain (Fig. 2.14). To detect colorless or unknown antimicrobials, bioassays using different *Streptomyces* species (Table 2.1) carrying either the *ermE*-crp* construct or the empty vector were conducted using an array of indicator strains (Biohazard Level 1: *E. coli*, *B. subtilis*, *M. luteus*, and *S. aureus*; Biohazard Level 2: *P. aeruginosa*, MRSA, VRE, and *K. pneumonia*) (Table 2.1). Only *S. sp.* WAT1 showed an increased antimicrobial activity against the Gram-positive non-pathogenic bacteria *B. subtilis*, *M. luteus*, and *S. aureus* (Fig. 2.15). The lack of effectiveness of Crp overexpression on the antimicrobial activity of all other species may be due to the protein level rise not being high enough to enhance secondary metabolite production or it could be that Crp target sites were already saturated with native levels of Crp. Alternatively, some novel or overproduced metabolite(s) may not be readily detectable using our bioassays. To circumvent this problem, high performance liquid chromatography coupled with mass spectrometry (HPLC-MS) was used in conducting comparative metabolic profiling of the wild-type and the Crp-overexpression strains.

With the extraction and separation conditions used, there was no change in the levels of known secondary metabolites produced by any of the tested *Streptomyces* species, *e.g.* Red (*S. coelicolor*), pristinamycin (*S. pristinaespiralis*) and chloramphenicol (*S. venezuelae*); however, some interesting differences were observed for compounds of unknown identity in the extracts of some species. Following 3 days of growth, a species with an *m/z* of 1017.641 showed a 2.6-fold reduction in the overexpression strain of *S. sp.* WAT1 (Fig. 2.16). This peak disappeared in the extracts harvested on day 7; instead, another substance with an *m/z* of 1033.636 was eluted in the same time range and showed a 5.2-fold decrease (Fig. 2.16). Since their molecular ion mass differ by 16 and they are both eluted at approximately 10:40 min, the latter may be the oxidized product of the former. *S. sp.* SPB74 is the species that produced the most abundant and significant changes in the comparative separation profiles. From the 3 day culture samples, the most drastically increased molecular ion (620.189 *m/z*) was upregulated by 22-fold, and the strongest repression was observed for a compound with an *m/z* of 1651.644 (Fig. 2.17). In the day 7 extracts, species with *m/z* of 471.270, 638.200, and 654.197 were detected at 10.8, 33, and 13.8 times higher levels, respectively, in the strain containing *ermE*-crp* relative to its empty plasmid control strain. A molecular ion with *m/z* of 620.189 was still present in the day 7 samples, but it was eluted at a much lower quantity (Fig 2.17). Given that ions 620.189, 638.200, and 654.197 *m/z* were eluted one after another, it is possible that species 638.200 *m/z* is the molecular ion 620.189 *m/z* plus an NH_4^+ adduct and that the 654.197 species is the oxidized product of the 638.200 *m/z* species (Fig. 2.17). Several attempts were made initially to predict the formulae of these molecules using the Data Analysis software (Bruker Daltonics) and this gave rise to a great number of possibilities for each molecule. Moreover, searching online chemical databases such as CAS (Chemical Abstracts Service) resulted in hundreds and thousands of compound structures for each candidate formula, making it challenging to elucidate the identity of these molecules given the information we currently have. Towards this end, in the future, these

up- or down-regulated species are worth more in depth characterization using high resolution mass spectrometry and nuclear magnetic resonance (NMR).

2.4 Discussion and Future Directions

Crp, as a founding member of the CRP/FNR family of regulators, predominantly functions as a positive transcriptional activator (Korner *et al*, 2003). In addition to the classical catabolite repression pathway established in *E. coli*, this regulator also controls a much broader range of cellular functions including primary and secondary metabolism, stress resistance, cell motility, and pathogenesis in many other bacteria (Fic *et al*, 2009). *E. coli* Crp is the most thoroughly studied member of this family and serves as the paradigm for the characterization of its homologs in other microorganisms.

S. coelicolor Crp, as reported previously, is important for germination and morphological differentiation (Derouaux *et al*, 2004b). According to observations described here, *crp* mutant spore germination was delayed by at least 6 h; however, even after 12 h, ~ 95% of the spores had failed to initiate germination. In addition to its inability to germinate, *crp* mutants also exhibited impaired vegetative growth: colony biomass was dramatically reduced and colonies appeared significantly smaller compared to the wild-type. Although retarded in vegetative growth, the *crp* mutant exhibited robust sporulation and was more resistant to heat stress; this correlated very well with its thicker spore wall compared to the wild-type strain. The absence of Crp also tended to favour sporulation and spore dispersal, as loss of the *crp* gene resulted in more abundant free spores than seen for the wild-type, while continual induction of Crp led to an increase in the number of unseptated aerial filaments relative to the wild-type. By manipulating the timing of Crp induction in a strain carrying the inducible *tipA-crp* construct, we were able to dissect the functions of Crp during sporulation and spore germination/vegetative outgrowth. The strain in which *crp* expression was induced during spore harvest, but not afterwards, exhibited wild-type levels of germination, whereas induction of Crp exclusively during subsequent replating only partially restored germination. This suggested that the reservoir of Crp from the previous life cycle (during sporulation/spore harvest) is indispensable and, at the same time, is sufficient for promoting normal germination. Alternatively, Crp activity during sporulation may contribute to the production of spore wall that is amenable to germination. Regardless of the cause of germination-competent spores, a new supply of Crp is required to sustain wild-type levels of regrowth, as evidenced by the drastic difference in colony appearance on agar plates. Collectively, these results suggest that Crp activity favours vegetative growth, since it promotes germination and colony enlargement early in the life cycle, and prevents morphological differentiation and aerial hyphae septation during later development. These findings prompted us to investigate the mechanism of Crp-mediated regulation at a molecular level.

Crp is broadly conserved in the actinomycetes, and multiple sequence alignments revealed 91.3% sequence identity in the DNA-binding motifs of three actinomycete Crp-

like proteins: *M. tuberculosis* Rv3676, *C. glutamicum* GlxR, and *S. coelicolor* Crp. Both the mycobacterial and corynebacterial proteins have been shown to bind the sequences related to the canonical binding motif [TGTGA$(N)_6$TCACA] using gel mobility shift assays (Agarwal *et al*, 2006; Rickman *et al*, 2005; Toyoda *et al*, 2011). However, interaction between *S. coelicolor* Crp and such sequences has never been established *in vitro* (Derouaux *et al*, 2004a; Derouaux *et al*, 2004b). Consequently, we undertook an unbiased *in vivo* approach to identify Crp targets by performing ChIP-chip assays. Compared to the conventional ways of characterizing protein-DNA interactions, ChIP-chip analyses allow us to capture transient interactions and generate a genome wide map of Crp association sites in a single experiment (Buck & Lieb, 2004). Interestingly, over 50% of the sequences coprecipitated with Crp were localized within ORFs; this is an unusual position for binding by a transcription regulator. Transcription activators typically bind upstream of promoters, where they help to recruit RNA polymerase to the promoter either by interacting with subunits of the RNA polymerase, or by changing the conformation of the DNA and this is unlikely to be the mechanism by which Crp and its intragenic association sequences act. There is, however, a specific class of short DNA sequences (30-200 bp) known as enhancers that are recognized by activator proteins, and binding to these sequences serves to regulate transcription units over distances of up to 15 kb in prokaryotes and 60 kb in eukaryotes (Buck *et al*, 2000; Sasse-Dwight & Gralla, 1990). Such enhancer elements are widespread in eukaryotes, but are much less common in bacterial gene regulation, since only promoters specific for σ^{54} are known to be regulated by enhancers (Buck *et al*, 2000). In our case, the intragenic binding sequences identified for Crp are not likely to be typical σ^{54}-associated enhancer elements, as σ^{54} - and its cognate activators - are not found in any *Streptomyces* species (Bentley *et al*, 2002). Although the biological rationale for binding sites both in intergenic regions and within coding sequences is not clear, it is not unprecedented. A ChIP-chip study of the *E. coli* Crp protein revealed a very noisy background, with thousands of low affinity association sites, and 68 unambiguous high affinity peaks, of which five were located in coding regions and 34 were positioned at, or close to, promoters (Grainger *et al*, 2005). Besides directly affecting transcription initiation, Crp could also function in other ways: (1) Crp may interact with *cis*-elements in the coding regions to block transcription of the genes, similar to the scenario where *E. coli* Crp binds downstream of promoters and blocks the progression of RNA polymerase (Kolb *et al*, 1993; Musso *et al*, 1977); (2) Crp may regulate flanking downstream genes (the coding region may overlap with the regulatory sequences of adjacent genes); (3) Crp binding may influence the local structure of the chromosome by bending the DNA and facilitating transcription initiation (followed by dislodging of Crp). We also can not exclude the possibility that these associated sites may correspond to *cis*-acting elements activating expression of yet-to-be-identified non-coding RNAs or antisense RNAs.

The dispersed genome-wide binding/association pattern of both the *E. coli* and *S. coelicolor* Crp is reminiscent of that of nucleoid-associated proteins like integration host factor (IHF). Nucleoid-associated proteins play an essential role in chromosome organization, replication, segregation, repair and expression (Browning *et al*, 2010).

Among the various nucleoid-associated proteins, IHF is most similar to Crp in the way it bends DNA sharply (Browning *et al*, 2010). It associates with DNA in both sequence-specific and –nonspecific manners to compact the chromosome and regulate gene expression, frequently in concert with other transcription regulators (Browning *et al*, 2010). This may partly account for the challenge in validating the selected ChIP-chip targets using *in vitro* strategies. In addition, inside living cells, the flanking DNA sequences and structures may be crucial for Crp to properly interact with its target DNA. It is also conceivable that Crp functions in conjunction with a co-factor to fully carry out its function.

ChIP-chip assays have been successfully used to examine the genomic distribution of many transcription regulators and to generate binding motifs for these proteins (Tomljenovic-Berube *et al*, 2010; Toyoda *et al*, 2011). However, with *S. coelicolor* Crp, the lack of prior knowledge of candidate target genes and the difficulty in verifying direct binding using traditional *in vitro* techniques has made it difficult to derive a consensus binding sequence, even though 'Crp' induction-ChIP assays coupled with qPCR had indirectly validated several ChIP-chip targets. It is possible that we have not found the correct buffer conditions to support Crp-DNA binding or it may be that the complex gets disrupted during electrophoresis. In some instances, such difficulties have been overcome by performing DNase I footprinting assays (Leblanc & Moss, 2009), where the Crp-DNA complex would not be required to survive electrophoresis; instead, the complex is exposed to DNase I *in situ* before the digested fragments are separated on a denaturing gel to visualize the pattern of protection mediated by Crp (Galas & Schmitz, 1978). Alternatively, surface plasmon resonance (SPR) has proven effective in some cases (Majka & Speck, 2007). In SPR, DNA is immobilized on the surface of a sensor chip while the mobile phase containing Crp (either purified protein or crude cell extract) flows over the top of the chip. If Crp binds to the target DNA, the mass of the immobilized material will be changed and this is recorded as the change in refractive index of the solution adjacent to the sensor surface (Stockley & Persson, 2009). Another option involves immobilizing purified Crp on the bottom surface of wells in a microtitre plate and adding fluorescently labeled DNA into the wells. Following incubation and washing, the signals of bound DNA may be recorded using a microplate reader (Zhang *et al*, 2003). The last two methods have the advantage of recording interactions in real time, which may be particularly important for transient interactions. Moreover, fluorescent microplate-based assays also allow high throughput examination of many DNA probes at the same time. In addition to *in vitro* binding assays, ChIP targets may be also validated using *in vivo* assays. Bacterial one-hybrid assays involve fusing a transcription factor, *i.e.* Crp in our case, to a subunit of RNA polymerase and transforming this construct into a host strain containing another vector carrying a library of random oligonucleotide sequences cloned upstream of reporter genes. Only when Crp interacts with a target DNA sequence, will the RNA polymerase be recruited to the promoter and initiate transcription of the reporter genes (Meng *et al*, 2005). By sequencing all the oligonucleotides showing positive interaction and making an alignment, we would be able to delineate a Crp binding sequence.

In the event that we are able to successfully establish Crp-DNA binding in gel mobility shift assays, we could then also address DNA bending Crp using circular permutation assays. The Crp binding motif (identified via DNase I footprinting and bacterial one-hybrid assay or generated by making alignment of ChIP binding sites DNA sequences) would be positioned at different sites along the length of probes radioactively labeled at one end (Herrera *et al*, 2012). When examined using gel shift assays, if Crp really bends the target DNA sequence, the probe with a binding site in the middle will migrate most slowly as DNA bending at this position produces shortened end-to-end distance, while the ones carrying the binding motif at one end will run most rapidly due to their long end-to-end distances (Kahn, 1999). DNA bending can also be detected utilizing fluorescence resonance energy transfer (FRET) from the donor fluorophore, which is used to label one end of the DNA fragment, to the acceptor fluorophore, which is used to label the other end (Khrapunov *et al*, 2006). Upon excitation of the donor with an external light source, the acceptor will be excited and emits light at a different wavelength only when the two ends of the DNA are brought into close proximity (10-100 Å).

Although the binding sequence of Crp and its regulatory mechanism are currently undeciphered, we have made significant progress towards understanding the downstream regulatory network affected by Crp. Classifying Crp associated genes on the basis of corresponding gene function led to the identification of several major target groups, including genes involved in primary metabolism, secondary metabolism, gene regulation and ribosome biogenesis, and morphological development. This is consistent with our transcriptome profiling data, which also had transcriptional regulators and secondary metabolism-associated genes ranked as the two largest subcategories.

The role of Crp as a master regulator controlling a number of other transcription factors is well correlated with published literature on its orthologs (den Blaauwen & Postma, 1985; Fic *et al*, 2009; Petersen & Young, 2002; Rickman *et al*, 2005; Toyoda *et al*, 2011), but its impact on secondary metabolism in *S. coelicolor* has never been thoroughly investigated, although some studies have noted reduced Act production by surface-grown *crp* mutant cultures (Derouaux *et al*, 2004b; Piette *et al*, 2005). Using quantitative assays, I have, for the first time, demonstrated diminished Act and CDA levels and a reproducible delay in Red production in the *crp* deletion mutant. yCPK was not examined due to the lack of an appropriate protocol. Insight into the molecular mechanisms underlying these phenotypes was provided by ChIP-chip and microarray experiments. All four antibiotic biosynthetic clusters (*act*, *red*, *cda*, and *cpk*) were upregulated following Crp induction and each contains several Crp binding sites, with at least one site falling within their cognate pathway-specific regulatory genes. Additionally, four other secondary metabolic clusters were also affected by Crp induction, with *SCO6429-6438* being upregulated and *SCO5222-5223* downregulated, suggesting that Crp may act as both a positive and a negative regulator of secondary metabolism. Several pleiotropic antibiotic regulatory genes were also affected by Crp induction, *e.g. relC*, *ppk*, and *afsS*. *relC*, whose expression was stimulated by Crp, codes for ribosomal protein L11 and is required for the

activation of the ribosome-bound ppGpp synthetase RelA (Chakraburtty & Bibb, 1997; Kang *et al*, 1998). Crp may therefore boost secondary metabolism in a ppGpp-dependent manner. Crp may also contribute to increasing the pool of free monophosphate by enhancing *ppk* gene expression, and this may have a negative effect on secondary metabolism (Chouayekh & Virolle, 2002). Our data did not show any direct connection between Crp and *afsS*, however, there is a Crp binding site within the coding region of the AfsK kinase inhibitor gene *kbpA*, whose gene product can block the AfsK/R/S signaling pathway (Umeyama & Horinouchi, 2001). In addition, *SCO4878*, which appeared in both the ChIP-chip and transcriptome profiling experimental data sets, is directly regulated by PhoP and displays reduced expression in the *afsS* mutant strain, suggesting cross regulation between Crp, AfsS, and PhoP-regulated pathways. Based on transcriptome profiling experiments, Crp shares 38 targets with PhoP and 25 targets with AfsS respectively and 11 genes are overlapped by the regulons of all three transcription factors (Fig. 2.18, Table 2.7).

Given the drastic developmental abnormalities of the *crp* mutant, it was not surprising to find Crp association sites within genes whose products contribute to aerial hyphae differentiation and sporulation. Although *crp* is capable of raising aerial hyphae and sporulating (albeit abnormally), chaplin (*chpC*, *chpD*, and *chpH*) and rodlin (*rdlA* and *rdlB*) genes were all identified as Crp associated genes in the ChIP-chip experiments. The chaplins are major components of the hydrophobic sheath enclosing aerial hyphae, whereas the rodlins are proposed to organize the hydrophobic chaplins into paired rodlets on the surface of aerial hyphae and spores (Claessen *et al*, 2002). Other than the chaplins and rodlins, products of the *ram* gene cluster are also required for aerial hyphae erection on rich medium (Kodani *et al*, 2004). Two genes in this cluster, *ramC* and *ramS*, were activated by Crp induction. *ramS* encodes a lantibiotic pre-peptide that is modified by the product of *ramC*, yielding the surfactant peptide SapB that helps aerial hyphae break surface tension in aqueous environments, allowing their extension into the air (Elliot *et al*, 2007; Kodani *et al*, 2004). Regarding the target genes important for sporulation, *mreB* is essential for lateral cell wall synthesis in rod-shaped bacteria, but Mre proteins in *S. coelicolor* cooperate to direct the synthesis of thickened spore wall resistant to detrimental environments (Kleinschnitz *et al*, 2011). Both *mreB* (*SCO2611*) and *mreC* (*SCO2610*) possess Crp associated sequences within their coding regions and *mreC* was also upregulated by Crp induction in the microarray studies. This seems to contradict our observation that the absence of Crp leads to thicker spore wall. However, ten cell wall hydrolase genes (*SCO3097*, *SCO4108*, *SCO4561*, *SCO4582*, *SCO4847*, *SCO5029*, *SCO5660*, *SCO6131*, *SCO6773*, and *SCO6884*) were either associated with or upregulated by Crp, suggesting they may ensure efficient exit from dormancy by promoting hydrolysis of the spore wall (Haiser *et al*, 2009).

According to the data obtained in this study, the impact of Crp activity is far from restricted to development and antibiotic production. For instance, *S. coelicolor* encodes 12 universal stress proteins (USPs), six of which displayed lower transcription levels when Crp was induced. Their precise role in *S. coelicolor* is still unknown, but the

deletion or downregulation of their orthologs in *E. coli* and *M. tuberculosis* impaired the ability of these bacteria to survive adverse conditions and prolonged growth arrest (Kvint *et al*, 2003; O'Toole & Williams, 2003). Most recently, the USPs of several *Mycobacterium* species were found to be part of the regulon of DosS/T/R, a two component system mediating hypoxic response and regulating dormancy in *Mycobacteria* (Park *et al*, 2003). It is noteworthy that the *S. coelicolor* ortholog of the response regulator gene *dosR* (*SCO0204*) was repressed by Crp, indicating that the USPs may be repressed by Crp in a DosR-dependent manner. This notion is further supported by the fact that the orthologs of four other hypoxia induced genes (*SCO0177*, *SCO0214*, *SCO125*, *SCO0616*), whose expression also requires DosR, were downregulated following Crp induction as well. Although neither the USPs nor DosR has been characterized in *S. coelicolor*, it is possible that Crp is involved in mounting a cellular response to oxygen deprivation. To address this, cobalt chloride (a hypoxia inducer) could be added to the cultures of wild-type, the *crp* mutant and the Crp overexpression strains to mimic hypoxic stress (Tsuzuki *et al*, 2012). Phenotypic analyses such as cell morphology, antibiotic production and SEM/TEM could then be conducted to evaluate whether Crp overproduction and deletion lead to any difference in cell survival and development under hypoxic conditions.

We expected to identify *cya* (*SCO4928*) as a repressed target of Crp, due to the substantially higher cAMP levels observed in the absence of Crp and the remarkable decrease in cAMP levels upon Crp expression. However, this gene was neither downregulated in the microarray assays, nor associated with Crp in ChIP-chip experiments, suggesting that Crp may indirectly affect cAMP synthesis. The negative feedback regulation of cAMP production by Crp is not unprecedented, as previous research has shown *crp* deletion to give rise to over 20 and 50 times higher cAMP levels in *E. coli* and *Salmonella typhimurium*, respectively (Daniel, 1984; den Blaauwen & Postma, 1985). In these instances, enhanced cAMP production is mediated through elevated phosphorylation of Enzyme II A of the phosphotransferase system (PTS), which in turn activates adenylate cyclase, the enzyme synthesizing cAMP (Takahashi *et al*, 1998). Phosphorylation of Enzyme II A (encoded by *crr/SCO1390*) depends on phosphoryl carrier protein (HPr, encoded by *ptsH/SCO5841*) and Enzyme I (EI, encoded by *ptsI/SCO1391*). Neither *ptsH* nor *ptsI* was identified in our microarray or ChIP-chip assays. However, the upstream region of the *crr* gene has a Crp association site, although its transcript level was not affected by Crp induction. The PTS system of *S. coelicolor* is not involved in carbon catabolite repression exerted by glucose, but instead responds to N-acetylglucosamine (GlcNAc) (Kamionka *et al*, 2002; Nothaft *et al*, 2003). This system is controlled by a GntR family transcription factor, DasR, which represses the PTS expression in the presence of low GlcNAc levels (Rigali *et al*, 2006). The gene encoding DasR, *SCO5231*, was not identified in our study, but the downstream gene *SCO5230* was activated upon Crp induction. *SCO5230* encodes an integral membrane protein and is potentially a functional partner of DasR as predicted based on their genomic context (Jensen *et al*, 2009). Taken together, Crp may negative regulate cAMP synthesis directly by repressing the transcription of the PTS component genes (*ptsH, ptsI, crr*) or indirectly

by activating DasR. Since GlcNAc was not used as a carbon source in our experiment, we tend to favor the notion that Crp may inhibit phosphorylation of Enzyme II A by negatively regulating *ptsH* and *ptsI*. Therefore, it will be useful to check the transcription levels of the PTS component genes in the *crp* (pIJ6902*crp*) strain after Crp induction. Direct versus indirect regulation by Crp could be distinguished using *in vitro* binding assays. At the same time, the phosphorylation state of Enzyme II A in the wild-type, *crp* deletion mutant and the *tipA-crp* carrying strain could be analyzed using SDS-PAGE and immunoblotting as *E. coli* Enzyme II A displayed reduced mobility upon phosphorylation (Takahashi *et al*, 1998). Alternatively, phosphorylated Enzyme II A can be detected using antibodies developed by immunizing rabbits with a synthetic phosphopeptide encompassing the Enzyme II A phosphorylation site and its surrounding amino acids (Marshall, 1995).

Other gene clusters positively regulated by Crp include: the pentose phosphate pathway (*SCO1921-1926*), an anabolic process that oxidizes glucose to generate pentoses and NADPH; the *dnaK* operon (*dnaK–grpE–dnaJ–hspR*/*SCO3668-3671*), involved in the heat shock response; the riboflavin synthesis operon (*ribBMAH/SCO1443-1440*), whose product is a precursor of coenzymes FAD and FMN; ribosome biogenesis genes (*SCO4701-4728*); and the *dhsA-D* cluster (*SCO4855-4858*), whose products contribute to the citric acid cycle (Bentley *et al*, 2002; Kallifidas *et al*, 2010). In addition, Crp also promotes the biosynthesis of amino acids like arginine (SCO1576-1579), cysteine and methionine (*ssuABC/SCO6094-6096* and *cysNDCHI/SCO6097-6102*) (Lee *et al*, 2005). One of the genes repressed following Crp induction, *SCO4562*, belongs to a cluster coding for subunits of respiratory chain NADH dehydrogenase. This cluster is one of the most strongly conserved regions between the *S. coelicolor* and *M. tuberculosis* genomes (Bentley *et al*, 2002). In the future, all of the above targets are worth further investigation to fully clarify the regulatory role for Crp.

Considering the immense difference in antibiotic production observed for wild-type *S. coelicolor* compared with the *crp* mutant, we were expecting to improve the synthesis of known antibiotics and, more importantly, to stimulate novel compounds by overexpressing Crp in both *S. coelicolor* and other *Streptomyces* species. However, neither the antibiotic production assays for *S. coelicolor*, nor bacterial bioassays for other Crp overexpressing *Streptomyces* species yielded significant differences. Using comparative metabolic profiling, we did not detect any significant differences for the known or predicted secondary metabolites made by any of the examined *Streptomyces* species. This may be due to the minor increase in Crp expression levels in the strain carrying *ermE*-crp*, since the strain has only one copy of the *ermE*-crp* construct integrated in its chromosome. This fact may also account for the discrepancy between our morphological observation and previous literature. Piette *et. al.* (2005) reported that when Crp was expressed from a high-copy-number plasmid, the colony failed to initiate aerial growth, whereas in our hands, the *ermE*-crp* containing strain was able to form aerial hyphae and sporulate. Alternatively, the characterized secondary metabolites may not be readily extractable or ionizable under the conditions used (*e.g.* actinorhodin was not

effectively extracted or ionized). Despite the lack of effect of Crp overexpression on known secondary metabolites, we were still able to successfully demonstrate a major impact of Crp on a number of other uncharacterized compounds. In the future, the *ermE*-crp* construct may be replaced by a high copy number plasmid expressing Crp and this strain could then be subjected to antibiotic bioassays and comparative metabolic profiling, although one downside to this is that the overexpression stain will need continuous antibiotic selection and this could affect the indicator strains. High throughput conjugation may be set up to engineer the Crp overexpression construct into a much greater number of different *Streptomyces* species than the ones looked at in this study. In this way, bioassays could be performed in a high throughput format using indicator strains expressing a reporter gene, *e.g.* an *E. coli* variant expressing the Lux bioluminescence system of *Vibrio fischeri*, where the signals could be accurately quantified (Gverzdys, 2011). With the aim of identifying novel bioactive agents, various culture conditions, solvents and extraction procedures should be tested to extract as many compounds as possible. Molecules showing differential production profiles in the wild-type and overexpression strains may be isolated via preparative HPLC and further characterized using high resolution MS and NMR. The purified samples will also be tested for anti-fungal and anti-cancer activities. They will serve as good candidates for discovery of novel antimicrobial and anti-cancer agents.

The research presented here provides us with a thorough phenotypic analysis of the impact of Crp throughout the life cycle of *S. coelicolor*, and has allowed us to make significant progress towards understanding the downstream regulatory network affected by Crp. More importantly, for the first time, we uncovered a connection between Crp and antibiotic biosynthesis as well as potential cross regulations between Crp and other transcription factors involved in development and primary/secondary metabolisms. The vast versatility of Crp governed targets supports a role for Crp as a master regulator in *S. coelicolor*.

2.5 Figures and Tables:

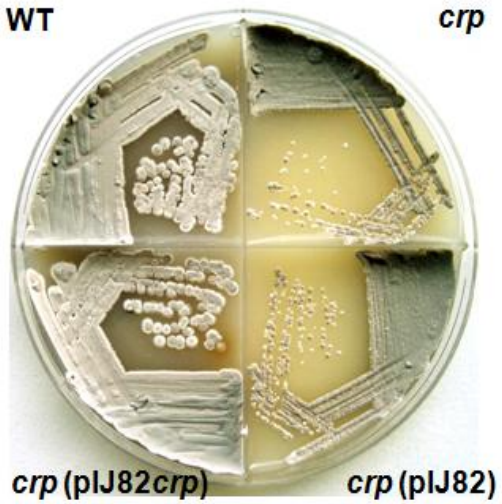

Fig. 2.1 Morphology of wild-type, *crp* mutant and *crp* (pIJ82*crp*) after 6 days' growth on MS agar. *crp* was cloned, along with its native promoter, into an integrating vector and the construct was conjugated into the *crp* mutant. The resulting strain was compared with the wild-type, mutant, and empty-plasmid-containing mutant strains.

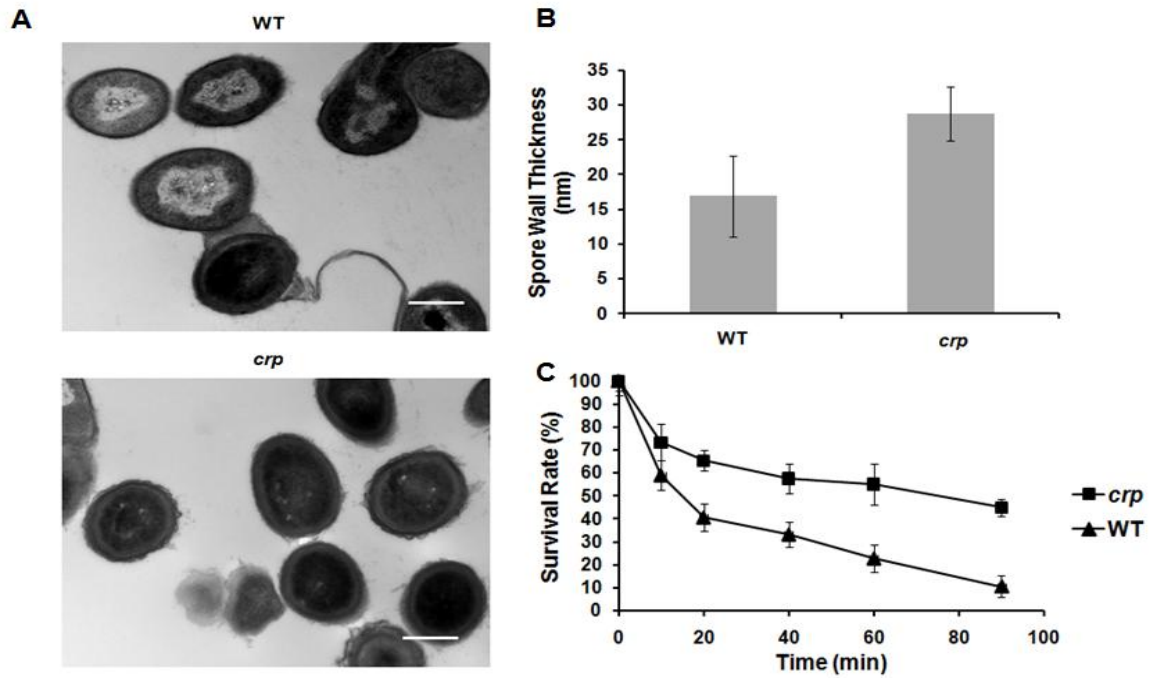

Fig. 2.2 Increased spore wall thickness and resistance to heat stress for the *crp* mutant strain. (A) TEM micrographs of wild-type (WT) (upper panel) and *crp* mutant (lower panel) grown on MS agar for 6 days; (Bars = 500 nm). (B) Quantification of spore wall thickness in wild-type and *crp* mutant strains. Around 100 spores of each strain were examined and the error bars indicate the standard deviations. (C) Heat shock assays of wild-type and *crp* mutant spores. Approximately 300 spores of each strain were incubated at 60°C for different lengths of time before being spread on DNA agar and cultured for 2 (wild-type) and 5 (*crp*) days. The data represent three independent experiments and the error bars indicate the standard errors.

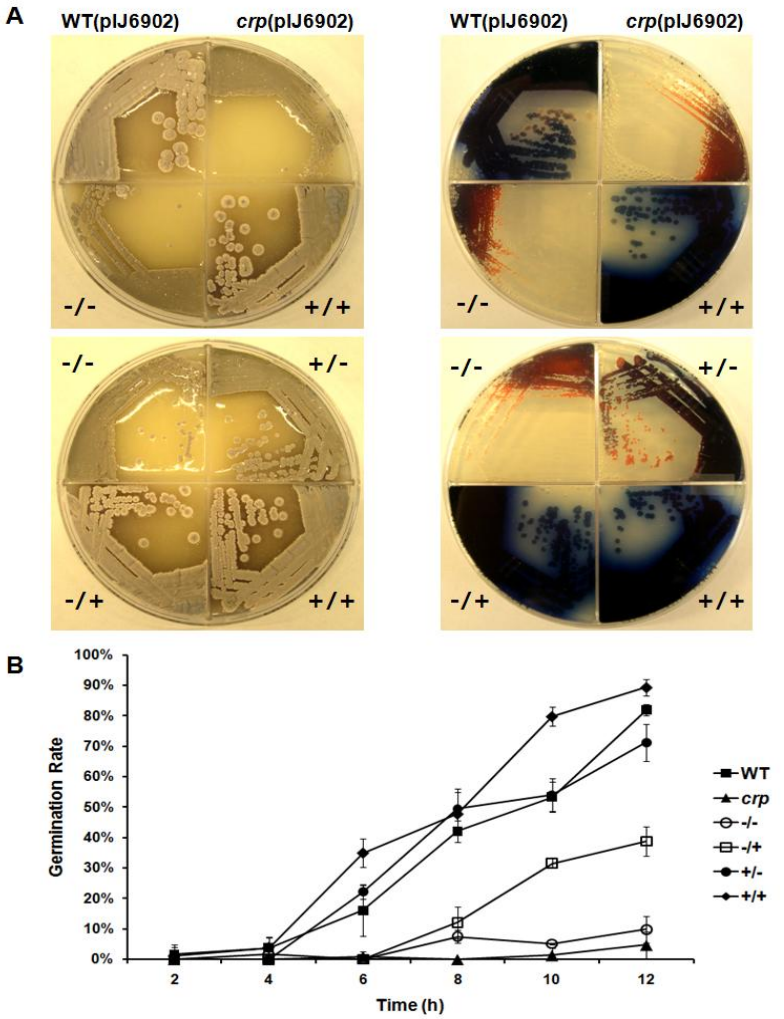

Fig. 2.3 The deletion of *crp* impairs germination and vegetative growth of *S. coelicolor*. *crp* was cloned downstream of a thiostrepton-inducible promoter *tipA* and the construct was transformed into the *crp* mutant. Spores of the strain were harvested both in the presence and absence of thiostrepton and each spore stock was replated in the presence or absence of thiostrepton. Wild-type (WT) and *crp* strains containing pIJ6902 were used as controls. (-/-): no Crp induction at any stage; (+/+): continual Crp induction throughout growth; (-/+): Crp induced exclusively during regrowth; (+/-): Crp induced exclusively during spore harvest. (A) Colony morphology of *crp* (pIJ6902*crp*) upon *crp* induction at different stages. The strains were cultured for 6 days on MS agar (left panels) or 4 days on R2YE agar (right panels). (B) Germination assay of wild-type, *crp* and *crp* (pIJ6902*crp*) strains. Spores were cultured on MS agar overlaid with cellophane disks for 12 h. A portion of the cellophane disk was excised every 2 h and examined using phase contrast microscopy. Images were captured and germinated spores were scored using Image J. The data shown were obtained from three independent experiments with four images taken for each timepoint (~ 200 spores were examined for each strain). The error bars represent standard deviations between three experiments.

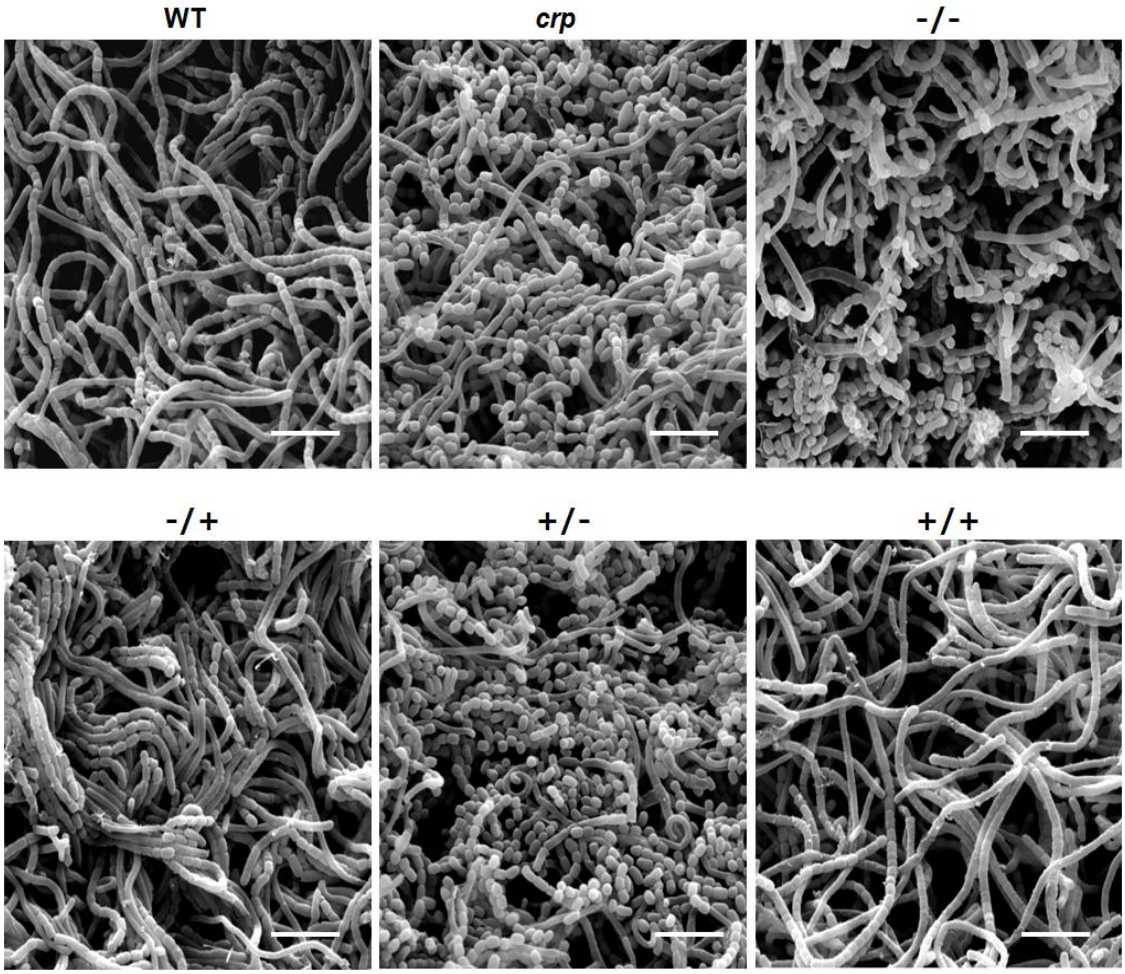

Fig. 2.4 SEM of wild-type (WT), *crp* mutant, and *crp* (pIJ6902*crp*) strains grown on MS agar for 6 days. For the meaning of '+' and '-', please refer to Fig. 2.3. White bars = 5 µm.

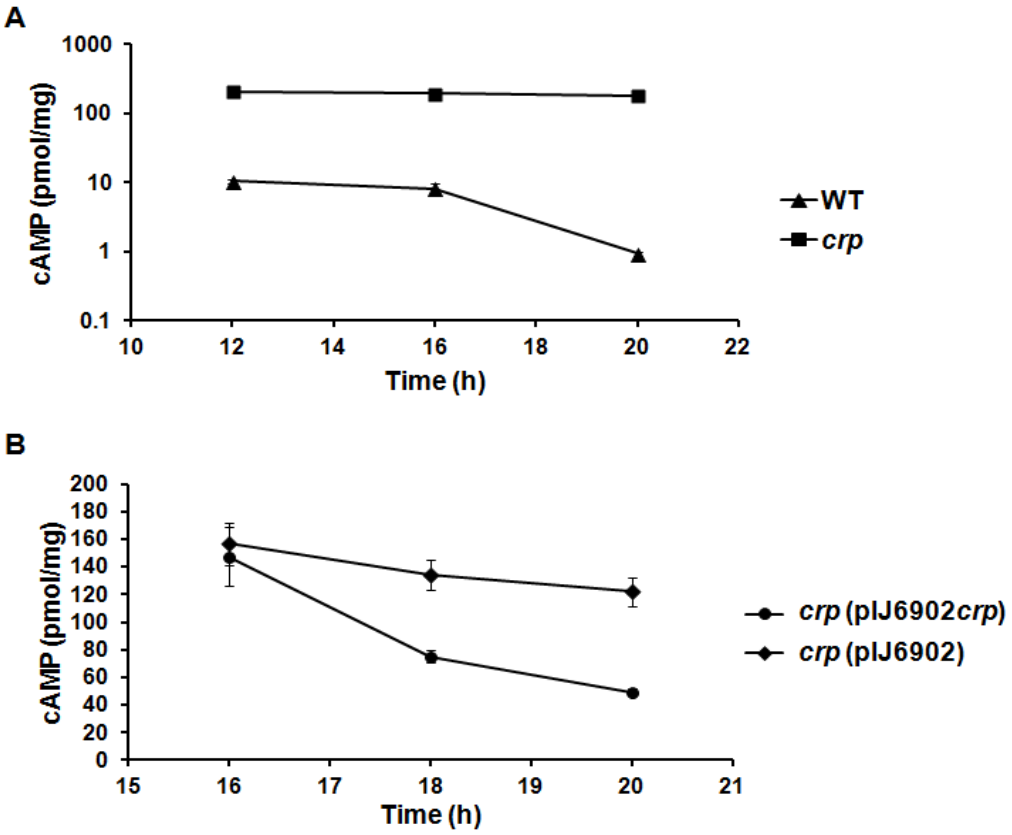

Fig. 2.5 Impact of Crp on cAMP levels. (A) Liquid cultures of wild-type (WT) and *crp* mutant strains were tested every 4 h starting at 12 h. *crp* mutant cultures were set up 48 h ahead of the wild-type. (B) Cultures of *crp* (pIJ6902crp) and *crp* (pIJ6902) strains were induced with thiostrepton at 16 h after an aliquot of the uninduced culture was taken for analysis. The cultures were then checked 2 and 4 h after induction. All results were normalized with the biomass of the mycelia. Only extracellular cAMP levels are shown since intracellular cAMP levels were not detectable. Three independent cultures were set up for each strain, with duplicate aliquots taken from each culture at each time point. The error bars indicate the standard deviations between three experiments.

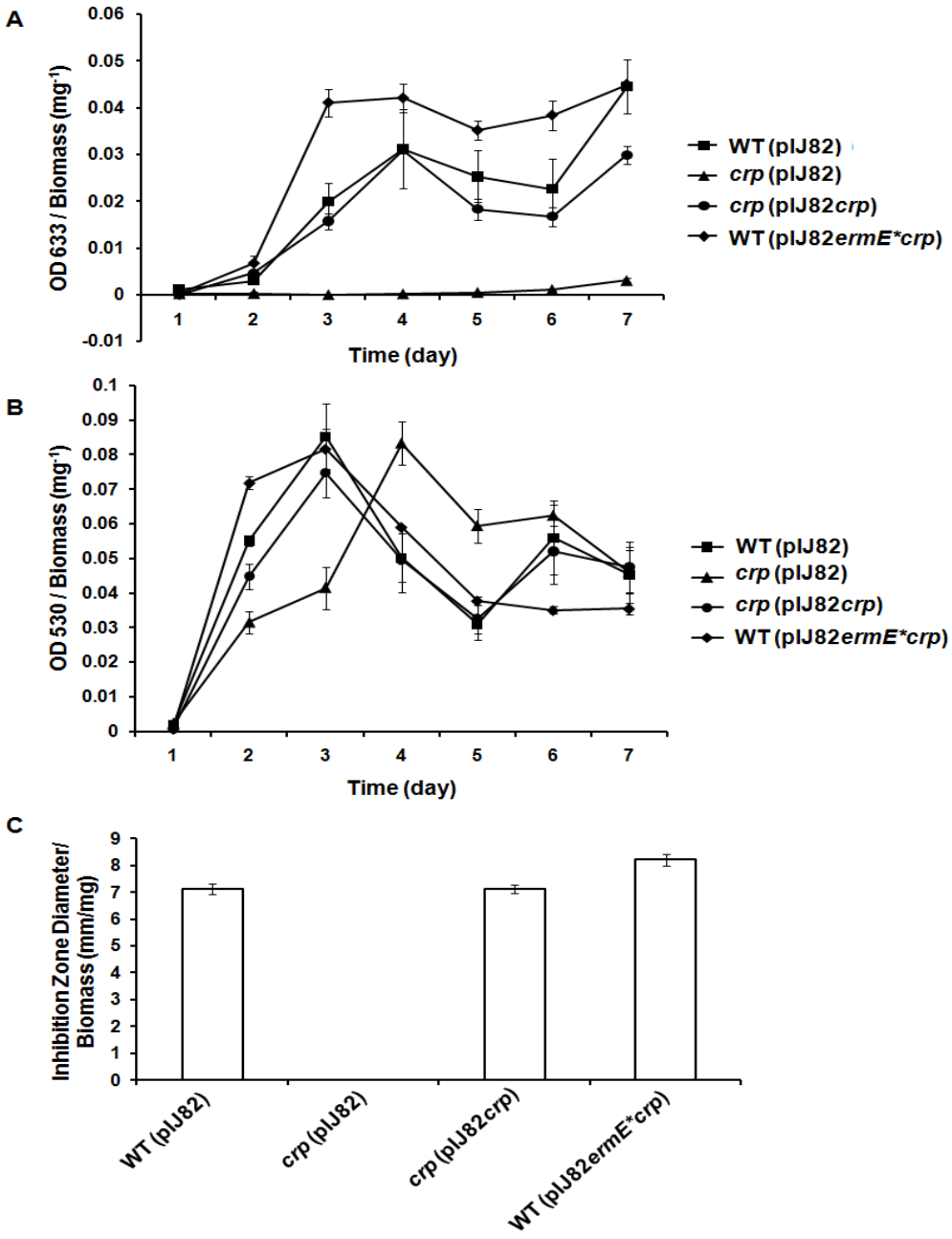

Fig. 2.6 Antibiotic production by wild-type (WT), *crp*, and *crp* mutants carrying *crp* transcribed from its native promoter [*crp*(pIJ82*crp*)] or *ermE promoter.** (A) actinorhodin; (B) undecylprodigiosin; (C) calcium-dependent antibiotic. The results of (A) and (B) are representative of three independent cultures with duplicate aliquots examined for each culture at each time point. The results for (C) were obtained from four independent assays. The data were normalized relative to the biomass of the mycelia. The error bars denote the standard errors for these experiments.

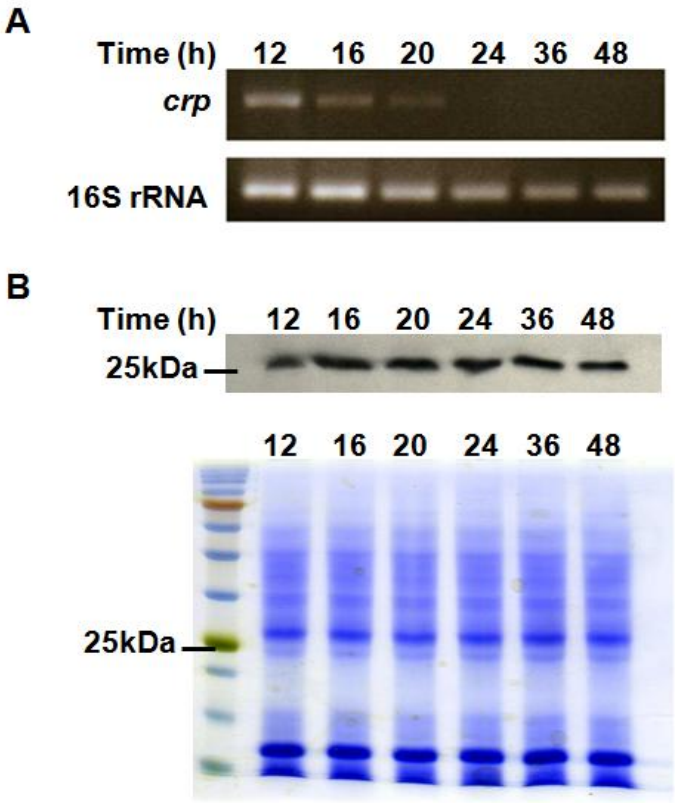

Fig. 2.7 Transcriptional and translational levels of Crp in wild-type cells over a 48 h time course. (A) RNA samples prepared from cells harvested at different time points were subjected to reverse transcription and PCR to evaluate *crp* transcript abundance. 16S rRNA was included as a control. (B) Cell extracts of the wild-type strain were resolved using SDS-PAGE. Crp protein was detected using immunoblotting. Coomassie blue staining was performed in parallel to ensure equal sample loading in all lanes.

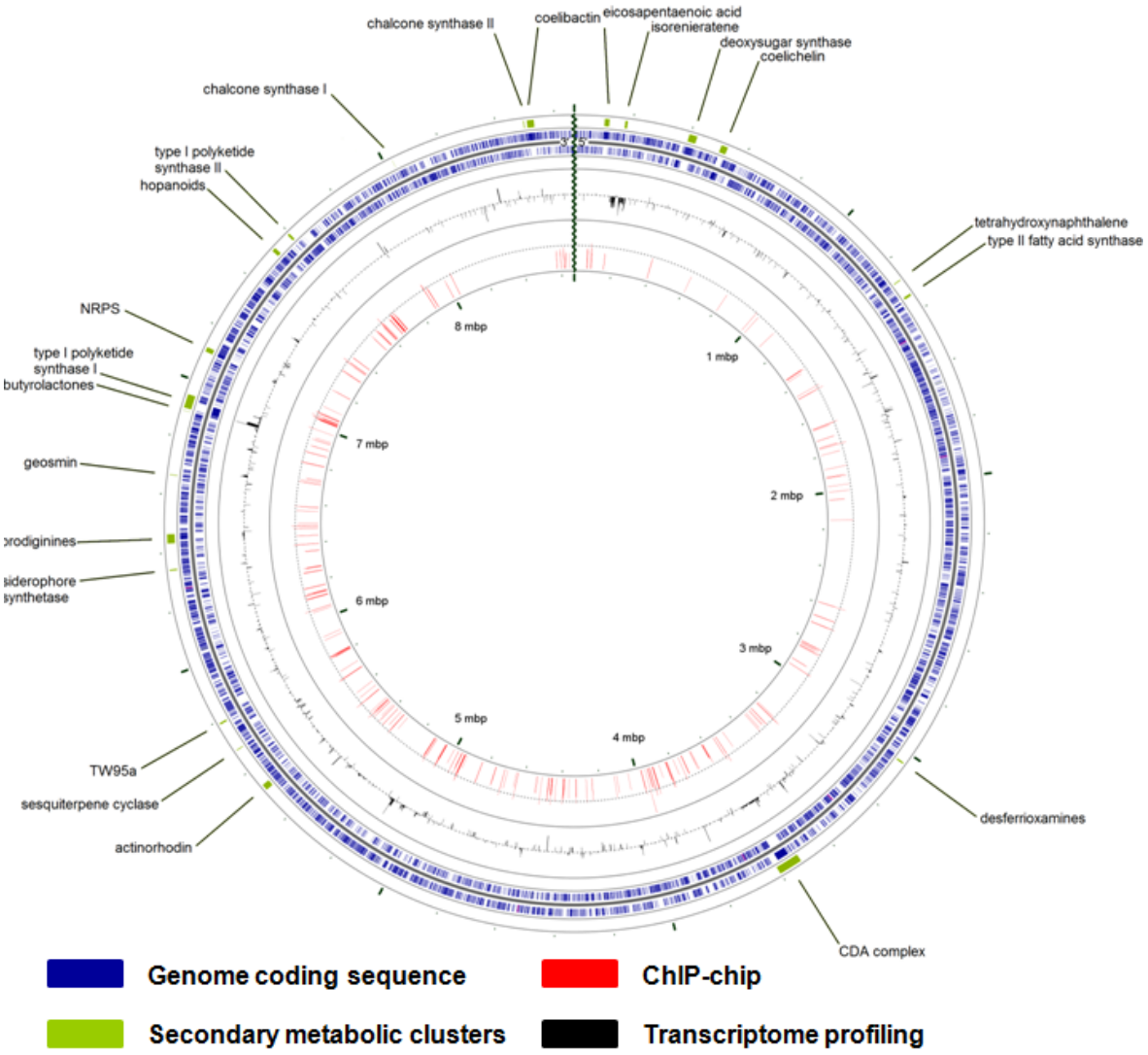

Fig. 2.8 Compilation map of the *S. coelicolor* chromosome showing the Crp ChIP-chip and transcriptome profiling data. Coding sequences in the genome are shaded in blue. The characterized or predicted secondary metabolic clusters are highlighted in green in the outermost circle (Bentley *et al*, 2002). The genome distribution of the Crp association sites and the Crp targets identified in transcriptome profiling are marked as red lines in the innermost circle and black lines in the second innermost circle, respectively. The black lines pointing out represent genes upregulated by Crp induction while those pointing in indicate downregulated genes. This map was created by Charles Yin using *CGView* software and was modified by myself (Stothard & Wishart, 2005).

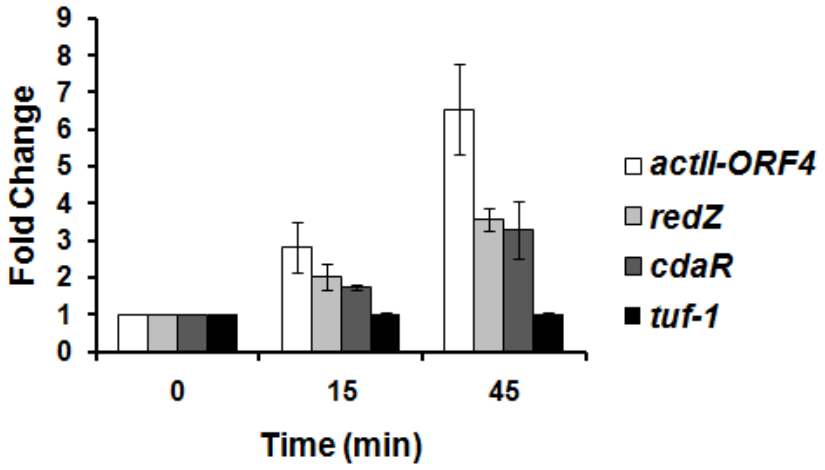

Fig. 2.9 Validation of *actII-ORF4*, *redZ*, *cdaR* as Crp-associated genes using 'Crp induction'-ChIP assays. The *crp* (pIJ6902*crp*) strain induced for different lengths of time were cross-linked and subjected to chromatin immunoprecipitation as described. The bound DNA was eluted and used as template to amplify secondary metabolism pathway-specific regulatory genes using qPCR. *tuf-1* gene did not coprecipitate with Crp in ChIP-chip assays and hence was used as a negative control. Three independent cultures were set up with four technical replicates for each culture. The error bars show the standard deviations between three experiments.

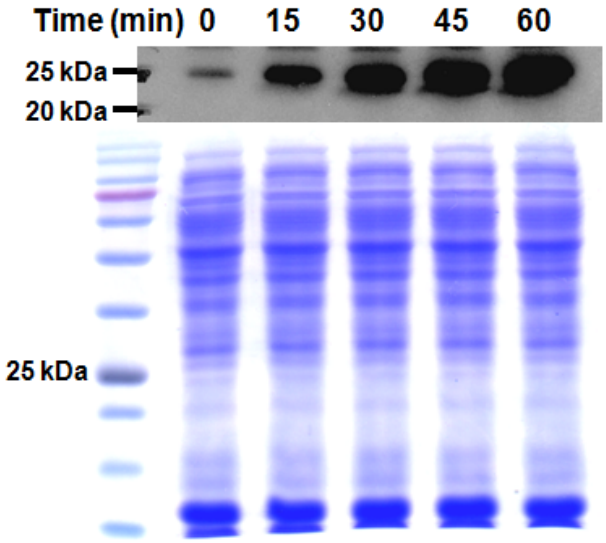

Fig. 2.10 Crp expression from the *tipA-crp* construct upon induction with thiostrepton. The strain *crp* (pIJ6902*crp*) was cultured and induced as described in Section 2.2.13. Cell extract was subjected to analysis using SDS-PAGE following by Coomassie blue staining as a protein loading control (bottom) and immunoblotting to determine Crp expression levels (top). In the immunoblots, the band at around 25 kDa is the signal of Crp.

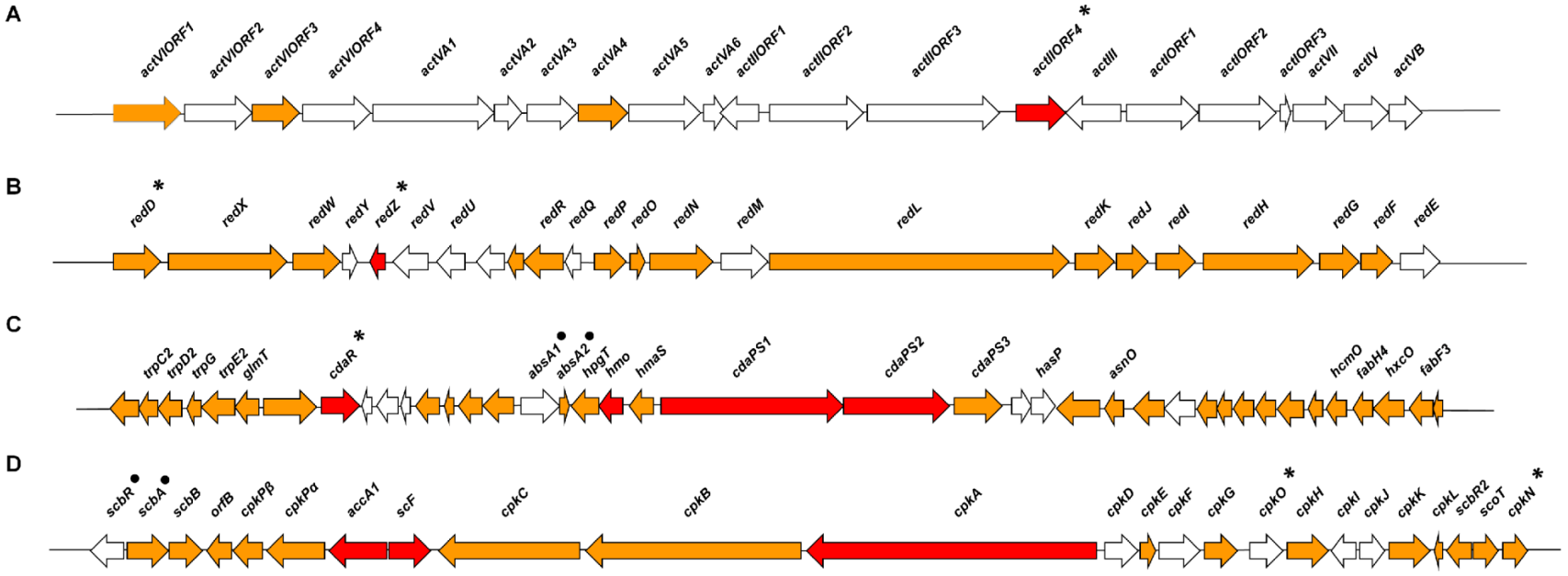

Fig. 2.11 Schematic representation of antibiotic biosynthesis clusters showing ChIP-chip targets and transcriptome profiling targets: (A) *act*, (B) *red*, (C) *cda*, and (D) *cpk*. The ChIP-chip targets are shaded in red for all four clusters and the upreguated genes are shaded in orange. The pathway-specific regulatory genes are marked with asterisks. The global regulatory gene names are marked with black dots.

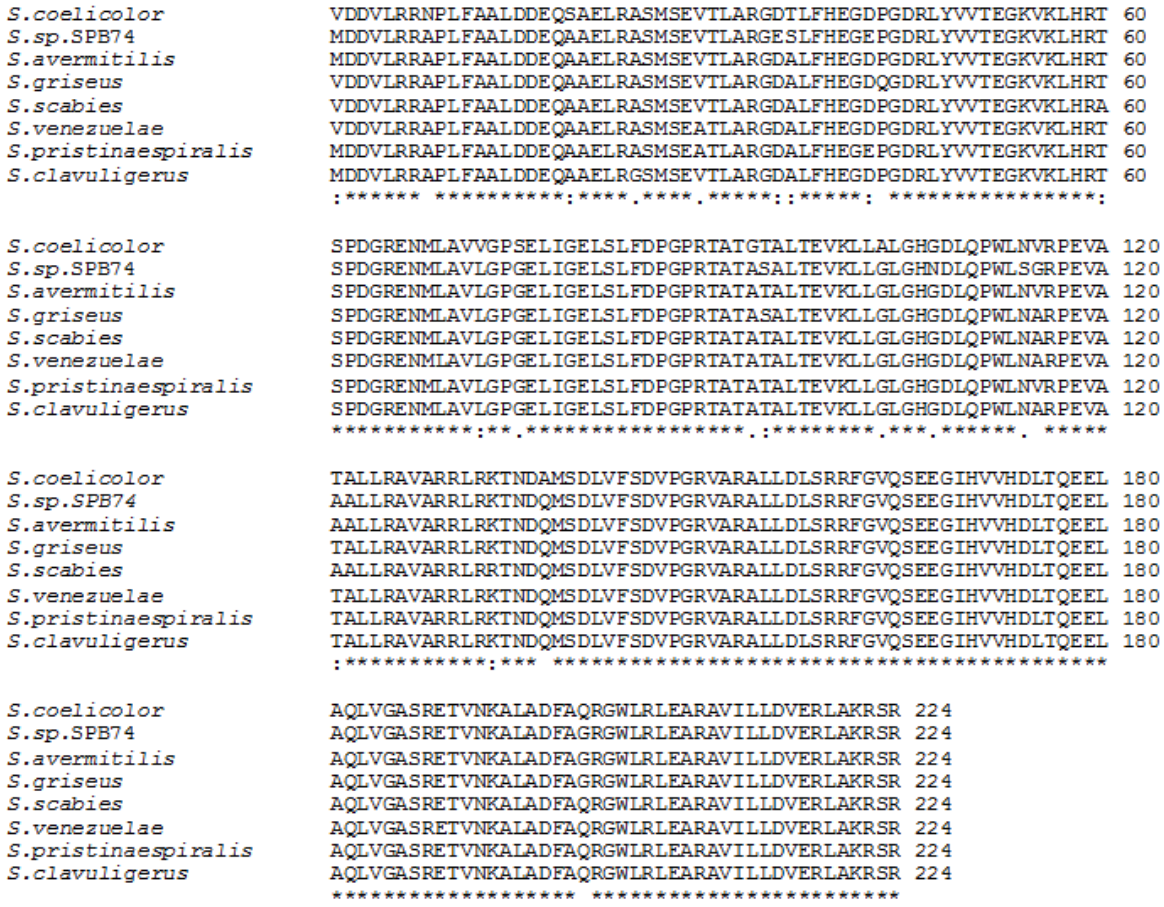

Fig. 2.12 Multiple sequence alignment of Crp family proteins of *Streptomyces* species. The Crp protein sequences were obtained from StrepDB, NCBI, and Broad Institute websites. (*) indicates a fully conserved residue; (:) indicates a strongly conserved residue; (.) indicates a weakly conserved residue. The alignment was generated using Clustal W2 on the EMBL-EBI webserver (Larkin *et al*, 2007).

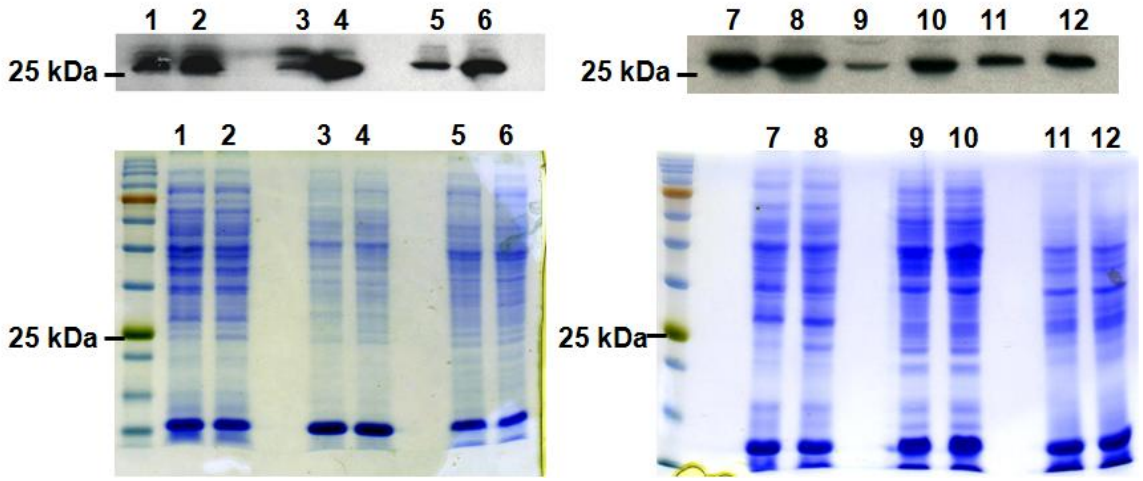

Fig. 2.13 Overexpression of Crp in *Streptomyces* species carrying *ermE*-crp*. Wild-type strains containing either the empty vector or the *ermE*-crp* construct were cultured in liquid R5 medium for 16 h before cell extracts were obtained. Equal amounts of cell extract from both strains were resolved using SDS-PAGE, followed by Coomassie blue staining as a protein loading control (bottom) and immunoblotting to determine Crp expression levels (top). In the immunoblots, the band at around 25 kDa is the signal of Crp. The loading order is as follows: 1, *S. sp.* WAT1 (pIJ82); 2, *S. sp.* WAT1 (pIJ82*ermE*crp*); 3, *S. sp.* SPB74 (pIJ82); 4, *S. sp.* SPB74 (pIJ82ermE**crp*); 5, *S. sp.* Ja2b (pIJ82); 6, *S. sp.* Ja2b (pIJ82ermE**crp*); 7, *S. coelicolor* (pIJ82); 8, *S. coelicolor* (pIJ82*ermE*crp*); 9, *S. venezuelae* (pIJ82); 10, *S. venezuelae* (pIJ82*ermE*crp*); 11, *S. pristinaespiralis* (pIJ82); 12, *S. pristinaespiralis* (pIJ82*ermE*crp*).

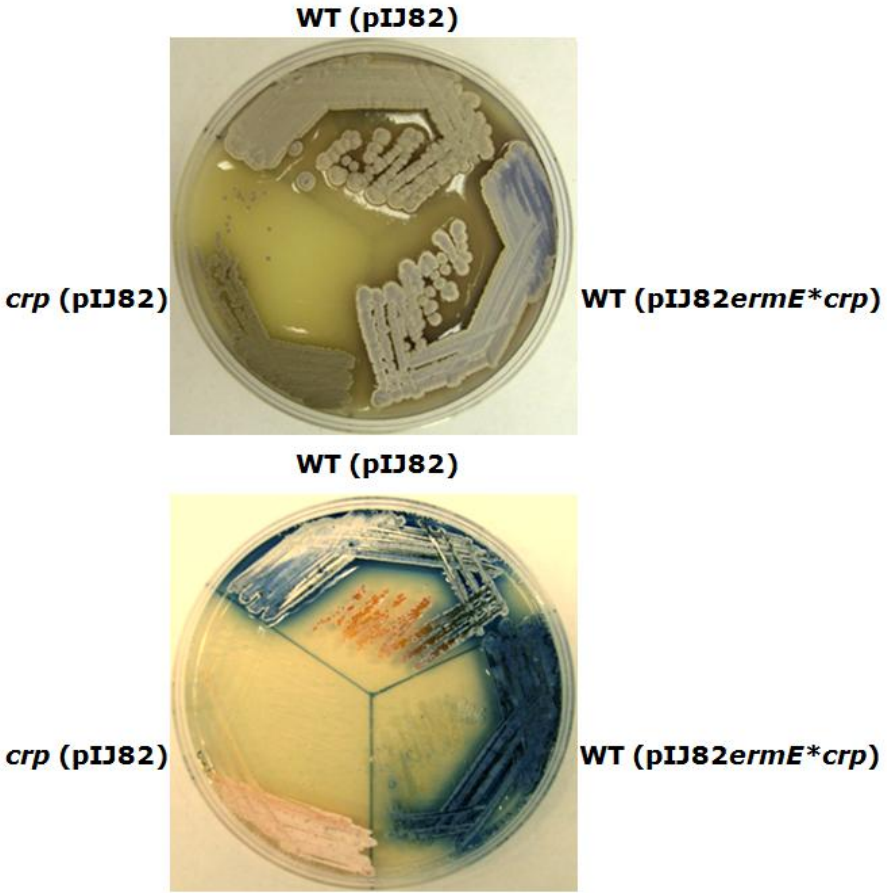

Fig. 2.14 Increased actinorhodin production in *S. coelicolor* (pIJ82*ermE*crp*). The cultures were grown on MS agar for 7 days (top) and R2YE agar for 4 days (bottom).

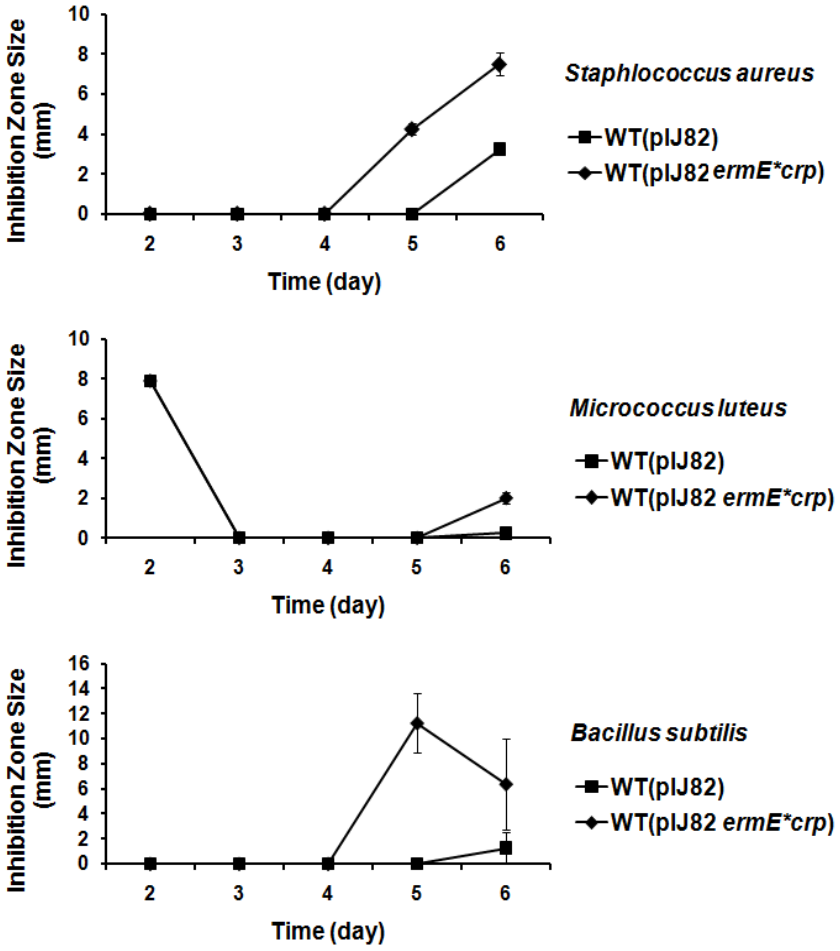

Fig. 2.15 Bioassays showing increased antimicrobial activity of *S. sp.* WAT1 (pIJ82*ermE*crp*) strain against Gram positive bacteria *S. aureus*, *M. luteus*, and *B. subtilis*. *Streptomyces* spores were spotted on DNA agar and incubated for different numbers of days before being overlaid with soft agar containing the indicator strains. Following overnight cultivation, sizes of zones of inhibition were measured. The results were obtained from four replicates and the error bars show the standard deviations between experiments.

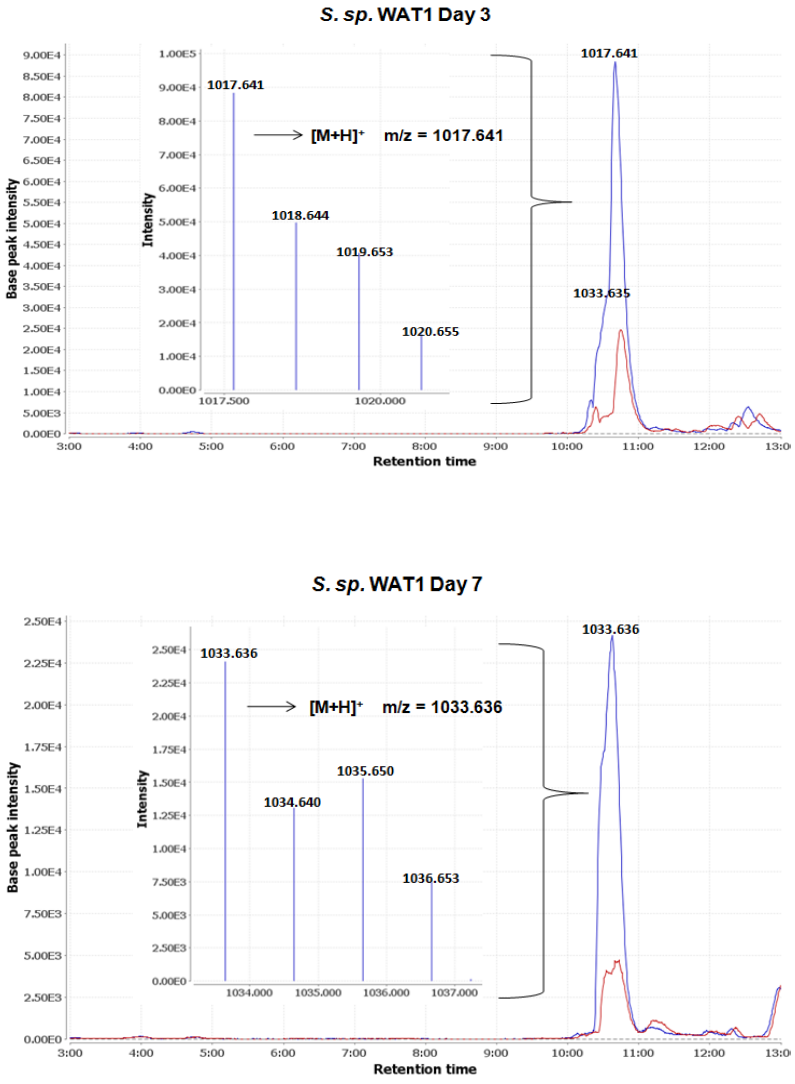

Fig. 2.16 Base peak chromatograms and mass spectra of the downregulated peaks upon Crp overexpression in *S. sp.* WAT1. In the base peak chromatograms, the wild-type strain carrying the empty plasmid is indicated with blue color and the overexpression strain is indicated with red color. Mass spectra are overlaid on top of the chromatograms. [M+H] indicates the molecular ions in the mass spectra. The numbers on top of the peaks are the mass/charge ratios of the peaks.

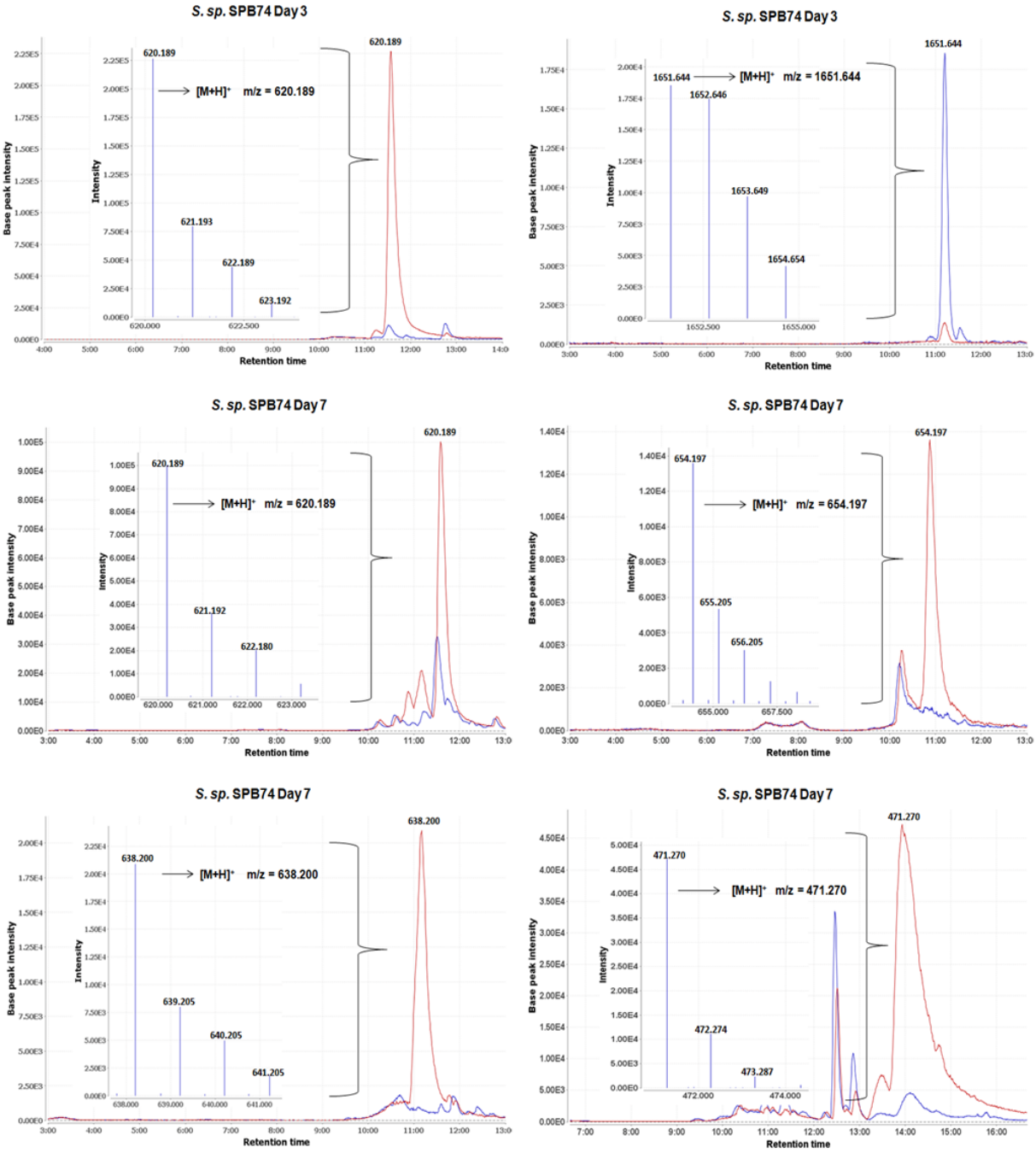

Fig. 2.17 Base peak chromatograms and mass spectra of the up- and downregulated peaks upon Crp overexpression in *S. sp.* SPB74. Please refer to Fig. 2.16 for the color scheme of the chromatograms and the meaning of all labelings.

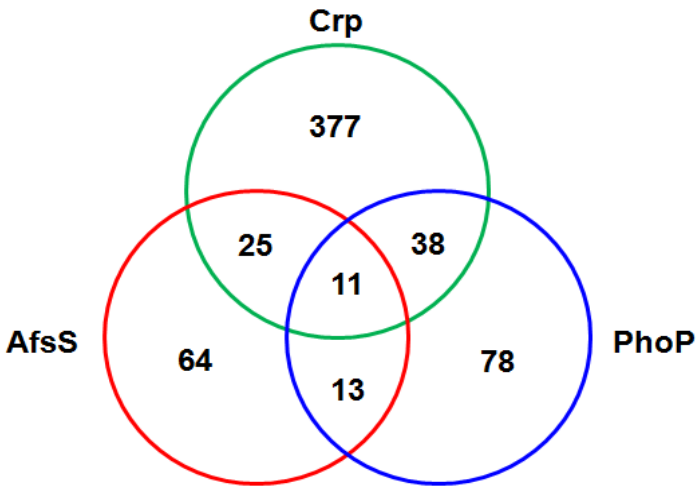

Fig. 2.18 A venn diagram showing the overlap between the regulons of Crp, AfsS and PhoP. The information on the target genes of AfsS and PhoP was obtained from transcriptome profiling data (Lian *et al*, 2008; Rodriguez-Garcia *et al*, 2007)

Table 2.1 Strains and plasmids used in this study

Strain	Genotype/characteristics/use	References
Streptomyces		
S. coelicolor M145	SCP1- SCP2-	(Kieser, 2000)
S. coelicolor crp	M145 *SCO3571::aac(3)IV*	This study
S. sp. Ja2b	Wild isolate	Gift from G. Wright
S. pristinaespiralis ATCC25486	Wild-type	(Blanc *et al*, 1995)
S. sp. SPB74	Wild-type	(Scott *et al*, 2008)
S. venezuelae ATCC 10712	Wild-type	(Stuttard & Dwyer, 1981)
S. sp. WAT1	Wild isolate	Gift from G. Wright
Escherichia coli		
DH5α	*fhuA2,* Δ*(argF-lacZ)U169, phoA, glnV44, Φ80,* Δ*(lacZ)M15, gyrA96, recA1, relA1, endA1, thi-1, hsdR17*	(Hanahan, 1983)
SE DH5α	Highly-competent (Subcloning Efficiency™) DH5α cells	Invitrogen
ET12567	*dam*, *dcm*, *hsdS*, *cat*, *tet*; carries plasmid pUZ8002	(MacNeil et al, 1992)
BL31 (DE3)	F–, *omp*T, *hsd*SB(rB–, mB–), *gal, dcm,* (DE3)	(Studier & Moffatt, 1986)
Others		
Bacillus subtilis		Gift from J. Nodwell
Staphylococcus aureus		Gift from J. Nodwell
Klebsiella pneumoniae		Biosafety Level 2 lab of Center for Microbial Chemical Biology (CMCB) at McMaster University
Micrococcus luteus		Same as above
Pseudomonas aeruginosa		Same as above
Methicillin-resistant *Staphylococcus aureus*		Same as above
Vancomycin-resistant *Enterococcus*		Same as above
Plasmids		
pIJ82	Integrative cloning vector; *ori* pUC18, *hyg*, *oriT*, RK2, *int* ΦC31, *attP* ΦC31	Gift from H. Kieser
pIJ6902	Integrative P*tipA* expression vector; *ori* pUC18, *oriT,* RK2, *int* ΦC31, *attP, tsr, aac*(*3*)*IV*	(Huang *et al*, 2005)
pET15b	Over expression of His_6-tagged proteins	Novagen
pET15b*crp*	His_6-tagged Crp overexpression plasmid	C. Hanke
pIJ82*crp*	Complementation construct carrying the *crp* gene downstream of its native promoter	This study
pIJ6902*crp*	Inducible construct with the *crp* gene under the control of a *tipA* promoter	This study
pMC500	pUC57 derivative carrying the *ermE** promoter	J. Swiercz

pIJ82*ermE*crp*	Constitutive overexpression plasmid carrying the *crp* gene cloned downstream of the *ermE** promoter	This study
pT77*strepII-virB8*	Overexpression construct carrying the Strep-tagged *Brucella suis* VirB8 overexpression plasmid.	Part II of the thesis
pCR2.1-TOPO	TOPO cloning vector, *ampR*, *kanR*	Invitrogen
TOPO-*crp*	pCR2.1-TOPO carrying the *crp* gene	This study

Table 2.2 Oligonucleotides used in this study

Name	Sequence (5'- 3')	Comments
P-Crp F	CAGTTCTAGAGGAATCGTGACGTGC	Complementation construct cloning
P-Crp R	CAGTGGATCCTGGAGTACTACCTCG	Complementation construct cloning
Crp F	CAGTCATATGGACGACGTTCTGCGG	Crp-inducible construct cloning
Crp R	CAGTGAATTCTGGAGTACTACCTCG	*crp*-inducible construct cloning
tipA F	TCAGAGAAGGGAGCGGAAGTGAGCT	*crp*-inducible construct cloning
*ermE**-Crp F	TACTACGAATTCCACAGGAGGTCAGTCAGGTGGACGACGTTCTGCGG	*crp* overexpression construct
*ermE**-Crp F	TACTACGGATCCTCAGCGGGAGCGCTTGGC	*crp* overexpression construct
M13 FWD	GTAAAACGACGGCCAGT	Routine cloning
M13 REV	CAGGAAACAGCTATGAC	Routine cloning
RT Crp F	CACGGAAGGCAAGGTCAAGC	RT-PCR of *crp* transcript
RT Crp R	AGCTCCTCCTGCGTCAGGTCG	RT-PCR of *crp* transcript
16SrRNA F	CCTGATGCCGAGCCGATTGT	RT-PCR of 16S rRNA transcript, positive control
16SrRNA R	GGGGCATTTTGCCGAGTTCC	RT-PCR of 16S rRNA transcript, positive control
RT 2776 F	ATGCACGAGGCACCGGAGCTGACGA	Microarray validation, RT-qPCR
RT 2776 R	GGTCGAGCAGCGTGTCCACGCGGTC	Microarray validation, RT-qPCR
RT 3217 F	GACGGCGGGCGGAGGCACTCACGAC	Microarray and ChIP-chip validation, RT-qPCR
RT 3217 R	TCAACCACTGCTCGCGGCTACGCCC	Microarray and ChIP-chip validation, RT-qPCR
RT 3230 F	TGAGAACTCTTCGGTTCGGCACGGT	Microarray validation, RT-qPCR
RT 3230 R	TGTACGGCACCCCGTCCGGGTCCTC	Microarray validation, RT-qPCR

RT 4921 F	CGCAAGGTGCTCATCGCCAATCGTG	Microarray validation, RT-qPCR
RT 4921 R	GGACCGCCTGCGCGAACTCGGCGTT	Microarray validation, RT-qPCR
RT 5079 F	CCGATGAGAACAAGCCCGTACTGGT	Microarray validation, RT-qPCR
RT 5079 R	ATCTTGCCCTGCCGCTCCTCGCCCT	Microarray validation, RT-qPCR
RT 5085 F	AGCGACGACCGGGAATTGGTCCTGA	Microarray and ChIP-chip validation, RT-qPCR
RT 5085 R	TTCCAGGTGCGCGATATTGCTTTCG	Microarray and ChIP-chip validation, RT-qPCR
RT 5144 F	ACACCGTGCTCTACGAAGTGAGCGA	Microarray validation, RT-qPCR
RT 5144 R	TGTAGTGCTCGCGCACCGTGCTCAT	Microarray validation, RT-qPCR
RT 5222 F	ATGCTTTCCCACACGGCACCACAGC	Microarray validation, RT-qPCR
RT 5222 R	ATGAGGTCCGTGTAGCACAGACCGT	Microarray validation, RT-qPCR
RT 5260 F	GCACTCCCGGCTAGCAGCGGTCGGT	Microarray validation, RT-qPCR
RT 5260 R	TCTTGCCGGAGTCGTCCTTGAACTC	Microarray validation, RT-qPCR
RT 5535 F	CGCAGCACGCCAAGGGCAAGCTGAC	Microarray validation, RT-qPCR
RT 5535 R	TGTCCATGATCTTGTGGATCTTCGT	Microarray validation, RT-qPCR
RT 5877 F	GGGGGAGTGCTTGCCACGATGGACC	Microarray validation, RT-qPCR
RT 5877 R	TACGGGTCCCAATATGTTGATTTCC	Microarray validation, RT-qPCR
RT 5878 F	GACAACGACACAGCCACAGCCACGG	Microarray validation, RT-qPCR
RT 5878 R	CCACGTCGACCTCGATGTCACCGAG	Microarray validation, RT-qPCR
RT 6271 F	TGCTCATCGCCAATCGTGGCGAAAT	Microarray validation, RT-qPCR

RT 6271 R	CGGCGTTCTCCGAGAGGAATCCGTA	Microarray validation, RT-qPCR
RT 6272 F	AGGTGTCCCGCCGGAAACTCATGAA	Microarray validation, RT-qPCR
RT 6272 R	TGACGACGCGGATGTAGTCGGGGTT	Microarray validation, RT-qPCR
RT 6275 F	CACCAGTGAGAACAATTCCGTAGCA	Microarray validation, RT-qPCR
RT 6275 R	AGATGCCGAAGAATCCGGCGTCGAA	Microarray validation, RT-qPCR
4662 F	GACAACGACACAGCCACAGCCACGG	ChIP-chip validation, qPCR
4662 R	CCACGTCGACCTCGATGTCACCGAG	ChIP-chip validation, qPCR
5881 F	GTACACCGGATCCTGCGCAAACTGG	ChIP-chip validation, qPCR
5881 R	CGTCGAGGAGATGAGCGGCGAACTC	ChIP-chip validation, qPCR

Table 2.3 Functional classification of the genes regulated by Crp-associated sites in ChIP-chip assays. Numbers of genes in each category is shown on the right.

Metabolism	**69**
Carbohydrate transport and metabolism	19
Amino acid transport and metabolism	9
Lipid metabolism	7
Nucleotide transport and metabolism	4
Coenzyme metabolism	6
Energy production and conversion	4
Secondary metabolism	20
Transcription and translation	**61**
Transcription regulators	35
Translation, ribosomal and RNA metabolism and transport	17
Signal transduction	4
Posttranslational modification, protein turnover, chaperones	5
Development	**20**
Aerial hyphae differentiation	10
Aerial hyphae septation and sporulation	10
DNA replication, recombination, and repair	**30**
Others	**25**
Cell envelope biogenesis and remodelling	6
Defense mechanisms	13
Oxidoreductases	6
Transferases	3
Hydrolases	6
Hypothetical proteins	**188**

Table 2.4 ChIP-chip targets involved in secondary metabolism and morhogenesis.

Region bound[a]	Enrichment ratio[b]	Strand[c]	Regulated genes[d]	Function of regulated genes	Distance[e]
Secondary Metabolism					
SCO0380-SCO0381	2.38	+	SCO0381	Putative glycosyl transferase for deoxysugar synthesis	-401
SCO0381	2.04	+	SCO0381	Putative glycosyl transferase for deoxysugar synthesis	186
SCO1273	1.96	-	SCO1273	Putative reductase in type II fatty acid synthase	650
SCO1273-SCO1274	2.07	+	SCO1273	Putative reductase in type II fatty acid synthase	-94
SCO3217	3.20	+	SCO3217	CdaR, pathway specific activator for CDA synthesis	847
SCO3229	2.81	-	SCO3229	HmaS, 4-hydroxymandelate synthase	90
			SCO3230	CdaPSI, CDA peptide synthetase I	-359
SCO3230	2.03	+	SCO3230	CdaPSI, CDA peptide synthetase I	282
SCO5084	1.86	+	SCO5084	ActII-3, actinorhodin export	1706
SCO5085-1	2.77	+	SCO5085	ActII-4, actinorhodin biosynthesis pathway regulator	358
SCO5085-2	1.60	+	SCO5085	ActII-4, actinorhodin biosynthesis pathway regulator	38
SCO5221-SCO5222	2.89	+	SCO5222	EizA lyase, sesquiterpene cyclase	-177
SCO5222-1	3.01	+	SCO5222	EizA lyase, sesquiterpene cyclase	896
SCO5222-2	2.88	+	SCO5222	EizA lyase, sesquiterpene cyclase	356
SCO5223	2.47	+	SCO5223	Putative cytochrome P450, sesquiterpene cyclase	736
SCO5881-1	2.44	-	SCO5881	RedZ, undecylprodgiosin biosynthesis pathway specific regulator	130
SCO5881-2	2.16	-	SCO5881	RedZ, undecylprodgiosin biosynthesis pathway specific regulator	590
SCO6271-SCO6272	1.83	+	SCO6271	AccA1, acyl-CoA carboxylase complex, cryptic polyketide biosynthesis	-51
			SCO6272	Scf, secreted FAD-binding protein, cryptic polyketide biosynthesis	-257
SCO6272	2.06	+	SCO6272	Scf, secreted FAD-binding protein, cryptic polyketide biosynthesis	43
SCO6275-SCO6276	1.80	+	SCO6275	CpkA, type I polyketide synthase, cryptic polyketide biosynthesis	-68
			SCO6276	Monooxygenase, cryptic polyketide biosynthesis	-257
SCO6276	2.13	+	SCO6276	Monooxygenase, cryptic polyketide biosynthesis	85
SCO6283	1.81	+	SCO6283	Nucleoside-diphosphate-sugar epimerases, cryptic polyketide biosynthesis	136
			SCO6282	3-oxoacyl-[acyl-carrier protein] reductase, cryptic polyketide biosynthesis	-253
SCO6429	2.27	+	SCO6429	Putative NRPS	139
Aerial Hyphae Development					
SCO0935	2.58	+	SCO0935	Putative secreted protein similar to Sortase E	456
SCO1675	2.11	-	SCO1674	ChpC, long chaplin	-283
			SCO1675	ChpH, short chaplin	190
SCO2717-SCO2718	2.79	+	SCO2717	ChpD, short chaplin	-228
SCO2718	2.73	-	SCO2718	RdlA, rodlet layer formation	70
			SCO2719	RdlB, rodlet layer formation	-290
SCO3548	2.30	-	SCO3548	Putative anti-sigma factor	330
SCO3549	1.54	-	SCO3548	Putative anti-sigma factor	-141
			SCO3549	BldG,putative anti-sigma factor antagonist	310
SCO4075-SCO4076	1.99	+	SCO4075	RagA, ABC transport protein, ATP-binding subunit	-260
SCO4280-SCO4281	2.07	+	SCO4281	CdgB, a diguanylate Cyclase	-305
Sporulation					
SCO1541	2.48	+	SCO1541	SsgB, *Streptomyces* sporulation and cell division protein	30
SCO1792	2.22	-	SCO1793	Similar to sporulation control protein Spo0M from *Bacillus subtilis*	-636
SCO2610-SCO2611	2.18	+	SCO2610	MreC, rod shape-determining protein	-133
SCO2611	1.53	-	SCO2611	MreB, rod shape-determining protein	90
SCO2611-SCO2612	1.82	+	SCO2611	MreB, rod shape-determining protein	-98
SCO3034-SCO3035	2.19	+	SCO3034	WhiB	-43
SCO3557-SCO3558	2.40	+	SCO3557	Putative septum site determining protein	-101
SCO3854	2.47	+	SCO3854	SrgA, septation inhibitor protein	125
SCO4543	1.56	-	SCO4543	WhiJ	790
			SCO4542	WhiJ antagonist protein	-72
SCO7105-SCO7106	2.31	+	SCO7106	WblJ, similar to WhiB	-113

a, genomic context of Crp-associated sequences;
b, Crp-DNA interaction affinity;
c, sense (+) or antisense (-) strand the sequences are found in;
d, genes regulated by Crp-associated sequences;
e, the distances from the center of Crp-associated sequences to the start codons of regulated genes.
Negative values indicate association sites upstream of start codons while positive values indicate association sites within ORFs.
Secondary metabolic cluster genes are highlighted in blue for actinorhodin, pink for undecylprodigiosin, cyan for calcium-dependent antibiotic, and yellow for cryptic polyketide.
The region bound is shown only once in the first column if it regulates more than one gene.

Table 2.5 Functional classification of target genes up- and down-regulated by Crp induction in transcriptome profiling assays. Numbers of genes in each category are shown on the right.

Upregulated genes 360	
Metabolism 187	
Carbohydrate transport and metabolism	20
Lipid metabolism	18
Amino acid transport and metabolism	38
Inorganic ion transport and metabolism	14
Nucleotide transport and metabolism	3
Cofactor and vitamin metabolism	7
Electron transport	6
Secondary metabolism	73
Redox reactions	8
Transcription and translation 58	
Transcription regulators	21
Translation, ribosomal structure and biogenesis	18
Signal transduction	12
Posttranslational modification, protein turnover and chaperones	5
Transcript processing	2
Development 7	
Aerial hyphae differentiation	4
Aerial hyphae septation and sporulation	3
DNA replication, recombination, and repair 1	
Others 29	
Cell envelope biogenesis and remodelling	11
Detoxification and drug resistance	14
Transferases	4
Hypothetical proteins 78	

Downregulated genes 91	
Metabolism 32	
Carbohydrate transport and metabolism	7
Lipid metabolism	1
Amino acid transport and metabolism	10
Inorganic ion transport and metabolism	4
Cofactor and vitamin metabolism	1
Energy production and conversion	2
Secondary metabolism	6
Redox reactions	1
Transcription and translation 12	
Transcription regulators	4
Translation, ribosomal structure and biogenesis	1
Signal transduction	6
Posttranslational modification, protein turnover and chaperones	1
Development 1	
Aerial hyphae septation and sporulation	1
Others 7	
Dehydrogenases	5
Defense mechanisms	2
Hypothetical proteins 39	

Table 2.6 Overlapped targets between ChIP-chip and transcriptome profiling. The bound region is shown only once in the first column if it regulates more than one gene. Secondary metabolic cluster genes are highlighted in blue for actinorhodin, cyan for calcium-dependent antibiotic, and yellow for cryptic polyketide.

Bound region in ChIP-chip	Overlapped targets with transcriptome profiling	Effect of Crp induction on targets transcription	Function
SCO0165-0166	SCO0165	Down	Hypothetical protein
	SCO0166	Down	Putative regulator, similar to polyphosphate kinase
SCO0166	SCO0166	Down	Putative regulator, similar to polyphosphate kinase
SCO1753	SCO1753	Up	Putative integral membrane protein
SCO1905-1906	SCO1905	Up	Hypothetical protein, contains potential RNA-binding domain
SCO2610-2611	SCO2610	Up	MreC, rod shape-determining protein
SCO3217	SCO3217	Up	CdaR, pathway specific activator for CDA synthesis
SCO3229	SCO3229	Up	HmaS, 4-hydroxymandelate synthase
	SCO3230	Up	CdaPSI, CDA peptide synthetase I
SCO3230	SCO3230	Up	CdaPSI, CDA peptide synthetase I
SCO3538	SCO3538	Up	Hypothetical protein
SCO3538-3539	SCO3538	Up	Hypothetical protein
SCO4515-1	SCO4515	Up	Putative membrane protein
SCO4515-2	SCO4515	Up	Putative membrane protein
SCO4561-4562-1	SCO4562	Down	NuoA, NADH dehydrogenase subunit
SCO4561-4562-2	SCO4562	Down	NuoA, NADH dehydrogenase subunit
SCO4878	SCO4878	Up	Putative glycosyltransferase
SCO4921-4922	SCO4921	Up	AccA2, acyl-CoA carboxylase complex A subunit
SCO4978-4979	SCO4979	Up	PckA, phosphoenolpyruvate carboxykinase
SCO5085-1	SCO5085	Up	ActII-4, actinorhodin biosynthesis pathway regulator
SCO5085-2	SCO5085	Up	ActII-4, actinorhodin biosynthesis pathway regulator
SCO5221-5222	SCO5222	Down	EizA, putative lyase
SCO5223	SCO5223	Down	Putative cytochrome P450
SCO5260-5261	SCO5260	Up	Solute binding protein of ABC transporter
	SCO5261	Up	Putative malate oxidoreductase
SCO5439	SCO5439	Up	Hypothetical protein
SCO5544-5545	SCO5544	Up	CvnA1, putative membrane conservon protein
SCO6197-6198	SCO6198	Up	Putative secreted protein
SCO6198	SCO6198	Up	Putative secreted protein
SCO6271-6272	SCO6271	Up	AccA1, acyl-CoA carboxylase complex, cryptic polyketide biosynthesis
	SCO6272	Up	Scf, secreted FAD-binding protein, cryptic polyketide biosynthesis
SCO6272	SCO6272	Up	Scf, secreted FAD-binding protein, cryptic polyketide biosynthesis
SCO6275-6276	SCO6275	Up	CpkA, type I polyketide synthase, cryptic polyketide biosynthesis
SCO6381-6382	SCO6382	Up	Putative secreted protein, similar to β-lactamase class A
SCO6382	SCO6382	Up	Putative secreted protein, similar to β-lactamase class A
SCO6429	SCO6429	Up	Putative NRPS
SCO6510	SCO6510	Up	Hypothetical protein
SCO6591-6592	SCO6591	Up	Putative secreted protein
SCO6592	SCO6592	Up	Putative secreted protein
SCO6593	SCO6593	Up	Hypothetical protein
SCO6773-6774	SCO6774	Up	Hypothetical protein
SCO7227	SCO7227	Down	Putative secreted protein

Table 2.7 Overlap between the regulons of Crp, PhoP, and AfsS based on transcriptome analysis. The lists of genes affected by PhoP and AfsS were obtained from (Lian *et al*, 2008; Rodriguez-Garcia *et al*, 2007)

Crp & AfsS & PhoP	Crp & PhoP	Crp & AfsS
Oxidative stress and iron metabolism		
SCO0379		
Nitrogen metabolism		
SCO2210, SCO5583, SCO5584	SCO2198, SCO4086	
Phosphate metabolism		
SCO4139, SCO4140,		SCO0324, SCO4144, SCO7722
Amino acids metabolism, transcription, and translation		
SCO4293	SCO0992, SCO3615, SCO3906, SCO3907, SCO4654, SCO4655, SCO4701, SCO4702, SCO4703, SCO4704, SCO4706, SCO4708, SCO4711, SCO4716, SCO4718, SCO5776, SCO5777	SCO2910, SCO2911, SCO2912, SCO4164, SCO4165, SCO6094, SCO6095, SCO6096, SCO6097, SCO6098, SCO6099, SCO6100 SCO6101, SCO6102
Cell envelope biosynthesis		
SCO4879, SCO4880, SCO4881	SCO4875	SCO4878
Secondary metabolism		
SCO5877	SCO5878, SCO5888, SCO5890, SCO5893, SCO5896, SCO5897, SCO6273	SCO5072, SCO5074, SCO5079, SCO6281
Respiration		
	SCO0212, SCO0213, SCO0217, SCO0922, SCO0923, SCO3946	
Morphological development		
	SCO5189	
Carbohydrate metabolism		
		SCO0546
Lipid metabolism		
		SCO6691

REFERENCES:

Abramoff MD, Magalhaes PJ, Ram SJ (2004) Image Processing with ImageJ. *Biophotonics International* **11:** 36-42

Agarwal N, Raghunand TR, Bishai WR (2006) Regulation of the expression of *whiB1* in *Mycobacterium tuberculosis*: role of cAMP receptor protein. *Microbiology* **152:** 2749-2756

Ainsa JA, Bird N, Ryding NJ, Findlay KC, Chater KF (2010) The complex *whiJ* locus mediates environmentally sensitive repression of development of *Streptomyces coelicolor* A3(2). *Antonie Van Leeuwenhoek* **98:** 225-236

Ainsa JA, Parry HD, Chater KF (1999) A response regulator-like protein that functions at an intermediate stage of sporulation in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **34:** 607-619

Ainsa JA, Ryding NJ, Hartley N, Findlay KC, Bruton CJ, Chater KF (2001) WhiA, a Protein of Unknown Function Conserved among Gram-Positive Bacteria, Is Essential for Sporulation in*Streptomyces coelicolor* A3(2) *J. Bacteriol.* **182**: 5470-5478

Anderson T, Brian P, Riggle P, Kong R, Champness W (1999) Genetic suppression analysis of non-antibiotic-producing mutants of the *Streptomyces coelicolor absA* locus. *Microbiology* **145 (Pt 9):** 2343-2353

Arias P, Fernandez-Moreno MA, Malpartida F (1999) Characterization of the pathway-specific positive transcriptional regulator for actinorhodin biosynthesis in *Streptomyces coelicolor* A3(2) as a DNA-binding protein. *J Bacteriol* **181:** 6958-6968

Bai G, Schaak DD, McDonough KA (2009) cAMP levels within *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG increase upon infection of macrophages. *FEMS Immunol Med Microbiol* **55:** 68-73

Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D, Bateman A, Brown S, Chandra G, Chen CW, Collins M, Cronin A, Fraser A, Goble A, Hidalgo J, Hornsby T, Howarth S, Huang CH, Kieser T, Larke L, Murphy L, Oliver K, O'Neil S, Rabbinowitsch E, Rajandream MA, Rutherford K, Rutter S, Seeger K, Saunders D, Sharp S, Squares R, Squares S, Taylor K, Warren T, Wietzorrek A, Woodward J, Barrell BG, Parkhill J, Hopwood DA (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417:** 141-147

Berdy J (2005) Bioactive microbial metabolites. *The Journal of antibiotics* **58:** 1-26

Berg OG, von Hippel PH (1988) Selection of DNA binding sites by regulatory proteins. II. The binding specificity of cyclic AMP receptor protein to recognition sites. *J Mol Biol* **200:** 709-723

Bibb MJ (2005) Regulation of secondary metabolism in streptomycetes. *Curr Opin Microbiol* **8:** 208-215

Bibb MJ, Hopwood DA, Freeman RF (1977) Physical and genetical characterisation of a second sex factor, SCP2, for *Streptomyces coelicolor* A3(2) *Molecular & General Genetics* **154:** 155-156

Bibb MJ, Molle V, Buttner MJ (2000) sigma(BldN), an extracytoplasmic function RNA polymerase sigma factor required for aerial mycelium formation in *Streptomyces coelicolor* A3(2). *J Bacteriol* **182:** 4606-4616

Bignell DRD, Warawa JL, Strap JL, Chater KF, Leskiw BK (2000) Study of the *bldG* locus suggests that an anti-anti-sigma factor and an anti-sigma factor may be involved in *Streptomyces coelicolor* antibiotic production and sporulation. *Microbiology* **146**: 2161-2173.

Blanc V, Lagneaux D, Didier P, Gil P, Lacroix P, Crouzet J (1995) Cloning and analysis of structural genes from *Streptomyces pristinaespiralis* encoding enzymes involved in the conversion of pristinamycin IIB to pristinamycin IIA (PIIA): PIIA synthase and NADH:riboflavin 5'-phosphate oxidoreductase. *J Bacteriol* **177:** 5206-5214

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72:** 248-254

Brian P, Riggle PJ, Santos RA, Champness WC (1996) Global negative regulation of *Streptomyces coelicolor* antibiotic synthesis mediated by an *absA*-encoded putative signal transduction system. *J Bacteriol* **178:** 3221-3231

Browning DF, Grainger DC, Busby SJ (2010) Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. *Curr Opin Microbiol* **13:** 773-780

Buck M, Gallegos MT, Studholme DJ, Guo Y, Gralla JD (2000) The bacterial enhancer-dependent σ^{54} (σ^{N}) transcription factor. *J Bacteriol* **182:** 4129-4136

Buck MJ, Lieb JD (2004) ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments. *Genomics* **83:** 349-360

Burger A, Sichler K, Kelemen G, Buttner M, Wohlleben W (2000) Identification and characterization of the mre gene region of *Streptomyces coelicolor* A3(2). *Mol Gen Genet* **263:** 1053-1060

Busby S, Ebright RH (1999) Transcription activation by catabolite activator protein (CAP). *J Mol Biol* **293:** 199-213

Bystrykh LV, Fernandez-Moreno MA, Herrema JK, Malpartida F, Hopwood DA, Dijkhuizen L (1996) Production of actinorhodin-related "blue pigments" by *Streptomyces coelicolor* A3(2). *J Bacteriol* **178:** 2238-2244

Capstick, D.S., Jomaa, A., Hanke, C., Ortega, J. and Elliot, M.A (2011) Dual amyloid domains promote differential functioning of the chaplin proteins during *Streptomyces* aerial morphogenesis. *Proc. Natl. Acad. Sci. USA.* **108**: 9821-26

Carballido-Lopez R (2006) The bacterial actin-like cytoskeleton. *Microbiol Mol Biol Rev* **70:** 888-909

Chakraburtty R, Bibb M (1997) The ppGpp synthetase gene (relA) of *Streptomyces coelicolor* A3(2) plays a conditional role in antibiotic production and morphological differentiation. *J Bacteriol* **179:** 5854-5861

Chang HM, Chen MY, Shieh YT, Bibb MJ, Chen CW (1996) The *cutRS* signal transduction system of *Streptomyces lividans* represses the biosynthesis of the polyketide antibiotic actinorhodin. *Mol Microbiol* **21:** 1075-1085

Chater KF (1975) Construction and phenotypes of double sporulation deficient mutants in *Streptomyces coelicolor* A3(2). *J Gen Microbiol* **87:** 312-325

Chater KF (2001) Regulation of sporulation in *Streptomyces coelicolor* A3(2): a checkpoint multiplex? *Curr Opin Microbiol* **4:** 667-673

Chater KF, Biro S, Lee KJ, Palmer T, Schrempf H (2010) The complex extracellular biology of *Streptomyces*. *FEMS Microbiol Rev* **34:** 171-198

Chater KF, Bruton CJ (1985) Resistance, regulatory and production genes for the antibiotic methylenomycin are clustered. *EMBO J* **4:** 1893-1897

Chater KF, Bruton CJ, Plaskitt KA, Buttner MJ, Mendez C, Helmann JD (1989) The developmental fate of *S. coelicolor* hyphae depends upon a gene product homologous with the motility sigma factor of *B. subtilis*. *Cell* **59:** 133-143

Chong PP, Podmore SM, Kieser HM, Redenbach M, Turgay K, Marahiel M, Hopwood DA, Smith CP (1998) Physical identification of a chromosomal locus encoding

biosynthetic genes for the lipopeptide calcium-dependent antibiotic (CDA) of *Streptomyces coelicolor* A3(2). *Microbiology* **144 (Pt 1):** 193-199

Chouayekh H, Virolle MJ (2002) The polyphosphate kinase plays a negative role in the control of antibiotic production in *Streptomyces lividans*. *Mol Microbiol* **43:** 919-930

Claessen D, Rink R, de Jong W, Siebring J, de Vreugd P, Boersma FG, Dijkhuizen L, Wosten HA (2003) A novel class of secreted hydrophobic proteins is involved in aerial hyphae formation in *Streptomyces coelicolor* by forming amyloid-like fibrils. *Genes Dev* **17:** 1714-1726

Claessen D, Stokroos I, Deelstra HJ, Penninga NA, Bormann C, Salas JA, Dijkhuizen L, Wosten HA (2004) The formation of the rodlet layer of streptomycetes is the result of the interplay between rodlins and chaplins. *Mol Microbiol* **53:** 433-443

Claessen D, Wosten HA, van Keulen G, Faber OG, Alves AM, Meijer WG, Dijkhuizen L (2002) Two novel homologous proteins of *Streptomyces coelicolor* and *Streptomyces lividans* are involved in the formation of the rodlet layer and mediate attachment to a hydrophobic surface. *Mol Microbiol* **44:** 1483-1492

Cohn M, Monod J (1953) Specific inhibition and induction of enzyme biosynthesis. In Adaptation in micro-organisms, Srd symposium of the society for general microbiology. pp. 132-149. Edited by E. F. Gale and R. Davies. Cambridge University Press, London.

D'Alia D, Eggle D, Nieselt K, Hu WS, Breitling R, Takano E (2011) Deletion of the signalling molecule synthase ScbA has pleiotropic effects on secondary metabolite biosynthesis, morphological differentiation and primary metabolism in *Streptomyces coelicolor* A3(2). *Microb Biotechnol* **4:** 239-251

Daniel J (1984) Enzyme III stimulation of cyclic AMP synthesis in an *Escherichia coli* crp mutant. *J Bacteriol* **157:** 940-941

Daniel RA, Errington J (2003) Control of cell morphogenesis in bacteria: two distinct ways to make a rod-shaped cell. *Cell* **113:** 767-776

Davis NK, Chater KF (1992) The *Streptomyces coelicolor whiB* gene encodes a small transcription factor-like protein dispensable for growth but essential for sporulation. *Mol Gen Genet* **232:** 351-358

den Blaauwen JL, Postma PW (1985) Regulation of cyclic AMP synthesis by enzyme IIIGlc of the phosphoenolpyruvate:sugar phosphotransferase system in *crp* strains of *Salmonella typhimurium*. *J Bacteriol* **164:** 477-478

den Hengst CD, Tran NT, Bibb MJ, Chandra G, Leskiw BK, Buttner MJ (2010) Genes essential for morphological development and antibiotic production in *Streptomyces coelicolor* are targets of BldD during vegetative growth. *Mol Microbiol* **78:** 361-379

Derouaux A, Dehareng D, Lecocq E, Halici S, Nothaft H, Giannotta F, Moutzourelis G, Dusart J, Devreese B, Titgemeyer F, Van Beeumen J, Rigali S (2004a) Crp of *Streptomyces coelicolor* is the third transcription factor of the large CRP-FNR superfamily able to bind cAMP. *Biochem Biophys Res Commun* **325:** 983-990

Derouaux A, Halici S, Nothaft H, Neutelings T, Moutzourelis G, Dusart J, Titgemeyer F, Rigali S (2004b) Deletion of a cyclic AMP receptor protein homologue diminishes germination and affects morphological development of *Streptomyces coelicolor*. *J Bacteriol* **186:** 1893-1897

Deutscher J (2008) The mechanisms of carbon catabolite repression in bacteria. *Curr Opin Microbiol* **11:** 87-93

Edwards DH, Errington J (1997) The *Bacillus subtilis* DivIVA protein targets to the division septum and controls the site specificity of cell division. *Mol Microbiol* **24:** 905-915

Elliot M., Ellio M, Damji F, Chater K, Leskiw B (1998) The *bldD* gene of *Streptomyces coelicolor* A3(2): a regulatory gene involved in morphogenesis and antibiotic production. *J. Bacteriol.* **180**:1549–1555

Elliot MA, Bibb MJ, Buttner MJ, Leskiw BK (2001) BldD is a direct regulator of key developmental genes in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **40:** 257-269

Elliot MA, Karoonuthaisiri N, Huang J, Bibb MJ, Cohen SN, Kao CM, Buttner MJ (2003) The chaplins: a family of hydrophobic cell-surface proteins involved in aerial mycelium formation in *Streptomyces coelicolor*. *Genes Dev* **17:** 1727-1740

Elliot MA, M.J B, Nodwell JR (2007) Multicellular Development in *Streptomyces*. In *Myxobacteria:* Multicellularity and Differentiation, Whitworth DE (ed), 24, p 419. Washington, D.C.: ASM Press

Ensign JC (1978) Formation, properties, and germination of actinomycete spores. *Annu Rev Microbiol* **32:** 185-219

Fernandez-Moreno MA, Caballero JL, Hopwood DA, Malpartida F (1991) The *act* cluster contains regulatory and antibiotic export genes, direct targets for translational control by the *bldA* tRNA gene of *Streptomyces*. *Cell* **66:** 769-780

Fic E, Bonarek P, Gorecki A, Kedracka-Krok S, Mikolajczak J, Polit A, Tworzydlo M, Dziedzicka-Wasylewska M, Wasylewski Z (2009) cAMP receptor protein from *Escherichia coli* as a model of signal transduction in proteins--a review. *J Mol Microbiol Biotechnol* **17:** 1-11

Flardh K (2003) Essential role of DivIVA in polar growth and morphogenesis in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **49:** 1523-1536

Flardh K, Buttner MJ (2009) *Streptomyces* morphogenetics: dissecting differentiation in a filamentous bacterium. *Nat Rev Microbiol* **7:** 36-49

Flardh K, Leibovitz E, Buttner MJ, Chater KF (2000) Generation of a non-sporulating strain of *Streptomyces coelicolor* A3(2) by the manipulation of a developmentally controlled *ftsZ* promoter. *Mol Microbiol* **38:** 737-749

Floriano B, Bibb M (1996) *afsR* is a pleiotropic but conditionally required regulatory gene for antibiotic production in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **21:** 385-396

Fuchs EL, Brutinel ED, Jones AK, Fulcher NB, Urbanowski ML, Yahr TL, Wolfgang MC (2010) The *Pseudomonas aeruginosa* Vfr regulator controls global virulence factor expression through cyclic AMP-dependent and -independent mechanisms. *J Bacteriol* **192:** 3553-3564

Galas DJ, Schmitz A (1978) DNAse footprinting: a simple method for the detection of protein-DNA binding specificity. *Nucleic Acids Res* **5:** 3157-3170

Gentry DR, Hernandez VJ, Nguyen LH, Jensen DB, Cashel M (1993) Synthesis of the stationary-phase sigma factor sigma s is positively regulated by ppGpp. *J Bacteriol* **175:** 7982-7989

Gerhardt P, Pankratz HS, Scherrer R (1976) Fine structure of the *Baccilus thuringiensis* spore. *Appl Environ Microbiol* **32:** 438-440

Gottelt M, Kol S, Gomez-Escribano JP, Bibb M, Takano E (2010) Deletion of a regulatory gene within the *cpk* gene cluster reveals novel antibacterial activity in *Streptomyces coelicolor* A3(2). *Microbiology* **156:** 2343-2353

Grainger DC, Hurd D, Harrison M, Holdstock J, Busby SJ (2005) Studies of the distribution of *Escherichia coli* cAMP-receptor protein and RNA polymerase along the *E. coli* chromosome. *Proc Natl Acad Sci U S A* **102:** 17693-17698

Grantcharova N, Lustig U, Flardh K (2005) Dynamics of FtsZ assembly during sporulation in *Streptomyces coelicolor* A3(2). *J Bacteriol* **187:** 3227-3237

Gust B, Challis GL, Fowler K, Kieser T, Chater KF (2003) PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proc Natl Acad Sci U S A* **100:** 1541-1546

Guthrie EP, Flaxman CS, White J, Hodgson DA, Bibb MJ, Chater KF (1998) A response-regulator-like activator of antibiotic synthesis from *Streptomyces coelicolor* A3(2) with an amino-terminal domain that lacks a phosphorylation pocket. *Microbiology* **144 (Pt 3):** 727-738

Gverzdys T (2011) The Development of Protocols to Engineer and Screen *Streptomyces* in High Throughput to Test for the Activation of Cryptic Clusters by the Heterologous Expression of Pleiotropic Regulators. M.Sc. Thesis, Biochemistry and Biomedical Sciences, McMaster University, Open Access Dissertations and Theses

Haiser HJ, Yousef MR, Elliot MA (2009) Cell wall hydrolases affect germination, vegetative growth, and sporulation in *Streptomyces coelicolor*. *J Bacteriol* **191:** 6501-6512

Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166:** 557-580

Hempel AM, Wang SB, Letek M, Gil JA, Flardh K (2008) Assemblies of DivIVA mark sites for hyphal branching and can establish new zones of cell wall growth in *Streptomyces coelicolor*. *J Bacteriol* **190:** 7579-7583

Herrera MC, Daddaoua A, Fernandez-Escamilla A, Ramos JL (2012) Involvement of the global Crp regulator in cyclic AMP-dependent utilization of aromatic amino acids by *Pseudomonas putida*. *J Bacteriol* **194:** 406-412

Hesketh A, Sun J, Bibb M (2001) Induction of ppGpp synthesis in *Streptomyces coelicolor* A3(2) grown under conditions of nutritional sufficiency elicits *actII-ORF4* transcription and actinorhodin biosynthesis. *Mol Microbiol* **39:** 136-144

Hodgson DA (1982) Glucose Repression of Carbon Source Uptake and Metabolism in *Streptomyces coelicolor* A3(2) and its Perturbation in Mutants Resistant to 2-Deoxyglucose *Microbiology* **128:** 2417-2430

Hodgson DA (2000) Primary metabolism and its control in streptomycetes: a most unusual group of bacteria. *Adv Microb Physiol* **42:** 47-238

Hopwood DA (1999) Forty years of genetics with *Streptomyces*: from *in vivo* through *in vitro* to *in silico*. *Microbiology* **145 (Pt 9):** 2183-2202

Hopwood DA, Bibb MJ, Chater KF, Kieser T, Bruton CJ, Kieser HM, Lydiate DJ, Smith CP, Ward JM, Schrempf H (1985) Genetic manipulation of *Streptomyces* – A laboratory manual., Norwich, U.K: The John Innes Foundation.

Horinouchi S, Beppu T (1992) Regulation of secondary metabolism and cell differentiation in *Streptomyces*: A-factor as a microbial hormone and the AfsR protein as a component of a two-component regulatory system. *Gene* **115:** 167-172

Huang J, Lih CJ, Pan KH, Cohen SN (2001) Global analysis of growth phase responsive gene expression and regulation of antibiotic biosynthetic pathways in *Streptomyces coelicolor* using DNA microarrays. *Genes Dev* **15:** 3183-3192

Huang J, Shi J, Molle V, Sohlberg B, Weaver D, Bibb MJ, Karoonuthaisiri N, Lih CJ, Kao CM, Buttner MJ, Cohen SN (2005) Cross-regulation among disparate antibiotic biosynthetic pathways of *Streptomyces coelicolor*. *Mol Microbiol* **58:** 1276-1287

Ishizuka H, Horinouchi S, Kieser HM, Hopwood DA, Beppu T (1992) A putative two-component regulatory system involved in secondary metabolism in *Streptomyces* spp. *J Bacteriol* **174:** 7585-7594

Jakimowicz P, Cheesman MR, Bishai WR, Chater KF, Thomson AJ, Buttner MJ (2005) Evidence that the *Streptomyces* developmental protein WhiD, a member of the WhiB family, binds a [4Fe-4S] cluster. *J Biol Chem* **280:** 8309-8315

Jensen LJ, Kuhn M, Stark M, Chaffron S, Creevey C, Muller J, Doerks T, Julien P, Roth A, Simonovic M, Bork P, von Mering C (2009) STRING 8--a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res* **37:** D412-416

Jyothikumar V, Tilley EJ, Wali R, Herron PR (2008) Time-lapse microscopy of *Streptomyces coelicolor* growth and sporulation. *Appl Environ Microbiol* **74:** 6774-6781

Kahn JD (1999) Methods for analyzing DNA bending. *Methods Mol Biol* **94:** 109-123

Kaiser BK, Stoddard BL. (2011) DNA recognition and transcriptional regulation by the WhiA sporulation factor. *Scientific Reports*. Nature Publishing Group, Vol. 1.

Kallifidas D, Thomas D, Doughty P, Paget MS (2010) The σ^R regulon of *Streptomyces coelicolor* A3(2) reveals a key role in protein quality control during disulphide stress. *Microbiology* **156:** 1661-1672

Kamionka A, Parche S, Nothaft H, Siepelmeyer J, Jahreis K, Titgemeyer F (2002) The phosphotransferase system of *Streptomyces coelicolor*. *Eur J Biochem* **269:** 2143-2150

Kampfer P (2006) The Family *Streptomycetaceae*, Part I: Taxonomy. In *The Prokaryotes*, Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E (eds), Vol. 3, 1.1.7, pp 538-604. Springer

Kang SG, Jin W, Bibb M, Lee KJ (1998) Actinorhodin and undecylprodigiosin production in wild-type and *relA* mutant strains of *Streptomyces coelicolor* A3(2) grown in continuous culture. *FEMS Microbiol Lett* **168:** 221-226

Kato JY, Miyahisa I, Mashiko M, Ohnishi Y, Horinouchi S (2004) A single target is sufficient to account for the biological effects of the A-factor receptor protein of *Streptomyces griseus*. *J Bacteriol* **186:** 2206-2211

Kelemen GH, Brian P, Flardh K, Chamberlin L, Chater KF, Buttner MJ (1998) Developmental regulation of transcription of *whiE*, a locus specifying the polyketide spore pigment in *Streptomyces coelicolor* A3 (2). *J Bacteriol* **180:** 2515-2521

Khrapunov S, Brenowitz M, Rice PA, Catalano CE (2006) Binding then bending: a mechanism for wrapping DNA. *Proc Natl Acad Sci U S A* **103:** 19217-19218

Kieser T, M. Bibb, M. Buttner, K. Chater, and D. Hopwood (2000) Practical *Streptomyces* genetics., Norwich, United Kingdom: John Innes Foundation.

Kim ES, Hong HJ, Choi CY, Cohen SN (2001) Modulation of actinorhodin biosynthesis in *Streptomyces lividans* by glucose repression of *afsR2* gene transcription. *J Bacteriol* **183:** 2198-2203

Kim HJ, Calcutt MJ, Schmidt FJ, Chater KF (2000) Partitioning of the linear chromosome during sporulation of *Streptomyces coelicolor* A3(2) involves an *oriC*-linked *parAB* locus. *J Bacteriol* **182:** 1313-1320

Kim HJ, Kim TH, Kim Y, Lee HS (2004) Identification and characterization of *glxR*, a gene involved in regulation of glyoxylate bypass in *Corynebacterium glutamicum*. *J Bacteriol* **186:** 3453-3460

Kleinschnitz EM, Heichlinger A, Schirner K, Winkler J, Latus A, Maldener I, Wohlleben W, Muth G (2011) Proteins encoded by the *mre* gene cluster in *Streptomyces coelicolor* A3(2) cooperate in spore wall synthesis. *Mol Microbiol* **79:** 1367-1379

Kodani S, Hudson ME, Durrant MC, Buttner MJ, Nodwell JR, Willey JM (2004) The SapB morphogen is a lantibiotic-like peptide derived from the product of the developmental gene *ramS* in *Streptomyces coelicolor*. *Proc Natl Acad Sci U S A* **101:** 11448-11453

Kolb A, Busby S, Buc H, Garges S, Adhya S (1993) Transcriptional regulation by cAMP and its receptor protein. *Annu Rev Biochem* **62:** 749-795

Kornberg A, Rao NN, Ault-Riche D (1999) Inorganic polyphosphate: a molecule of many functions. *Annu Rev Biochem* **68:** 89-125

Korner H, Sofia HJ, Zumft WG (2003) Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: exploiting the metabolic spectrum by controlling alternative gene programs. *FEMS Microbiol Rev* **27:** 559-592

Kvint K, Nachin L, Diez A, Nystrom T (2003) The bacterial universal stress protein: function and regulation. *Curr Opin Microbiol* **6:** 140-145

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227:** 680-685

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23:** 2947-2948

Lawlor EJ, Baylis HA, Chater KF (1987) Pleiotropic morphological and antibiotic deficiencies result from mutations in a gene encoding a tRNA-like product in *Streptomyces coelicolor A3*(2). *Genes Dev* **1:** 1305-1310

Leblanc B, Moss T (2009) DNase I footprinting. *Methods Mol Biol* **543:** 37-47

Lee EJ, Karoonuthaisiri N, Kim HS, Park JH, Cha CJ, Kao CM, Roe JH (2005) A master regulator σ^B governs osmotic and oxidative response as well as differentiation via a network of sigma factors in *Streptomyces coelicolor*. *Mol Microbiol* **57:** 1252-1264

Lee HN, Im JH, Lee MJ, Lee SY, Kim ES (2009) A putative secreted solute binding protein, SCO6569 is a possible AfsR2-dependent down-regulator of actinorhodin biosynthesis in *Streptomyces coelicolor*. *Process Biochemistry* **44:** 373-377

Leskiw BK, Lawlor EJ, Fernandez-Abalos JM, Chater KF (1991) TTA codons in some genes prevent their expression in a class of developmental, antibiotic-negative, *Streptomyces* mutants. *Proc Natl Acad Sci U S A* **88:** 2461-2465

Li YQ, Chen PL, Chen SF, Wu D, Zheng J (2004) A pair of two-component regulatory genes *ecrA1/A2* in *Streptomyces coelicolor*. *J Zhejiang Univ Sci* **5:** 173-179

Lian W, Jayapal KP, Charaniya S, Mehra S, Glod F, Kyung YS, Sherman DH, Hu WS (2008) Genome-wide transcriptome analysis reveals that a pleiotropic antibiotic regulator,

AfsS, modulates nutritional stress response in *Streptomyces coelicolor* A3(2). *BMC genomics* **9:** 56

Ma H, Kendall K (1994) Cloning and analysis of a gene cluster from *Streptomyces coelicolor* that causes accelerated aerial mycelium formation in *Streptomyces lividans J. Bacteriol.* **176**: 3800-3811

MacNeil DJ, Gewain KM, Ruby CL, Dezeny G, Gibbons PH, MacNeil T (1992) Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene* **111:** 61-68

Majka J, Speck C (2007) Analysis of protein-DNA interactions using surface plasmon resonance. *Adv Biochem Eng Biotechnol* **104:** 13-36

Malpartida F, Niemi J, Navarrete R, Hopwood DA (1990) Cloning and expression in a heterologous host of the complete set of genes for biosynthesis of the *Streptomyces coelicolor* antibiotic undecylprodigiosin. *Gene* **93:** 91-99

Margot P, Wahlen M, Gholamhoseinian A, Piggot P, Karamata D (1998) The *lytE* gene of *Bacillus subtilis* 168 encodes a cell wall hydrolase. *J Bacteriol* **180:** 749-752

Marshall CJ (1995) Specificity of receptor tyrosine kinase signaling: transient versus sustained extracellular signal-regulated kinase activation. *Cell* **80:** 179-185

Mazza P, Noens EE, Schirner K, Grantcharova N, Mommaas AM, Koerten HK, Muth G, Flardh K, van Wezel GP, Wohlleben W (2006) MreB of *Streptomyces coelicolor* is not essential for vegetative growth but is required for the integrity of aerial hyphae and spores. *Mol Microbiol* **60:** 838-852

McCormick JR, Su EP, Driks A, Losick R (1994) Growthand viability of *Streptomyces coelicolor* mutant for the cell division gene *ftsZ. Mol. Microbiol.* **14**: 243–254

McCormick JR, Losick R (1996) Cell division gene *ftsQ* is required for efficient sporulation but not growth and viability in *Streptomyces coelicolor* A3(2). *J Bacteriol* **178:** 5295-5301

McKenzie NL, Nodwell JR (2007) Phosphorylated AbsA2 negatively regulates antibiotic production in *Streptomyces coelicolor* through interactions with pathway-specific regulatory gene promoters. *J Bacteriol* **189:** 5284-5292

Meng X, Brodsky MH, Wolfe SA (2005) A bacterial one-hybrid system for determining the DNA-binding specificity of transcription factors. *Nat Biotechnol* **23:** 988-994

Molle V, Buttner MJ (2000) Different alleles of the response regulator gene *bldM* arrest *Streptomyces coelicolor* development at distinct stages. *Mol Microbiol* **36:** 1265-1278

Molle V, Palframan WJ, Findlay KC, Buttner MJ (2000) WhiD and WhiB, homologous proteins required for different stages of sporulation in *Streptomyces coelicolor* A3(2). *J Bacteriol* **182:** 1286-1295

Mukamolova GV, Turapov OA, Young DI, Kaprelyants AS, Kell DB, Young M (2002) A family of autocrine growth factors in *Mycobacterium tuberculosis*. *Mol Microbiol* **46:** 623-635

Musso RE, Di Lauro R, Adhya S, de Crombrugghe B (1977) Dual control for transcription of the galactose operon by cyclic AMP and its receptor protein at two interspersed promoters. *Cell* **12:** 847-854

Nguyen KT, Willey JM, Nguyen LD, Nguyen LT, Viollier PH, Thompson CJ (2002) A central regulator of morphological differentiation in the multicellular bacterium *Streptomyces coelicolor*. *Mol. Microbiol.* **46**:1223–1238.
Nguyen KT, Tenor J, Stettler H, Nguyen LT, Nguyen LD, Thompson CJ (2003) Colonial differentiation in *Streptomyces coelicolor* depends on translation of a specific codon within the *adpA* gene. *J Bacteriol* **185:** 7291-7296

Nieselt K, Battke F, Herbig A, Bruheim P, Wentzel A, Jakobsen OM, Sletta H, Alam MT, Merlo ME, Moore J, Omara WA, Morrissey ER, Juarez-Hermosillo MA, Rodriguez-Garcia A, Nentwich M, Thomas L, Iqbal M, Legaie R, Gaze WH, Challis GL, Jansen RC, Dijkhuizen L, Rand DA, Wild DL, Bonin M, Reuther J, Wohlleben W, Smith MC, Burroughs NJ, Martin JF, Hodgson DA, Takano E, Breitling R, Ellingsen TE, Wellington EM (2010) The dynamic architecture of the metabolic switch in *Streptomyces coelicolor*. *BMC genomics* **11:** 10

Nodwell JR, McGovern K, Losick R (1996) An oligopeptide permease responsible for the import of an extracellular signal governing aerial mycelium formation in *Streptomyces coelicolor*. *Mol Microbiol* **22:** 881-893

Nodwell JR, Yang M, Kuo D, Losick R (1999) Extracellular complementation and the identification of additional genes involved in aerial mycelium formation in *Streptomyces coelicolor*. *Genetics* **151:** 569-584

Noens EE, Mersinias V, Traag BA, Smith CP, Koerten HK, van Wezel GP (2005) SsgA-like proteins determine the fate of peptidoglycan during sporulation of *Streptomyces coelicolor*. *Mol Microbiol* **58:** 929-944

Nothaft H, Dresel D, Willimek A, Mahr K, Niederweis M, Titgemeyer F (2003) The phosphotransferase system of *Streptomyces coelicolor* is biased for N-acetylglucosamine metabolism. *J Bacteriol* **185:** 7019-7023

O'Connor TJ, Kanellis P, Nodwell JR (2002) The *ramC* gene is required for morphogenesis in *Streptomyces coelicolor* and expressed in a cell type-specific manner under the direct control of RamR. *Mol. Microbiol.* **45**:45–57.

O'Rourke S, Wietzorrek A, Fowler K, Corre C, Challis GL, Chater KF (2009) Extracellular signalling, translational control, two repressors and an activator all contribute to the regulation of methylenomycin production in *Streptomyces coelicolor*. *Mol Microbiol* **71:** 763-778

O'Toole R, Williams HD (2003) Universal stress proteins and *Mycobacterium tuberculosis*. *Res Microbiol* **154:** 387-392

Ohnishi Y, Yamazaki H, Kato JY, Tomono A, Horinouchi S (2005) AdpA, a central transcriptional regulator in the A-factor regulatory cascade that leads to morphological development and secondary metabolism in *Streptomyces griseus*. *Biosci Biotechnol Biochem* **69:** 431-439

Onaka H, Nakagawa T, Horinouchi S (1998) Involvement of two A-factor receptor homologues in *Streptomyces coelicolor* A3(2) in the regulation of secondary metabolism and morphogenesis. *Mol Microbiol* **28:** 743-753

Parashar A, Colvin KR, Bignell DR, Leskiw BK (2009) BldG and SCO3548 interact antagonistically to control key developmental processes in *Streptomyces coelicolor*. *J Bacteriol* **191:** 2541-2550

Park HD, Guinn KM, Harrell MI, Liao R, Voskuil MI, Tompa M, Schoolnik GK, Sherman DR (2003) Rv3133c/DosR is a transcription factor that mediates the hypoxic response of *Mycobacterium tuberculosis*. *Mol Microbiol* **48:** 833-843

Pawlik K, Kotowska M, Chater KF, Kuczek K, Takano E (2007) A cryptic type I polyketide synthase (*cpk*) gene cluster in *Streptomyces coelicolor* A3(2). *Arch Microbiol* **187:** 87-99

Petersen S, Young GM (2002) Essential role for cyclic AMP and its receptor protein in *Yersinia enterocolitica* virulence. *Infect Immun* **70:** 3665-3672

Petrickova K, Petricek M (2003) Eukaryotic-type protein kinases in *Streptomyces coelicolor*: variations on a common theme. *Microbiology* **149:** 1609-1621

Piette A, Derouaux A, Gerkens P, Noens EE, Mazzucchelli G, Vion S, Koerten HK, Titgemeyer F, De Pauw E, Leprince P, van Wezel GP, Galleni M, Rigali S (2005) From dormant to germinating spores of *Streptomyces coelicolor* A3(2): new perspectives from the *crp* null mutant. *J Proteome Res* **4:** 1699-1708

Pluskal T, Castillo S, Villar-Briones A, Oresic M (2010) MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **11:** 395

Potuckova L, Kelemen GH, Findlay KC, Lonetto MA, Buttner MJ, Kormanec J (1995) A new RNA polymerase sigma factor, sigma F, is required for the late stages of morphological differentiation in *Streptomyces spp*. *Mol Microbiol* **17:** 37-48

Price B, Adamidis T, Kong R, Champness W (1999) A *Streptomyces coelicolor* antibiotic regulatory gene, *absB*, encodes an RNase III homolog. *J Bacteriol* **181:** 6142-6151

Redenbach M, Kieser HM, Denapaite D, Eichner A, Cullum J, Kinashi H, Hopwood DA (1996) A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb *Streptomyces coelicolor* A3(2) chromosome. *Mol Microbiol* **21:** 77-96

Rickman L, Scott C, Hunt DM, Hutchinson T, Menendez MC, Whalan R, Hinds J, Colston MJ, Green J, Buxton RS (2005) A member of the cAMP receptor protein family of transcription regulators in *Mycobacterium tuberculosis* is required for virulence in mice and controls transcription of the *rpfA* gene coding for a resuscitation promoting factor. *Mol Microbiol* **56:** 1274-1286

Rigali S, Nothaft H, Noens EE, Schlicht M, Colson S, Muller M, Joris B, Koerten HK, Hopwood DA, Titgemeyer F, van Wezel GP (2006) The sugar phosphotransferase system of *Streptomyces coelicolor* is regulated by the GntR-family regulator DasR and links N-acetylglucosamine metabolism to the control of development. *Mol Microbiol* **61:** 1237-1251

Rigali S, Titgemeyer F, Barends S, Mulder S, Thomae AW, Hopwood DA, van Wezel GP (2008) Feast or famine: the global regulator DasR links nutrient stress to antibiotic production by *Streptomyces*. *EMBO Rep* **9:** 670-675

Rodriguez-Garcia A, Barreiro C, Santos-Beneit F, Sola-Landa A, Martin JF (2007) Genome-wide transcriptomic and proteomic analysis of the primary response to phosphate limitation in *Streptomyces coelicolor* M145 and in a *ΔphoP* mutant. *Proteomics* **7:** 2410-2429

Ruban-Osmialowska B, Jakimowicz D, Smulczyk-Krawczyszyn A, Chater KF, Zakrzewska-Czerwinska J (2006) Replisome localization in vegetative and aerial hyphae of *Streptomyces coelicolor*. *J Bacteriol* **188:** 7311-7316

Rudd BA, Hopwood DA (1980) A pigmented mycelial antibiotic in *Streptomyces coelicolor*: control by a chromosomal gene cluster. *J Gen Microbiol* **119:** 333-340

Ruiz B, Chavez A, Forero A, Garcia-Huante Y, Romero A, Sanchez M, Rocha D, Sanchez B, Rodriguez-Sanoja R, Sanchez S, Langley E (2010) Production of microbial secondary metabolites: regulation by the carbon source. *Crit Rev Microbiol* **36:** 146-167

Ryding NJ, Anderson TB, Champness WC (2002) Regulation of the *Streptomyces coelicolor* calcium-dependent antibiotic by *absA*, encoding a cluster-linked two-component system. *J Bacteriol* **184:** 794-805

Ryding NJ, Kelemen GH, Whatling CA, Flardh K, Buttner MJ, Chater KF (1998) A developmentally regulated gene encoding a repressor-like protein is essential for sporulation in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **29:** 343-357

Ryu YG, Butler MJ, Chater KF, Lee KJ (2006) Engineering of primary carbohydrate metabolism for increased production of actinorhodin in *Streptomyces coelicolor*. *Appl Environ Microbiol* **72:** 7132-7139

Santos-Beneit F, Rodriguez-Garcia A, Martin JF (2011) Complex transcriptional control of the antibiotic regulator *afsS* in Streptomyces: PhoP and AfsR are overlapping, competitive activators. *J Bacteriol* **193:** 2242-2251

Santos-Beneit F, Rodriguez-Garcia A, Sola-Landa A, Martin JF (2009) Cross-talk between two global regulators in Streptomyces: PhoP and AfsR interact in the control of *afsS, pstS* and *phoRP* transcription. *Mol Microbiol* **72:** 53-68

Sasse-Dwight S, Gralla JD (1990) Role of eukaryotic-type functional domains found in the prokaryotic enhancer receptor factor σ54. *Cell* **62:** 945-954

Schmitt-John T, Engels JW (1992) Promoter constructions for efficient secretion expression in *Streptomyces lividans*. *Applied microbiology and biotechnology* **36:** 493-498

Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative C(T) method. *Nature protocols* **3:** 1101-1108

Schwedock J, McCormick JR, Angert ER, Nodwell JR, Losick R (1997) Assembly of the cell division protein FtsZ into ladder-like structures in the aerial hyphae of *Streptomyces coelicolor*. *Mol Microbiol* **25:** 847-858

Scott JJ, Oh DC, Yuceer MC, Klepzig KD, Clardy J, Currie CR (2008) Bacterial protection of beetle-fungus mutualism. *Science* **322:** 63

Setlow P (2003) Spore germination. *Curr Opin Microbiol* **6:** 550-556

Sevcikova B, Rezuchova B, Homerova D, Kormanec J (2010) The anti-anti-sigma factor BldG is involved in activation of the stress response sigma factor σ^H in *Streptomyces coelicolor* A3(2). *J Bacteriol* **192:** 5674-5681

Sharples GP, Williams ST (1976) Fine Structure of Spore Germination in Actinomycetes. *Journal of General Microbiology* **96:** 323-332

Sola-Landa A, Moura RS, Martin JF (2003) The two-component PhoR-PhoP system controls both primary metabolism and secondary metabolite biosynthesis in *Streptomyces lividans*. *Proc Natl Acad Sci U S A* **100:** 6133-6138

Soliveri JA, Gomez J, Bishai WR, Chater KF (2000) Multiple paralogous genes related to the *Streptomyces coelicolor* developmental regulatory gene *whiB* are present in *Streptomyces* and other actinomycetes. *Microbiology* **146 (Pt 2):** 333-343

Stackebrandt E, Woese CR (1981) Towards a phylogeny of the actinomycetes and related organisms. *Current Microbiology* **5:** 197-202

Stockley PG, Persson B (2009) Surface plasmon resonance assays of DNA-protein interactions. *Methods Mol Biol* **543:** 653-669

Stothard P, Wishart DS (2005) Circular genome visualization and exploration using CGView. *Bioinformatics* **21:** 537-539

Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warrener P, Hickey MJ, Brinkman FS, Hufnagle WO, Kowalik DJ, Lagrou M, Garber RL, Goltry L, Tolentino E, Westbrock-Wadman S, Yuan Y, Brody LL, Coulter SN, Folger KR, Kas A, Larbig K, Lim R, Smith K, Spencer D, Wong GK, Wu Z, Paulsen IT, Reizer J, Saier MH, Hancock RE, Lory S, Olson MV (2000) Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature* **406:** 959-964

Studier FW, Moffatt BA (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **189:** 113-130

Stuttard C, Dwyer M (1981) A new temperate phage of *Streptomyces venezuelae*: morphology, DNA molecular weight and host range of SV2. *Can J Microbiol* **27:** 496-499

Susstrunk U, Pidoux J, Taubert S, Ullmann A, Thompson CJ (1998) Pleiotropic effects of cAMP on germination, antibiotic biosynthesis and morphological development in *Streptomyces coelicolor*. *Mol Microbiol* **30:** 33-46

Takahashi H, Inada T, Postma P, Aiba H (1998) CRP down-regulates adenylate cyclase activity by reducing the level of phosphorylated IIAGlc, the glucose-speci®c phosphotransferase protein,
in *Escherichia coli*. *Mol Gen Genet.* **259**: 317-326

Takano E, Chakraburtty R, Nihira T, Yamada Y, Bibb MJ (2001) A complex role for the gamma-butyrolactone SCB1 in regulating antibiotic production in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **41:** 1015-1028

Takano E, Gramajo HC, Strauch E, Andres N, White J, Bibb MJ (1992) Transcriptional regulation of the *redD* transcriptional activator gene accounts for growth-phase-dependent production of the antibiotic undecylprodigiosin in *Streptomyces coelicolor* A3(2). *Mol Microbiol* **6:** 2797-2804

Takano E, Kinoshita H, Mersinias V, Bucca G, Hotchkiss G, Nihira T, Smith CP, Bibb M, Wohlleben W, Chater K (2005) A bacterial hormone (the SCB1) directly controls the expression of a pathway-specific regulatory gene in the cryptic type I polyketide biosynthetic gene cluster of *Streptomyces coelicolor*. *Mol Microbiol* **56:** 465-479

Takano E, Tao M, Long F, Bibb MJ, Wang L, Li W, Buttner MJ, Deng ZX, Chater KF (2003) A rare leucine codon in *adpA* is implicated in the morphological defect of *bldA* mutants of *Streptomyces coelicolor*. *Mol Microbiol* **50:** 475-486

Tomljenovic-Berube AM, Mulder DT, Whiteside MD, Brinkman FS, Coombes BK (2010) Identification of the regulatory logic controlling *Salmonella* pathoadaptation by the SsrA-SsrB two-component system. *PLoS Genet* **6:** e1000875

Toyoda K, Teramoto H, Inui M, Yukawa H (2011) Genome-wide identification of in vivo binding sites of GlxR, a cyclic AMP receptor protein-type regulator in *Corynebacterium glutamicum*. *J Bacteriol* **193:** 4123-4133

Tsuzuki T, Okada H, Cho H, Tsuji S, Nishigaki A, Yasuda K, Kanzaki H (2012) Hypoxic stress simultaneously stimulates vascular endothelial growth factor via hypoxia-inducible factor-1α and inhibits stromal cell-derived factor-1 in human endometrial stromal cells. *Hum Reprod* **27:** 523-530

Uguru GC, Stephens KE, Stead JA, Towle JE, Baumberg S, McDowall KJ (2005) Transcriptional activation of the pathway-specific regulator of the actinorhodin biosynthetic genes in *Streptomyces coelicolor*. *Mol Microbiol* **58:** 131-150

Ullmann A, Danchin A (1983) Role of cyclic AMP in bacteria. *Advances in Cyclic Nucleotide Research* **15:** 1-53

Umeyama T, Horinouchi S (2001) Autophosphorylation of a bacterial serine/threonine kinase, AfsK, is inhibited by KbpA, an AfsK-binding protein. *J Bacteriol* **183:** 5506-5512

Vogtli M, Chang PC, Cohen SN (1994) *afsR2*: a previously undetected gene encoding a 63-amino-acid protein that stimulates antibiotic production in *Streptomyces lividans*. *Mol Microbiol* **14:** 643-653

Waksman SA (1967) The ACTINOMYCETES: a summary of current knowledge, New York: The Ronald Press Company.

Waksman SA, Woodruff HB (1940) The Soil as a Source of Microorganisms Antagonistic to Disease-Producing Bacteria. *J Bacteriol* **40:** 581-600

Wang L, Yu Y, He X, Zhou X, Deng Z, Chater KF, Tao M (2007) Role of an FtsK-like protein in genetic stability in *Streptomyces coelicolor* A3(2). *J Bacteriol* **189:** 2310-2318

West SE, Sample AK, Runyen-Janecky LJ (1994) The *vfr* gene product, required for *Pseudomonas aeruginosa* exotoxin A and protease production, belongs to the cyclic AMP receptor protein family. *J Bacteriol* **176:** 7532-7542

White J, Bibb M (1997) *bldA* dependence of undecylprodigiosin production in *Streptomyces coelicolor* A3(2) involves a pathway-specific regulatory cascade. *J Bacteriol* **179:** 627-633

Wietzorrek A, Bibb M (1997) A novel family of proteins that regulates antibiotic production in streptomycetes appears to contain an OmpR-like DNA-binding fold. *Mol Microbiol* **25:** 1181-1184

Willemse J, Borst JW, de Waal E, Bisseling T, van Wezel GP (2011) Positive control of cell division: FtsZ is recruited by SsgB during sporulation of *Streptomyces*. *Genes Dev* **25:** 89-99

Willey J, Santamaria R, Guijarro J, Geistlich M, Losick R (1991) Extracellular complementation of a developmental mutation implicates a small sporulation protein in aerial mycelium formation by *Streptomyces coelicolor*. *Cell* **65:** 641-650

Willey J, Schwedock J, Losick R (1993) Multiple extracellular signals govern the production of a morphogenetic protein involved in aerial mycelium formation by *Streptomyces coelicolor*. *Genes Dev* **7:** 895-903

Wright GD (2007) The antibiotic resistome: the nexus of chemical and genetic diversity. *Nat Rev Microbiol* **5:** 175-186

Wright LF, Hopwood DA (1976a) Actinorhodin is a chromosomally-determined antibiotic in *Streptomyces coelicolar* A3(2). *J Gen Microbiol* **96:** 289-297

Wright LF, Hopwood DA (1976b) Identification of the antibiotic determined by the SCP1 plasmid of *Streptomyces coelicolor* A3(2). *J Gen Microbiol* **95:** 96-106

Yepes A, Rico S, Rodriguez-Garcia A, Santamaria RI, Diaz M (2011) Novel two-component systems implied in antibiotic production in *Streptomyces coelicolor*. *PLoS One* **6:** e19980

Zhang ZR, Palfrey D, Nagel DA, Lambert PA, Jessop RA, Santos AF, Hine AV (2003) Fluorescent microplate-based analysis of protein-DNA interactions. I: Immobilized protein. *Biotechniques* **35:** 980-982, 984, 986

PART II

CHAPTER 1:
GENERAL INTRODUCTION

1.1 SECRETION PATHWAYS IN GRAM NEGATIVE BACTERIA

Gram negative bacteria are surrounded by double layers of phospholipid membrane, where the cytoplasmic (inner) membrane is separated from the outer membrane by an intervening periplasmic space (Costerton *et al*, 1974). The double membranes provide an impermeable barrier that restricts the movement of compounds both into and out of the cell. Consequently, the export of proteins from the cytoplasm to either the periplasm or the outside of the cell requires dedicated translocation machineries. In bacteria, most co-translated proteins are integral membrane proteins (Ulbrandt *et al*, 1997), while the vast majority of secreted proteins are translocated across the cytoplasmic membrane posttranslationally via the Sec translocon (du Plessis *et al*, 2011). The core of the Sec translocon comprises a heterotrimeric complex SecYEG, which forms a transmembrane protein-conducting channel and acts in conjunction with a variety of ancillary proteins such as the cytoplasmic molecular chaperone SecB and the ATPase SecA (Bechtluft *et al*, 2007; Brundage *et al*, 1990; Hartl *et al*, 1990). Substrates destined for secretion via the Sec pathway are recognized by SecB, which binds the newly synthesized peptide to maintain it in an unfolded translocation competent state, and SecB subsequently targets the peptide to the Sec translocon through interaction with SecA. SecA, in turn, drives the translocation of these unfolded peptides across the cytoplasmic membrane in an ATP hydrolysis-dependent manner (du Plessis *et al*, 2011; Gold *et al*, 2007). Folded proteins are not translocated by the Sec machinery; instead, their export is mediated by the twin arginine translocase (TAT), which consists of three integral membrane proteins TatA, B, and C (Muller & Klosgen, 2005). Proteins destined for secretion by either the Sec secretion apparatus or the TAT-system share an N-terminal signal sequence that consists of a positively charged N-terminus, a hydrophobic core and a hydrophilic C-terminus, and the signal peptide of TAT-specific substrates also encompasses two consecutive invariant Arg residues between its N-terminal region and the hydrophobic core (Hegde & Bernstein, 2006; Natale *et al*, 2008). These residues interact specifically with the membrane bound TatBC complex. Once the substrates are targeted to the translocon, TatA is recruited to the complex and, in the presence of the proton motive force, induces a major conformational change in TatBC that pulls the substrates across the cytoplasmic membrane (Natale *et al*, 2008).

Due to the complexity of the Gram negative bacterial cell envelope, export into the periplasm is just the first step in protein secretion, with other mechanisms being required to help the substrates traverse the outer membrane. This process has been the subject of intense research for over five decades, with considerable insight coming from investigations into bacterial pathogenesis (Fronzes *et al*, 2009a). Successful pathogens secrete a wide variety of proteinaceous virulence factors, including adhesins, toxins and enzymes, and these must traverse not only the cytoplasmic membrane but also the outer

membrane in order to enter the extracellular environment or host cells. What has been learned from studying pathogens can be applied more broadly to Gram negative bacteria as a whole, where the demand to adapt to different scenarios has impelled the emergence of six protein secretion pathways, type I-VI secretion systems (Abdallah *et al*, 2007; Fronzes *et al*, 2009a), which are described below (Table 1.1).

Type I secretion systems (T1SSs) consist of a pore-forming outer membrane protein TolC, a membrane fusion protein MFP and an inner membrane ATP-binding cassette (ABC) transporter (Holland *et al*, 2005). The ABC transporter recognizes the signal sequences of substrates and supplies energy for substrate translocation through the transenvelope channel formed by TolC and MFP. The substrate secretion signal is an uncleaved C-terminal sequence of 46-60aa long and it contains no characteristic primary sequence motif (Holland *et al*, 2005). T1SSs are involved in the secretion of a wide range of cytotoxins and hydrolytic enzymes, and are exemplified by the α-hemolysin export system of *Escherichia coli* (Gentschev *et al*, 2002).

In the case of type II secretion systems (T2SSs), they comprise a multiprotein complex (usually 12-16 components) that spans both membranes (Filloux, 2004). The outer membrane components constitute a secretion pore, while the inner membrane components share similarity with type IV pilin (component of a short filamentous projection on bacterial cell surface, used for adhering to other bacterial cells and eukaryotic cells) and thus are termed 'pseudopili' (Johnson *et al*, 2006). Type II substrates are initially transported by the Sec or TAT translocon into the periplasm where their signal peptides are removed. The mature proteins are subsequently "pushed" across the outer membrane by the pseudopilus (Sandkvist, 2001). T2SSs were originally discovered in *Klebsiella oxytoca* to secret a starch hydrolyzing enzyme pullulanase (d'Enfert *et al*, 1987) and later research revealed that it is widely utilized by Gram negative pathogens to secret toxins, proteases, cellulases and lipases (Johnson *et al*, 2006).

While T2SSs are similar to type IV pili, the type III secretion systems (T3SSs) bear homology to the flagellar export apparatus, where substrates are directly injected from the bacterial cytoplasm into host cells through a needle-like structure composed of 20-30 protein components (Arnold *et al*, 2010). T3SSs are thus also called 'injectisome's. Substrate export is triggered by direct contact with host cells (Pettersson *et al*, 1996). The Ysc injectisome involved in 'Yop' (*Yersinia* outer proteins, virulence factors) protein secretion by *Yersinia spp.* is the archetype of T3SSs (Cornelis & Wolf-Watz, 1997). Similar systems are also found in *Pseudomonas aeruginosa*, *Salmonella enterica* serovar Typhimurium and *Shigella flexneri* (Kubori *et al*, 1998; Roy-Burman *et al*, 2001; Tamano *et al*, 2000). Effectors translocated by T3SS are quite diverse and do not show any typical domain organization, although they do have a Ser rich N-terminus (Arnold *et al*, 2010).

Type IV secretion systems (T4SSs) are versatile secretion machineries composed of 9-12 protein components that assemble into an ATP-energized double membrane-spanning

complex. They serve as the subject of this study and will be discussed in detail in the following sections.

Similar to T2SSs, the type V pathways (T5SSs) also rely on the Sec secretory pathway to export substrates into the periplasm, although in this case, the secreted protein itself, once in the periplasm, directs its export across the outer membrane and is thus referred to as an 'autotransporter' (Henderson *et al*, 2004). The secreted proteins usually contain an N-terminal signal sequence, a passenger domain (domain to be translocated), a linker region and a β domain (which adopts a β-barrel structure upon membrane integration), but in some cases, the β-barrel domain is produced as a separate protein (Henderson *et al*, 1998; Jacob-Dubuisson *et al*, 2001). Once inserted into the outer membrane, the β-barrel domain forms a pore in the outer membrane, through which the passenger domain traverses the outer membrane (Henderson & Nataro, 2001). In Gram negative bacteria, a vast number of proteases, esterases, toxins, and lipases have been identified as autotransporters and the prototype amongst them is the IgA1 protease produced by *Neisseria gonorrhoeae* that attacks the immunoglobulin of the human IgA1 isotype (Pohlner *et al*, 1987).

Finally, type VI secretion systems (T6SSs) were recently found to be involved in virulence factor excretion from pathogens like *Vibrio cholera, Salmonella typhimurium, Yersinia pestis* and *Pseudomonas aeruginosa* (Mougous *et al*, 2006; Parsons & Heffron, 2005; Pukatzki *et al*, 2006; Yen *et al*, 2008). They are assembled from 15-25 protein components and are proposed to secret effectors independent of the Sec-translocon (Cascales, 2008; Pukatzki *et al*, 2009). Although the macromolecular structure and the effector delivery mechanism have not been thoroughly resolved, two secreted T6SS structural components, the valine–glycine repeat protein G (VgrG) and hemolysin coregulated protein (Hcp), are structurally homologous to the components of the puncturing device (tail) of T4 bacteriophage (Leiman *et al*, 2009; Pell *et al*, 2009), indicating they may function in a similar way to T4 phage. This was most recently evidenced by the fact that two other T6SS components of *V. cholerae*, VipA and VipB, assemble into a tubular structure resembling the T4 contracted tail sheath and powers substrate translocation through a phage tail-like contraction mechanism (Basler *et al*, 2012).

There exists another specialized type VII secretion system (T7SS) in the Gram positive bacterium *Mycobacterium tuberculosis* for the secretion of two T-cell antigenic targets, ESAT-6 (early secreted antigenic target of 6 kDa) and CFP-10 (culture filtrate protein of 10 kDa) (Lewis *et al*, 2003; Pym *et al*, 2002). They are both essential for the virulence of *M. tuberculosis* and both lack a canonical signal sequence. T7SS substrates are usually secreted in pairs so as to form a tight dimer and stabilize each other (Fortune *et al*, 2005). They are recruited through the recognition of a C-terminal signal to the membrane, where they get translocated across a membrane spanning channel in an ATP-dependent manner (Abdallah *et al*, 2007). In addition to *Mycobacteria*, most of the T7SSs identified so far belong to the order of *Actinomycetales* (Abdallah *et al*, 2007).

1.2 TYPE IV SECRETION SYSTEMS ENCODED BY GRAM NEGATIVE BACTERIA

The type IV secretion system (T4SS) was first defined following an investigation into the homology shared between three transport systems: the tumorigenic DNA transfer complex of *Agrobacterium tumefaciens*, the conjugation system encoded by the *tra* region of the IncN group (Inc for incompatibility) plasmid pKM101, and the *ptl* toxin export system of *Bordetella pertussis* (Christie & Vogel, 2000). Since then, the T4SS family has expanded enormously, due in large part to advances in genome sequencing over the last decade. T4SSs are widely utilized by Gram negative bacteria to mediate the translocation of macromolecules across the cell envelope (Cascales & Christie, 2003; Grohmann *et al*, 2003). They are ancestrally related to conjugation machineries and can be classified into three subfamilies on the basis of substrate and target cells: conjugation systems, DNA uptake and release systems, and effector translocation systems (Cascales & Christie, 2003; Christie *et al*, 2005).

The largest of the three T4SS subfamilies encompasses the conjugation systems encoded by conjugative plasmids such as *E. coli* F (IncF), RP4 (IncP), pKM101 (IncN), and R388 (IncW) plasmids (Cascales & Christie, 2003). These systems mediate the horizontal transfer of plasmid-encoded metabolic genes and antibiotic resistance genes between bacteria and thus are important contributors to both bacterial fitness under changing environmental conditions, and to antibiotic resistance in clinically significant pathogens (Baron, 2005). The 'DNA uptake and release' subfamily is found in *Helicobacter pylori* (DNA uptake), and *N. gonorrhoeae* (DNA release). These systems deliver DNA substrates to, or facilitate the acquisition of DNA from, the extracellular milieu, and this process does not require direct contact with target cells (Cascales & Christie, 2003; Christie *et al*, 2005). The third and final subfamily includes the effector translocation systems of the phytophathogen *A. tumefaciens* and the mammalian pathogens *Brucella suis, H. pylori, Legionella pneumophila, B. pertussis*, and *Bartonella henselae* (Baron, 2005; Cascales & Christie, 2003) These Gram negative pathogens utilize T4SSs to deliver oncogenic DNA or protein effectors into the cytosol of eukaryotic cells (Baron *et al*, 2002).

At the same time, T4SSs can also be classified into two subgroups according to sequence similarity and gene organization of the constituents. As the T4SS family expands, it was noticed that in some systems, the protein components exhibit significant homology to the VirB components of *A. tumefaciens* VirB/D4 T4SS (Fig. 1.1). For instance, the T4SS component genes in *B. suis* are colinearly arranged in a single operon like the *virB* operon of *A. tumefaciens*, suggesting a common ancestral origin (Christie & Vogel, 2000). By contrast, the *dot/icm* system of *L. pneumophila* is unrelated to this paradigmatic system but bears homology to the *tra* locus of the IncI plasmid ColIb-P9 of *Shigella flexneri* (Christie & Vogel, 2000). Consequently, T4SSs are further divided into type IVA and type IVB subgroups, with Type IVA group including systems resembling

the VirB/D4 system of *A. tumefaciens*, while type IVB group containing secretion systems assembled from Tra homologs (Christie & Vogel, 2000). Our current understanding of T4SS system function is primarily based on studies on the type IVA group, and this is also the focus of my thesis work.

1.3 TYPE IV SECRETION SYSTEMS IN *A. TUMEFACIENS*

Since 1907, *A. tumefaciens* has been extensively utilized as a model organism for studying plant-pathogen interactions (Zupan *et al*, 2000). As a ubiquitous soil-borne pathogen, it infects a wide range of dicotyledonous plants and causes crown gall tumors (Zupan & Zambryksi, 1995). The galls, as the name implies, usually develop at the crown level, the point at the soil line where main roots join the stem, but can also form on roots or stems (Cooksey & Moore, 1980). Young galls are usually smooth and spongy and, as they age, will become more rough and woody (Cooksey & Moore, 1980). The cause of tumor/gall formation is the transfer of an oncogenic DNA fragment (T-DNA) from *A. tumefaciens* to the plant cells.

The genome of wild-type *A. tumefaciens* strain C58 includes a circular chromosome, a linear chromosome, and two plasmids: pAtC58 and pTiC58 (Allardet-Servent *et al*, 1993). The tumor-inducing Ti plasmid (pTiC58) encodes two T4SSs: Tra and VirB/D4 (Kuldau *et al*, 1990; Li *et al*, 1998; Zupan *et al*, 2000). The Trb system consists of 11 proteins and mediates the conjugative transfer of the Ti plasmid between agrobacteria cells (Li *et al*, 1998). The virulence-related VirB/D4 system, conversely, is required for the transfer of T-DNA (a fragment of oncogenic DNA excised from the Ti plasmid) into susceptible plant cells (Kuldau *et al*, 1990). A third T4SS is found on the other plasmid pAtC58, and is encoded by the *avhB* locus. This system is proposed to promote the conjugal transfer of pAtC58 between agrobacteria (Chen *et al*, 2002). All three T4SSs encoded by *A. tumefaciens* belong to the typeIVA subgroup. To date, the VirB/D4 system encoded by Ti plasmid is the most thoroughly studied T4SS.

The 30 kb virulence regulon of the Ti plasmid contains at least six essential loci (*virA*, *virB*, *virC*, *virD*, *virE*, *virG*) and two non-essential operons (*virF*, *virH*) participating in T4SS expression and assembly (Riva *et al*, 1998). The *virA* and *virG* regions each encompasses a single gene and are constitutively transcribed (Riva *et al*, 1998). VirA is a transmembrane sensor protein that self-phosphorylates in response to signaling molecules found in plant exudates (Riva *et al*, 1998; Zupan *et al*, 2000). Activated VirA transfers the phosphate to the cytoplasmic transcription factor VirG, which in turn activates the transcription of other *vir* genes (Riva *et al*, 1998). Phenolic compounds such as acetosyringone (AS) are the most potent inducers, while monosaccharides and acidic pH both amplify the acetosyringone response (Zupan *et al*, 2000). Upon virulence induction by VirG, the resulting protein products of the 11 *virB* genes (*virB1* through *virB11*) and the *virD4* gene assemble into a T4SS complex spanning both the inner and the outer cell membranes (Fig 1.1). At the same time, the 20 kb T-DNA fragment from the Ti plasmid is excised by the relaxase VirD2, which then remains covalently linked to

the 5′ end of the T-DNA strand (Zupan & Zambryksi, 1995). The resulting T-DNA is then coated by the single-stranded DNA binding protein VirE2 to form a 'T-complex', which is then directed to the T4SS for translocation into plant cells. Once in the plant cell, the T-complex is targeted to the nucleus by the nuclear localization sequences of VirD2 and VirE2 (Zupan & Zambryksi, 1995). Through a currently unknown mechanism, the T-DNA is integrated into the plant chromosome and is transcribed along with plant genes. Since the T-DNA codes for enzymes synthesizing the plant growth hormones auxin and cytokinin, T-DNA integration and gene expression ultimately disrupts hormonal balance, resulting in tumor formation (Riva *et al*, 1998).

1.4 THE ARCHITECTURE AND COMPONENTS OF THE VIRB/D4 TYPE IV SECRETION SYSTEM IN *A. TUMEFACIENS*

The components of the VirB/D4 T4SS can be grouped into three classes according to their subcellular localization and function (Baron *et al*, 2002) (Fig.1.2). The first group includes the major T-pilus constituent VirB2 and the minor component VirB5 (Backert *et al*, 2008). A small inner membrane associated protein VirB3 may also be included in this group, although its exact contribution to T4SS is not known (Beijersbergen *et al*, 1994; Grahn *et al*, 2000; Jones *et al*, 1994; Mossey *et al*, 2010). The T-pilus itself is a long, flexible, filamentous appendage found on the surface of *A. tumefaciens* cells following virulence induction (Lai & Kado, 2002). The major component VirB2 is a hydrophobic protein with a long signal sequence, which directs its export through the Sec pathway into the periplasm (Lai & Kado, 1998). Following signal sequence cleavage, VirB2 is cyclized through the formation of a head-to-tail peptide bond, before being incorporated into the T-pilus through an unknown mechanism (Eisenbrandt *et al*, 1999; Kalkum *et al*, 2002). The accumulation of VirB2 at the inner membrane is independent of all other *virB* genes; however, all of the other *virB* genes are required for T-pilus assembly (Eisenbrandt *et al*, 1999; Lai & Kado, 1998). The minor component VirB5 was reported to co-fractionate with VirB2 both in isolated T-pilus fractions and in a low-molecular-mass membrane bound complex solubilized by mild detergent (Krall *et al*, 2002; Schmidt-Eisenlohr *et al*, 1999). Mutational studies of VirB5 showed that it functions as an adhesin, contributing to the attachment of bacteria to host cells, and this is further supported by the fact that VirB5 is targeted to the tip of T-pilus (Aly & Baron, 2007; Yeo *et al*, 2003). The mechanism driving T-pilus assembly is still not clear and it remains to be clarified whether the T-pilus serves as a conduit for substrate transfer or simply acts as adhesin between agrobacteria and target cells.

The second group comprises the inner membrane proteins VirB6, VirB8 and VirB10, and the outer membrane-associated proteins VirB7 and VirB9. VirB7 is a small lipoprotein attached to the outer membrane via its N-terminus. It forms both homo- and hetero-dimers through disulfide bridges in its N-terminus (homodimer formation) and its C-terminus (heterodimer formation with VirB9) respectively (Baron *et al*, 1997). VirB7-VirB7 and VirB7-VirB9 interactions require the presence of VirB6, a highly hydrophobic protein with 5 transmembrane domains (Jakubowski *et al*, 2003). It directly contacts the

T-DNA through a large periplasmic loop and is essential for substrate transfer to VirB8, VirB2 and VirB9 (Cascales & Christie, 2004b; Jakubowski *et al*, 2004). VirB6 is also important for VirB3/VirB5 stabilization, but the mechanism underlying stabilization is unknown (Hapfelmeier *et al*, 2000). VirB8 and VirB10 are both bi-topic inner membrane proteins (Terradot *et al*, 2005). VirB8 plays a central role by making multiple contacts with other protein components (Das & Xie, 2000; Ward *et al*, 2002). It is postulated to be the nucleating factor for T4SS complex assembly since the targeting of most other VirB proteins to specific sites requires the presence of VirB8 (Judd *et al*, 2005). VirB10 interacts with multiple components at both membranes and transduces the energy released by ATP hydrolysis from the inner to the outer membrane by changing its conformation (Cascales & Christie, 2004a; Sivanesan *et al*, 2010). A recent study revealed the homologs of VirB7, VirB9 and VirB10 encoded by the conjugative plasmid pKM101 assemble into a large core complex with two layers (Fronzes *et al*, 2009b). The inner membrane layer (I) is composed of the N-termini of the VirB9 and VirB10 homologs while the outer membrane layer (O) consists of the VirB7 homologs and the C-terminal domains of the VirB9 and VirB10 homologs.

The third group consists of three predicted NTPases VirB4, VirB11 and VirD4, with each possessing a Walker A nucleotide-binding site. In support of this prediction, ATP-hydrolyzing activity has been demonstrated for homologous proteins belonging to all three groups (VirB4, VirB11 and VirD4 families) (Krause *et al*, 2000b; Tato *et al*, 2005; Arechaga *et al*, 2008; Durand *et al*, 2010). VirB11 is a peripheral inner membrane protein that exists in a dynamic equilibrium between the cytoplasm and the membrane, while VirD4 is also tethered to the inner membrane through its N-terminus (Krause *et al*, 2000a; Rashkova *et al*, 1997). X-ray structures of VirB11 and VirD4 homologs revealed them both to form hexameric ring structures (Hare *et al*, 2006; Yeo & Waksman, 2004). Although VirB4 tightly associates with the inner membrane in gradient density ultracentrifugation (Fullner *et al*, 1994), its membrane topology and oligomeric state have not been unequivocally elucidated (Dang & Christie, 1997; Draper *et al*, 2006). VirB4, VirB11, and VirD4 interact with each other and are believed to coordinate their NTPase activities to drive substrate translocation across the cell envelope through the secretion channel (Atumakuri *et al*, 2004; Fronzes *et al*, 2009a**;** Ward *et al*, 2002). VirD4 also acts as a coupling factor, bringing the T-DNA and its associated proteins to the T4SS (Atumakuri *et al*, 2004; Hormaeche *et al*, 2002).

VirB1 is the only non-essential T4SS component in *A. tumefaciens*: its deletion attenuates substrate transfer, but does not affect T4SS assembly (Berger & Christie, 1994). In the periplasm, VirB1 is processed to yield VirB1*, a 73 amino acid peptide that comprises the C-terminus of VirB1, before being secreted into the extracellular milieu where it promotes T-pilus biogenesis (Baron *et al*, 1997; Llosa *et al*, 2000; Zupan *et al*, 2007). The N-terminal domain of VirB1 bears homology to lytic transglycosylases and following processing, it remains in the periplasm and degrades the peptidoglycan to create space for T4SS assembly (Höppner *et al*, 2004).

1.5 THE RELEVANCE OF THE TYPE IV SECRETION SYSTEM IN MEDICINE

As mentioned in Section 1.3, agrobacteria-mediated genetic transformation has found broad utility in creating transgenic plants with desired traits (Lybarger & Sandkvist, 2004). Moreover, the horizontal gene transfer driven by T4SSs may also be beneficial for processes like environmental remediation, as the metabolic pathways encoded by transmissible, broad host range plasmids can promote the degradation of xenobiotics and detoxification of heavy metals pollutants (Baron, 2005). At the same time, the exchange of genetic information via T4SSs can also be detrimental to human health, as it may facilitate the spread of antibiotic resistance genes, aggravating the threat of multiple antibiotic resistance in pathogenic bacteria (Baron, 2005). Furthermore , T4SSs themselves are employed by a number of Gram negative pathogens to secrete effectors across their cell envelope and cause a great variety of diseases.

Intracellular pathogens:

Brucella spp.: *Brucella* are considered a potential biological weapon, as they are the causative agent of brucellosis. Brucellosis is the most common cause of abortions in cows (*B. abortus*), pigs (*B. suis*) and goats (*B. melitensis*) (Pappas *et al*, 2006). Among the ten species currently known, *B. melitensis*, *B. abortus*, *B. suis* and *B. canis* can also cause diseases in humans, where brucellosis syndrome is typically presented as intermittent fever and may lead to chronic disease with very severe complications (Pappas *et al*, 2006). *Brucella* can be transmitted by consuming unpasteurized dairy products, directly contacting infected animal tissues or inhaling aerosols of the bacteria (Pappas *et al*, 2006). It is the aerosolization characteristic that has lead to concerns about the 'weaponization' of *Brucella*. A very important aspect of *Brucella* pathogenesis is their ability to evade the host immune system and effectively multiply within host cells (Boschiroli *et al*, 2002b). The VirB T4SS was first identified as an important contributor to *Brucella* pathogenesis when researchers were studying mutants displaying attenuated survival in HeLa cells and macrophages (O'Callaghan *et al*, 1999). As in the case of the *virB* operon of *A. tumefaciens*, the 11 *virB*-like genes in the mutated region of *B. suis* genome are co-linearly arranged in a single operon, although *Brucella* has an extra ORF, *virB12*, that is absent in *A. tumefaciens* (O'Callaghan *et al*, 1999). Since *Brucella* do not naturally carry any plasmids, the VirB system seems likely to be dedicated to protein effector secretion during infection rather than plasmid conjugation and this was recently evidenced by the identification of two proteins, VceA and VceC, secreted by the VirB T4SS of *B. abortus* (Boschiroli *et al*, 2002a; de Jong *et al*, 2008). Similar to *A. tumefaciens*, where the *virB* system induction is enhanced by acidic pH, the *virB* operon of *B. suis* is also induced by phagosome acidification upon entry into eukaryotic cells (Boschiroli *et al*, 2002b). The *Brucella* T4SS helps to inhibit phagosome-lysosome fusion and modulates maturation of *Brucella*-containing vacuoles into organelles with many features characteristic of the endoplasmic reticulum (Pizarro-Cerdá *et al*, 2000; Boschiroli *et al*, 2002b; Naroeni *et al*, 2001).

Legionella pneumophila: *L. pneumophila* is the causative agent of legionnaire's disease. When inhaled by humans, it can infect, multiply within and kill human macrophages (Backert & Meyer, 2006). A functional Dot/Icm T4SS (type IVB) is crucial for bacterial intracellular replication and induction of macrophage apoptosis (Zink *et al*, 2002). *Legionella*, once engulfed through regular phagocytosis, are capable of evading lysosome fusion and mediating the recruitment of endoplasmic reticulum (ER) derived vesicles (Hubber & Roy, 2010). These specialized vacuoles provide the bacteria with a niche to survive and multiple within until the cells are eventually lysed, releasing the bacteria to start a new round of infection (Coers *et al*, 1999; Wiater *et al*, 1998). Over 140 protein effectors have been confirmed to be secreted by the Dot/Icm T4SS and most of them interact with host proteins to modulate *Legionella*-containing vacuole trafficking, ER recruitment, and host cell apoptosis (Hubber & Roy, 2010).

Bartonella henselae: The zoonotic pathogen *B. henselae* infects cats and can be transmitted to humans, leading to cat scratch disease (Keret *et al*, 1998). A common theme shared among *Bartonella spp.* is their ability to infect erythrocytes and cause persistent bacteremia (Keret *et al*, 1998). They also stimulate a pro-inflammatory response by activating NF-κB and render the infected cells refractory to apoptotic stimuli (Schmid *et al*, 2004). Moreover, *Bartonella* prevent endothelial cell apoptosis and promote capillary sprouting (Scheidegger *et al*, 2009; Schmid *et al*, 2006). All of these activities depend on a VirB/D4 T4SS encoded by *B. henselae* (Schmid *et al*, 2004). Like in *Brucella* and *Legionella,* it appears that proteinacious virulence factors are the primary substrates for the *Bartonella* T4SS. To date, seven T4SS substrates, the *Bartonella* effector proteins (Beps), have been identified (Schulein *et al*, 2005). These Beps virulence factors appear to be responsible for the T4SS induced phenotypes; however, their precise functions remain to be defined (Franz & Kempf, 2011).

Extracellular pathogens:

Helicobacter pylori: The Gram negative pathogen *H. pylori* colonizes human gastric mucosa in 50% of the world's population (Peek & Crabtree, 2006; Tegtmeyer *et al*, 2011). Infection by this bacterium can induce chronic gastritis, peptic ulcers and even cancer (Atherton & Blaser, 2009; Peek & Crabtree, 2006). The *cag* (cytotoxin-associated genes) pathogenicity island encodes a T4SS, the only effectors translocated by which are peptidoglycan and CagA (Tegtmeyer *et al*, 2011; Viala *et al*, 2004). Peptidoglycan is recognized by the host protein Nod1 and stimulates NF-κB-dependent pro-inflammatory responses (Viala *et al*, 2004). CagA, once injected, is phosphorylated by oncogenic host kinases and serves to activate a variety of signaling activities such as actin cytoskeletal rearrangements, disruption of cell-cell junctions, as well as proliferative, pro-inflammatory and anti-apoptotic nuclear responses (Backert *et al*, 2010). Apart from the virulence-related Cag system, *H. pylori* harbours another transformation competence ComB T4SS mediating DNA intake (Hofreuter *et al*, 2001).

Bordetella pertussis: *B. pertussis* colonizes the human respiratory epithelia and causes respiratory infections manifested as whooping cough with potentially fatal complications (Mattoo & Cherry, 2005). The pertussis toxin is an A-B type exotoxin composed of five subunits, where each subunit is secreted to the periplasm individually by the Sec translocon and subsequently assemble into a holotoxin in the periplasm (Hewlett, 2007). A *ptl* operon encoded T4SS is required for the translocation of the holotoxin into the extracellular milieu (Weiss *et al*, 1993), where the holotoxin targets host cells and delivers the catalytic S1 subunit into the host cell cytosol (Plaut & Carbonetti, 2008). The S1 subunit ADP-ribosylates G protein involved in cAMP signaling pathway, thus interfering with the function of G-protein coupled receptors, causing defects in migration of neutrophils, monocytes, and lymphocytes (Hewlett, 2007; Shrivasava & Miller, 2009).

The role for T4SSs in bacterial pathogenesis as discussed above provides just a glimpse into the breadth of T4SS function in bacteria. Considering its broad relevance in biomedical fields, greater insight into the T4SS apparatus assembly and the mechanisms underlying substrate translocation will provide us with new targets for designing novel therapeutic agents. Such an understanding will allow us to inhibit T4SS assembly or substrate translocation at pivotal steps, thereby disarming the bacteria. This approach would reduce the selective pressure accompanying most traditional antimicrobial treatments, and may ultimately provide better therapies for treating infections.

1.6 THESIS OBJECTIVES

Thus far, the mechanism underlying T4SS assembly, as well as the precise route by which substrate transfer occurs, has not been deciphered, in part due to the complexity inherent in this multi-component system. According to a previous study in our lab, two T4SS subcomplexes had been identified following solubilization of the membrane bound complex with the mild detergent dodecyl-β-D-maltoside (DDM) (Krall *et al*, 2002). The pilus-associated components VirB2 and VirB5, co-fractionated in a low molecular mass complex, whereas the core VirB6-VirB10 components were detected in a high molecular mass complex (Krall *et al*, 2002). Most of the VirB proteins interact with multiple components in this system, as indicated by the yeast two-hybrid assay (Ward *et al*, 2002). Given that proteins frequently rely on the interplay with their binding partners to fully perform their functions, the focus of this thesis is to evaluate the impact of binary protein interactions on the assembly and function of T4SS in *A. tumefaciens*. To this end, binary protein-protein interactions are first validated using biochemical approaches and this is followed by identification of residues essential for interactions. The mutant VirB proteins are subsequently expressed in *A. tumefaciens* and subjected to functional analyses to evaluate the importance of the interactions.

1.7 FIGURES AND TABLES

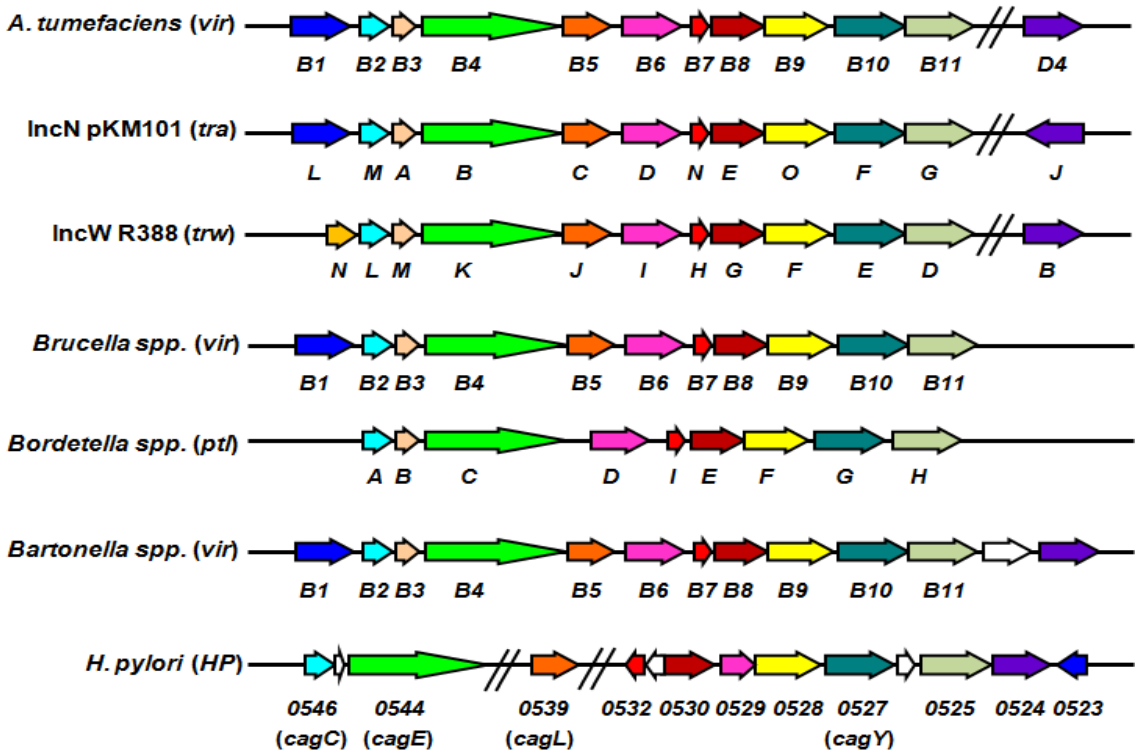

Fig. 1.1 Schematic representation of genome context of the loci encoding type IV secretion systems (type IVA) in Gram negative bacteria. The *A. tumefaciens* VirB/D4 T4SS mediates T-DNA transfer, VirE2 protein secretion and conjugation of other plasmids. The Tra and Trw systems encoded by plasmids pKM101 and R388, respectively, are responsible for conjugation of the plasmids themselves. The Vir, Ptl, and HP (Cag) systems encoded by *Brucella, Bordetella, Bartonella,* and *H. pylori* are involved in translocating proteinaceous virulence factors. Homologous proteins are shaded in the same color in different bacteria. The figure was created based on Christie & Vogel (2000).

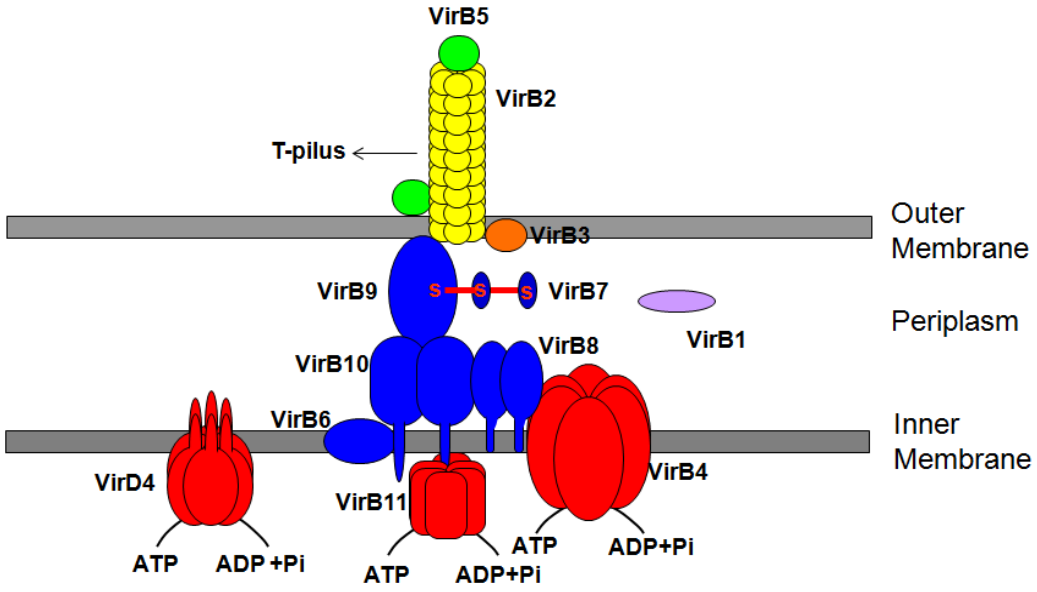

Fig. 1.2 Architectural model of *Agrobacterium tumefaciens* T4SS. Three NTPases, VirB4, VirB11, and VirD4, interact with each other and coordinate their NTPase activities to mediate the substrate translocation across the cell envelope, through the secretion channel composed of the core components VirB6-VirB10, to the outside of the cell with the aid of the T-pilus assembled from VirB2 and VirB5. Model was created based on Baron (2005).

Table 1.1 Summary of secretion systems in Gram negative bacteria.

Name	Number of components	Sec-dependent	Extracellular appendage	Energy source	Signal sequence	Substrates	Example
T1SS	3	No	No	ATP	Uncleaved C-terminal sequence of 46-60aa; poorly conserved	Cytotoxins, cell surface layer proteins, proteases and lipases	Haemolysin secretion system in *Escherichia coli*
T2SS	12-16	Yes	Pseudopilus; similar to type IV pilus	ATP	Sec N-terminal signal peptide; a positively charged N-terminus, a hydrophobic core and a hydrophilic C-terminus	Toxins, proteases, cellulases and lipases	Secretion of starch hydrolyzing enzyme pullulanase in *Klebsiella oxytoca*
T3SS	20-30	No	T3SS forms needle-like structure; resembles flagella	ATP	Ser rich N-terminus	Pathogen effector proteins and toxins	Ysc injectisome involved in 'Yop' (*Yersinia* outer proteins) secretion by *Yersinia spp.*
T4SS	9-12	No	Pilus/adhesins (T-pilus for *Agrobacterium tumefaciens*)	ATP	C-terminal sequence; maybe contain positively charged or hydrophobic residues	T-DNA, plasmid conjugation, toxins and protein effectors	T-DNA transfer system of *Agrobacterium tumefaciens*
T5SS	1-2	Yes	No	No	Sec N-terminal signal peptide; a positively charged N-terminus, a hydrophobic core and a hydrophilic C-terminus	Autotransporter; proteases, esterases, toxins, and lipases	IgA1 protease secretion by *Neisseria gonorrhoeae*
T6SS	15-25	No	Cell puncturing device similar to that of bacteriophage T4	T4 phage tail-like contraction	Not clear	Hemolysin coregulated protein (Hcp) and the valine-glycine repeat protein G (VgrG)	Hcp1 secretion by HIS-I system of *Pseudomonas aeruginosa*; VgrG secretion by VAS system of *Vibrio cholerae*

CHAPTER 2:
AN EVALUATION OF THE IMPACT OF VIRB4-VIRB8 INTERACTION ON TYPE IV SECRETION SYSTEM

2.1 INTRODUCTION:

VirB4 is the largest and the most widely conserved component of the T4SS. Despite extensive previous studies focused on VirB4-like proteins, little is known about their subcellular localization and atomic structures. Affinity chromatography indicated that *A. tumefaciens* VirB4 self-associates, and it is likely to form a hexamer according to an *in silico* model constructed for the C-terminus of *A. tumefaciens* VirB4 (Dang *et al*, 1999; Middleton *et al*, 2005). This prediction was later supported by studies showing that VirB4 homologs such as TraB, encoded by the IncN plasmid pKM101, and TrwK, encoded by the IncQ plasmid R388, both assemble into homohexamers (Arechaga *et al*, 2008; Durand *et al*, 2010). While TrwK predominantly purifies as a monomer, its catalytically competent form is a hexamer (Arechaga *et al*, 2008). Likewise, the TraB dimer extracted from the membrane fraction is not capable of ATP hydrolysis, whereas the soluble hexameric form displays ATPase activity (Durand *et al*, 2010). *A. tumefaciens* VirB4 possesses several hydrophobic segments of sufficient length to span the inner membrane, but these regions are all interrupted by charged residues, making it unlikely that VirB4 functions as an integral membrane protein (Kuldau *et al*, 1990). Despite this, *A. tumefaciens* VirB4 associates with the inner membrane in sucrose density gradient centrifugation (Dang & Christie, 1997; Shirasu *et al*, 1994) and has been suggested to have two periplasmic domains (58-84 aa and 450-514 aa) and four transmembrane (TM) domains, based on results of an alkaline phosphatase fusion study (Dang & Christie, 1997) (Fig. 2.1). However, this is contradictory to the results of a similar experiment done in another laboratory showing no membrane insertion by *A. tumefaciens* VirB4 (Das & Xie, 2000). Predictions using TMHMM server 2.0, HMMTOP server and DAS server do not support the presence of TM domains in VirB4 (Middleton *et al*, 2005).

Multiple sequence alignment of VirB4 family proteins identified two highly conserved motifs widely involved in nucleoside triphosphate (NTP) hydrolysis: a Walker A box (GXXGXGKT/S) (where G, K, T and S denote Gly, Lys, Thr and Ser respectively and X denotes any amino acid) and a Walker B box (hhhhDE) (where D and E denote Asp and Glu respectively, and h is any hydrophobic amino acid) (Hanson & Whiteheart, 2005; Walker *et al*, 1982) (Fig. 2.1). A protein harbouring a Walker A box can hydrolyze NTP only when the amino acids in the Walker A motif form a P-loop structure preceded by a β-sheet and followed by an α-helix (Ramakrishnan *et al*, 2002). The highly conserved, positively charged Lys residue is essential for NTP binding and directly interacts with the α- and β-phosphoryl groups to neutralize their negative charges, facilitating the removal of the γ-phosphate during hydrolysis (Fry *et al*, 1986). The highly conserved Asp and Glu residues in the Walker B motif are structurally proximal to the Walker A box so that the Asp residue can coordinate Mg^{2+} bound on substrate NTP and Glu activates water for the

hydrolysis reaction (Hanson & Whiteheart, 2005). These two motifs work in conjunction to keep the NTP appropriately positioned for hydrolysis (Fry *et al*, 1986). Thus far, amongst all VirB4-like proteins, only the ATP hydrolyzing activity of TraB and TrwK have been unequivocally determined (Arechaga *et al*, 2008; Durand *et al*, 2010). Moreover, the ATPase activity of wild-type TrwK is increased dramatically upon truncation of its C-terminus, but is decreased following the addition of synthetic peptides corresponding to the C-terminal sequence of TrwK, suggesting that the C-terminal domain has autoinhibitory activity (Pena *et al*, 2011). Due to poor protein solubility and low yield, the ATPase activity of other Vir4 homologs has not been convincingly demonstrated.

A previous study using an *A. tumefaciens virB4* deletion mutant indicated that VirB4 helps stabilize both VirB3 and VirB8 and is required for T-pilus formation and T-DNA transfer. It is also indispensable for the conjugative transfer of an IncQ plasmid pLS1 between agrobacteria strains, both when the T4SS was expressed in the donor (for mobilizing plasmid exit) and the recipient (for plasmid uptake) strains (Beijersbergen *et al*, 1992; Bohne *et al*, 1998; Yuan *et al*, 2005). Complementation of the *virB4* deletion mutant strain with wild-type *A. tumefaciens* VirB4 (VirB4a) and its *B. suis* homolog (VirB4s) showed that all defects could be fully restored by VirB4a and partially rescued by VirB4s, indicating that these two VirB4 proteins undergo similar interactions with other proteins (Yuan *et al*, 2005). When a Walker A box variant of VirB4 was subjected to the same complementation analyses, it had a hybrid phenotype: VirB3 and VirB8 stability was not affected; extracellular T-pilus formation and uptake of pLS1 upon VirB4 expression in the recipient cells proceeded as in the wild-type; but the mutant was unable to induce tumor formation or promote pLS1 transfer upon its production in the donor strain (Yuan *et al*, 2005). When the membrane bound T4SS system of *A. tumefaciens* was extracted with a mild detergent and separated using blue native-PAGE and gel filtration chromatography, the major T-pilus component VirB2 co-fractionated with the minor component VirB5 in a low molecular mass fraction, whereas VirB6-10 co-fractionated in a high molecular mass fraction (Krall *et al*, 2002). In the *virB4* mutant, however, the co-fractionation of VirB2 and VirB5 was abolished and this defect could be complemented by expressing wild-type VirB4a/VirB4s as well as their Walker A box variants *in trans*, indicating that NTPase activity is not required for physical assembly (Yuan *et al*, 2005). Collectively, these *in vivo* assay results prompted us to hypothesize that VirB4 supports the physical assembly of T4SS apparatus through interactions with other components, and at the same time drives substrate translocation in an NTP hydrolysis-dependent manner (Dang *et al*, 1999).

A yeast two-hybrid study identified VirB8 as an interacting partner of VirB4 (Ward *et al*, 2002). VirB8 is a bitopic inner membrane protein with a short cytoplasmic N-terminus and a large periplasmic C-terminus (Das & Xie, 2000). X-ray crystal structures for both *B. suis* VirB8 periplasmic domain (VirB8sp) and *A. tumefaciens* VirB8 periplasmic domain (VirB8ap) provided us with a wealth of information on their potential interaction sites with other proteins (Fig. 2.2). Both protein domains crystallize as dimers and each

monomer comprises five α-helices juxtaposed against a four-stranded antiparallel β-sheet (Bailey *et al*, 2006; Terradot *et al*, 2005). In solution, VirB8 exists in a dynamic monomer-dimer equilibrium, as purified StrepII tagged VirB8s (*B. suis* VirB8) self-associates in a concentration dependent manner, with high concentrations promoting dimer formation and low concentrations favouring existence as monomers (Paschos *et al*, 2006). Using a transfer DNA immunoprecipitation assay where T-DNA was co-precipitated with the VirB proteins it contacted, Dr. P.J. Christie's group defined a pathway for T-DNA transfer (Cascales & Christie, 2004b). Although VirB4 does not directly contact T-DNA, its deletion arrests the transfer from VirB11 to the core components VirB6 and VirB8 (Cascales & Christie, 2004b). As it appears that VirB4 interacts with VirB8, it is possible that VirB4 performs its function in coordination with VirB8. In the following chapter, VirB4-VirB8 association was examined using biochemical approaches, and interaction sites in VirB8 were identified. This was followed by assessing the impact of VirB4-VirB8 interaction on T4SS using *in vivo* assays. The ATP binding and hydrolytic activity of VirB4 was also checked using both ATPase and ATP-binding assays. As *A. tumefaciens* T4SS components are problematic for biochemical work due to poor solubility and low yield, all of the *in vitro* experiments were conducted using their *B. suis* homologs. From now on, all *A. tumefaciens* and *B. suis* VirB components names will be followed by either an "a" (*Agrobacterium)* or an "s" (*suis)* respectively, for the sake of simplicity.

2.2 MATERIALS AND METHODS

2.2.1 Bacterial strains, plasmids, and culture conditions

A. tumefaciens and *E. coli* strains used in this study are summarized in Table 2.1. Plasmids used in this study are listed in Table 2.2. Media, buffers and other materials are listed in Table 2.3. *E. coli* strains were grown at 37°C on solid or in liquid LB (Luria Bertani) medium (for composition of all media, please refer to Table 2.3). For virulence induction, *A. tumefaciens* was grown overnight in liquid YEB medium (Vervliet *et al*, 1975) at 26°C before being inoculated into AB minimal medium (ABMM) (Chilton *et al*, 1974) at a starting OD_{600} of 0.2 and cultivated at 20°C at 200 rpm for 5 h to pre-induce the cells. The cells were subsequently spread onto solid ABMM medium with (experimental group and positive control) or without (negative control) 200 μM acetosyringone (AS) (to induce virulence) and incubated at 20°C for 3-4 days before being harvested. When necessary, 500 μM IPTG was added to induce protein expression from the pTrc vector. Antibiotics were added to maintain plasmids when necessary. For *E. coli* cultures, antibiotics were added to the following final concentrations (μg/ml): carbenicillin, 100; kanamycin, 50; streptomycin, 50; spectinomycin, 50. For *Agrobacterium* cultures, antibiotics were added to YEB medium at the following final concentrations (μg/ml): streptomycin, 100; spectinomycin, 300.

2.2.2 Oligonucleotides

Oligonucleotides are listed in Table 2.4.

2.2.3 Plasmid construction

Manipulation of DNA including plasmid extraction, PCR amplification, digestion, ligation and sequencing were conducted following standard procedures (Maniatis *et al*, 1982). Restriction enzymes were purchased from New England Biolabs and MBI Fermentas and *E. coli* JM109 was used as a cloning host. To make pT7-7*strepII-virB4s*^{K464R}, the Lys residue (codon AAA) in the Walker A box was mutagenized to an Arg residue (AGA) using site-directed mutagenesis (Shenoy & Visweswariah, 2003). Primers carrying the desired mutation (AAA→AGA) were used for inverse PCR on the pT7-7*strepII-virB4s* backbone using a Mastercycler Gradient (Eppendorf). The PCR cycle was: 95°C for 5 min, followed by 30 cycles of 65°C for 30 s, 72°C for 4 min (+ 10s/each cycle), 95°C for 30 s; and a final 7 min at 72°C. The parental strands in 50 µl reaction were digested using 2 µl DpnI (10 U/µl). DNA was purified using a MinElute Gel Extraction Kit (QIAGEN), followed by electroporation (Eppendorf Electroporator 2510) into JM109 competent cells. The plasmids were extracted using QIAprep Spin Miniprep Kit (QIAGEN) and sent for sequencing. To mutagenize the Q223 residue of VirB8a to Glu/Lys in both pKT25*virB8ap* and pTrc*virB7avirB8a*, the codon CAG was changed to GAG/AAG. Inverse PCR was conducted on the backbones of pKT25 and pTrc200 respectively as described above, only using an annealing temperature of 60°C for pKT25 and an annealing temperature of 62°C for pTrc200.

To create bacterial-two hybrid assay constructs, the fragments encoding interaction partners [VirB4s-VirB8s/VirB8sp; VirB4a-VirB8a/VirB8ap; VirB4a (198-401aa)–VirB8ap; VirB4a (329-650aa)-VirB8ap; VirB4a (237-789aa)-VirB8ap] were PCR amplified and digested with *Xba*I/*Kpn*I. The fragments were purified as described above and either cloned into pKT25/pUT18C to create C-terminal fusions with T18/25 fragments for each partner, or cloned into pKNT25/pUT18 to create N-terminal fusions. The ligation products were electroporated into JM109 competent cells. Plasmids were isolated as described and sent for sequencing. To make *Agrobacterium* complementation constructs pTrc*virB7avirB8a, virB7a and virB8a* genes were co-amplified, digested with *Eco*RI/*Bam*HI, and cloned into pTrc200 vector following standard procedures.

2.2.4 SDS-PAGE and western blotting

To examine cellular protein levels, OD_{600} values for *E. coli* or *Agrobacterium* culture aliquots were measured and cells were collected by centrifugation at 10,000 × g for 5 min. Laemmli sample buffer (2×) was added according to the formula: volume of sample buffer (µl) = OD_{600} × 50 × volume of culture aliquot (ml). Cells were resuspended and boiled in sample buffer at 100°C for 5 min. For all liquid form samples, equal volumes of 2 × Laemmli sample buffer were mixed with protein samples followed by boiling at 100°C for 5 min. For SDS-PAGE, samples were resolved in linear 12.5% Laemmli gels (for proteins >20 kDa) (Laemmli, 1970) or 12.5% Schägger gels (for proteins <20 kDa)

(Schägger & Jagow, 1987) using a mini gel system (GE Healthcare). Proteins were transferred to 0.45 µm PVDF membrane (Immobilon-P, Millipore) in a vertical blot device (Trans Blot Cell, Bio-Rad) at 90 V for 1 h, or at 30V for 16 h at 4°C (Harlow & Lane, 1988). Immunoblotting was carried out using appropriate primary antibodies and anti-rabbit/mouse IgG horseradish peroxidase-linked secondary antibody. Proteins were detected using a chemiluminescence detection system (Lumi Light, Roche Diagnostics) and X-ray film (Kodak) (Harlow & Lane, 1988). To assess protein levels, gels were immersed in Coomassie Blue staining solution and microwaved for 30 s before being stained for 30 min with mild shaking. Destaining was performed overnight in destaining buffer (Neuhoff *et al*, 1985).

2.2.5 Protein overproduction and purification

To overexpress VirB4s, *E. coli* GJ1158 cells carrying plasmids pT7-7*strepII-virB4s* or pT7-7*strepII-virB4s*^{K464R} were grown in LBON medium (LB without NaCl) at 37°C until logarithmic phase ($O.D._{600}$=0.5-0.8), after which 0.3 M NaCl was added to induce protein overexpression (Bhandari & Gowrishankar, 1997). Induction was carried out for up to 20 h at 16 °C. To improve protein solubility, 10 mM benzyl alcohol (Sigma) and 1 mM β-mercaptolethanol (Sigma) were added to the culture 20 min before induction. Cells were harvested by centrifugation at 1,000 × g at 4°C for 15 min (Eppendorf) and lysed by four passages through a French Press cell (AMINCO) at 18,000 psi. PMSF (protease inhibitor) was added to the cell suspension to a final concentration of 0.5 mM prior to cell lysis. Resulting cell lysates were centrifuged at 12,000 × g at 4°C for 1 h and the supernatant was loaded onto a Streptavidin Sepharose matrix (Strep-tag MagStrep Kit, IBA). The flowthrough was discarded and the column was washed with three column volumes of buffer W (wash buffer) (for composition of all buffers and solutions, please refer to Table 2.3). Proteins were eluted using buffer E (elution buffer) containing 2.5 mM desthiobiotin. All of the elution fractions were examined using SDS-PAGE and Coomassie Blue staining. The fractions containing Strep tagged VirB4s (StrepII-VirB4s) were combined and enriched using a Nanosep centrifugal device with a cut-off value of 30 kDa (PALL) (as the molecular weight of VirB4s is 96 kDa). Following dialysis in 20 mM Tris-Cl pH 8.0, the sample was subjected to anion exchange chromatography on a HiTrap Q Sepharose Fast Flow column (GE Healthcare) in an AKTA FPLC purifier system at 4°C. Proteins were eluted using a gradient of NaCl in elution buffer (20 mM Tris pH8.0), with NaCl concentration increasing from 0 to 1 M over a volume of 20 ml at a flow rate of 0.5 ml/min. The fractions (0.5 ml/fraction) containing the greatest levels of StrepII-VirB4s were pooled and applied to a Superdex 200 10/300 GL column (GE Healthcare) for gel filtration chromatography. The elution fractions (1 ml fractions) were examined using SDS-PAGE and western blotting with anti-StrepII and anti-GroEL antibodies (anti-GroEL is used for checking GroEL contamination). The fractions containing pure StrepII-VirB4s were subjected to dialysis in storage buffer (50% glycerol, 50% buffer W).

StrepII-VirB8sp variants were overexpressed and purified in a similar fashion to StrepII-VirB4s. The two major differences were: 1 the induction temperature was 26°C; 2, the streptavidin affinity column purified samples were directly subjected to gel filtration chromatography on a Superdex 75 10/300 GL column (GE Healthcare). The elution fractions containing pure StrepII-VirB8s and its variants were subjected to dialysis in storage buffer (50% glycerol, 50% buffer W).

2.2.6 ATPase assay

Purified proteins were incubated in the presence of 1 mM ATP (excessive substrate) in NTPase buffer at 30°C. Ten microlitres were taken from the reaction at 0, 2, 5, 10, 20, 30, 45 and 60 min and mixed with 80 µl malachite green solution in a microtiter plate (Cogan *et al*, 1999). The malachite green solution was prepared by mixing solution A and solution B at a 3:1 ratio followed by stirring for at least 20 min before use. The mixtures of reaction and malachite green solution were incubated for 20 min, before being stopped with 10 µl solution C. The resulting OD_{660} was measured using a spectrophotometer (Spectramax Plus384, Molecular Devices) within 30 min of stopping the reaction. A standard curve was generated using 0, 1, 2, 4, 5, 6, 8, and 10 µl of 2 mM Na_2HPO_4 stock solution. OD_{660} values were then converted to the amounts of phosphate released on the basis of the standard curve, and the values were plotted versus the reaction time to get a reaction progress curve.

2.2.7 ATP-binding assay

Purified StrepII-VirB4s and StrepII-VirB4s^{K464R} were incubated with ATP-agarose slurry (Sigma) at 4°C with mild shaking for 90 min. The agarose beads were pelleted by centrifugation at 12,000 × g at 4°C for 5 min and washed with buffer W for 5 times. Bound proteins were eluted by adding Laemmli sample buffer and boiling at 100°C for 5 min. Eluted samples were analyzed using SDS-PAGE and western blotting with anti-StrepII antibody.

2.2.8 Gel filtration chromatography and blue native-PAGE

To investigate VirB4-VirB8 interaction using proteins overexpressed from the pT7-7*strepII-virB4s-virB8s/virB8sp* bicistronic constructs, the proteins were first isolated using affinity chromatography. Fractions containing these proteins were then loaded onto a Superdex 200 10/300 GL column (GE Healthcare) for gel filtration chromatography in an AKTA FPLC purifier system at 4°C. The flow rate was 0.5 ml/min. As negative controls, purified StrepII-VirB4s and StrepII-VirB8sp were run on a Superdex 200 10/300 GL column and a Superdex 75 10/300 GL column (GE Healthcare), respectively. Column calibrations were performed using high molecular weight and low molecular weight calibration kits (GE Healthcare) containing thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa), aldolase (158 kDa), albumin (67 kDa), ovalbumin (43 kDa), chymotrypsinogen (25 kDa), and ribonuclease A (13.7 kDa). The elution fractions (1 ml

fractions) were analyzed using SDS-PAGE and western blotting with anti-StrepII and anti-VirB8s antibodies. In a separate set of experiments, the affinity column-enriched protein samples were mixed with 5% Coomassie Blue G250 dye (1 µl dye per 20 µl sample) before being separated in a 10% blue native polyacrylamide gel for 7 h at 5 mA at 4°C in a mini gel system (GE Healthcare), followed by western blotting with anti-StrepII and anti-VirB8s antibodies (Schägger & Jagow, 1991).

To screen for VirB8sp mutants defective in binding VirB4s, 35 µl of affinity column enriched StrepII-VirB4s was mixed with 10 µl of purified VirB8sp and its 13 variants (0.01 µg/µl) and 10 µl of buffer W individually, or with 20 µl of buffer W as a control**.** VirB8sp alone controls were made by mixing 5 µl of the StrepIIVirB8sp proteins (0.01 µg/µl) with 15 µl of buffer W. The mixtures were then incubated at room temperature for 1 h, and 13 µl was subjected to blue native-PAGE as described above. StrepII-VirB4s alone and VirB8sp alone were also checked in parallel as negative controls. The samples were also loaded onto a Laemmli gel followed by immunoblotting using anti-VirB8s and anti-StrepII antibodies to ensure equal protein loading in each lane.

2.2.9 Bacterial two-hybrid assay

The bacterial two-hybrid (BTH) assay was performed according to a standard procedure (Karimova *et al*, 1998). Complementary constructs encoding T18/T25 fusion proteins were co-transformed into BTH101 competent cells. Double transformants were selected by cultivating on LB agar plates (carbenicillin, kanamycin). To examine potential interactions between proteins of interest, transformants were streaked onto LB agar plates (carbenicillin, kanamycin) supplemented with 40 µg/ml X-Gal and 500 µM IPTG and cultivated overnight at 26 °C, 30°C or 37 °C. For quantitative β-galactosidase assays, the double transformants were grown in liquid LB (carbenicillin, kanamycin) in the presence of 1000 µM IPTG overnight at 26°C, 30°C, or 37°C. Twenty microliters of each culture was mixed with 80 µl ONPG buffer in a microtiter plate and incubated at room temperature until the solution turned yellow. Stop buffer was added, following which both OD_{420} and OD_{550} were measured on a spectrophotometer (Spectramax Plus384, Molecular Devices). One hundred microlitres of the culture was added to separate wells in the same microtiter plates for OD_{600} readings. Interaction strength was quantified as Miller units = $1000 \times [(OD_{420} - 1.75 \times OD_{550})] / (\text{time of reaction in minutes} \times \text{volume of culture used in the assay in millilitres} \times OD_{600})$.

2.2.10 T-pilus isolation

T-pili were isolated as previously described (Schmidt-Eisenlohr *et al*, 1999). *A. tumefaciens* strains were cultivated on 15 cm ABMM agar plates for virulence induction as described above in Section 2.2.1. Cells were washed off using 50 mM Na-K-P (sodium, potassium, phosphate) buffer, and were then collected by centrifuging at 12,000 × g at 4°C for 1 h. The resulting cell pellet was resuspended in 1 ml Na-K-P buffer and passed 10 times through a 26-gauge needle (BD Biosciences) to shear off surface-associated

extracellular structures. The sample was then centrifuged at 16,000 × g for 45 min at 4°C to remove any intact cells. This was followed by ultracentrifugation in a Ti 90 rotor and an Optima L-90K preparative ultracentrifuge (Beckman Coulter) at 136,720 × g for 2 h at 4°C. The pellet was resuspended in 20 µl 2 × Laemmli sample buffer and examined using SDS-PAGE and western blotting with anti-VirB2 and anti-VirB5.

2.2.11 Conjugation assay

Conjugation assays were performed as described before (Yuan *et al*, 2005). *A. tumefaciens* donor strains carrying the IncQ plasmid pLS1 and the recipient strain UIA143 (pTiA6) were grown individually overnight at 26°C in YEB with appropriate antibiotic selection (carbenicillin, streptomycin, spectinomycin, erythromycin). The cells were pre-induced in liquid ABMM as described in Section 2.2.1. Donor and recipient cells were mixed at a ratio of 5:1 and spotted onto ABMM agar plates containing 500 µg/ml AS, followed by cultivation at 20°C for 3-4 days. When necessary, 500 µg/ml IPTG was added to induce protein expression from complementing plasmids. Each spot was washed off in 1 ml YEB medium, serially diluted, and cultivated on YEB agar plates at 26°C for 2-3 days to recover transconjugants. To differentiate between transconjugants and donor strains, antibiotics were added to YEB agar medium (erythromycin and carbenicillin for transconjugants, while donors were selected using carbenicillin, streptomycin and spectinomycin). The number of colony forming units representing transconjugants and donors were quantified and transfer efficiency was calculated using the formula: transfer efficiency = transconjugants / donor % (TK/D%).

2.2.12 Tumor induction

A. tumefaciens strains were inoculated into liquid ABMM and shaken at 20°C for 18 h in the presence of 200 µg/ml AS. When necessary, 500 µg/ml IPTG was added to induce protein expression from complementing plasmids. Cultures were diluted to an OD_{600} of 0.2, and 20 µl of each strain was loaded on a scratch made on a *Kalanchoë diagremontiana* leaf using a 26-gauge needle (BD Biosciences). The plants were grown at 20°C for up to 4 weeks to allow for tumor development.

2.3. RESULTS

2.3.1 VirB4 purification

VirB4 family proteins have long been known for their low solubility, and this is has greatly hindered their biochemical characterization (Rabel *et al*, 2003; Shirasu *et al*, 1994). When overexpressed in *E. coli*, the majority of Strep-tagged VirB4s (StrepII-VirB4s) aggregated to form inclusion bodies, with only a very small portion being found in the soluble cell extract. Given the unpredictable behaviour of proteins extracted from inclusion bodies, we chose to work with the soluble part of StrepII-VirB4s. To improve

the solubility of StrepII-VirB4s, the induction temperature was decreased to 16°C, and 10 mM benzyl alcohol and 1 mM β-mercaptoethanol were added to the culture 20 min before induction (Berrow *et al*, 2006; Bondos & Bicknell, 2003; Marco *et al*, 2005; Shigapova *et al*, 2005). As a membrane fluidizer, benzyl alcohol is able to induce an artificial heat shock response at a low temperature, leading to overproduction of chaperones, which helps to promote the proper folding of nascent polypeptides (Marco *et al*, 2005). β-mercaptoethanol acts as a reducing agent to prevent protein aggregation by inhibiting formation of non-native disulfide bridges (Bondos & Bicknell, 2003). Following cell lysis, soluble StrepII-VirB4s was subjected to a three-step purification procedure: streptavidin affinity chromatography, ultrafiltration and gel filtration chromatography. The protein appeared to be purified to 95% homogeneity as indicated by Coomassie blue staining (Fig. 2.3). When examined under the electron microscope after negative staining, the purified protein appeared to adopt a heptameric conformation (Fig. 2.4A, B). As VirB4s is not expected to form a heptameric complex, this raised the concern that the sample may be contaminated with the protein chaperone GroEL, which does forms heptameric structures (Radford, 2006). The presence of GroEL was confirmed with immunoblotting using GroEL specific antibodies (Fig 2.4C). Further analysis indicated that GroEL was co-eluted with StrepII-VirB4s from the gel filtration column. Since GroEL is capable of ATP hydrolysis, its presence would interfere with the subsequent ATPase assays, so additional procedures were applied to remove this contaminant.

As shown in Fig. 2.5, GroEL was co-eluted with StrepII-VirB4s from the streptavidin column (Fig. 2.5A); after which, anion exchange chromatography was performed on a Hi Trap Q Sepharose Fast Flow column to separate the two proteins before the sample was run through a gel filtration column. Since the estimated pI (isoelectric point) of StrepII-VirB4s is 6.75, while that of GroEL is 4.96, GroEL is more negatively charged than StrepII-VirB4s at pH 8.0 and thus should bind the anion exchange column more tightly. StrepII-VirB4s should therefore be eluted earlier than the GroEL contaminant. We followed the elution of each protein using immunoblotting, and found that the majority of StrepII-VirB4s was eluted in fractions B5~B3, whereas GroEL was detected predominantly in later fractions (Fig. 2.5B). Once fractions B5 and B4 were combined and run through a gel filtration column, the contaminant GroEL was no longer detectable in the elution fractions (Fig. 2.5C), suggesting that anion exchange chromatography, combined with our routine purification procedure, was sufficient to remove GroEL.

2.3.2 ATPase assay and ATP-binding assay

As described in Section 2.1 of the Introduction, the Walker A box of VirB4 is essential for nucleotide binding and hydrolysis and the Lys residue within this region is particularly important (Sung *et al*, 1988). Mutagenesis of this amino acid to an Arg residue abolished the ATPase activity of the yeast RAD3 protein (Sung *et al*, 1988). Furthermore, the same mutation in VirB4a blocked substrate translocation by *A. tumefaciens* T4SS (Yuan *et al*, 2005). We took advantage of this property, and used an

active site variant, where K464 was replaced by Arg (VirB4s^{K464R}), as a negative control in an ATPase assay. ATP hydrolysis by purified StrepII-VirB4s was measured using the malachite green assay (Cogan *et al*, 1999), a non-radioactive, colorimetric assay based on quantification of the green complex formed between malachite green, molybdate and free phosphate released during the course of a reaction. The color intensity of the resulting product could be measured using a spectrophotometer at a wavelength of 660 nm (Cogan *et al*, 1999). At 5 μM, the positive control (purified StrepII-VirB11s) showed a moderate level of ATPase activity, but the behaviour of the wild-type StrepII-VirB4s was basically the same as that of its active site variant, indicating that under the conditions used, StrepII-VirB4s does not display any obvious ATP hydrolyzing activity (Fig. 2.6).

ATP-binding ability by StrepII-VirB4s was also examined using ATP-agarose. ATP-agarose comprises ATP attached to agarose beads via its γ-phosphate, and has been effectively used to identify kinases, heat shock proteins, and other ATP-binding proteins (Haystead *et al*, 1993). Purified wild-type StrepII-VirB4s and StrepII-VirB4s^{K464R} were incubated with ATP-agarose. StrepII-VirB8sp (the periplasmic domain of VirB8s), which is not expected to bind ATP, was also tested in parallel as a negative control. Following extensive washing, both wild-type StrepII-VirB4s and its active site variant were detected in the eluted samples, while StrepII-VirB8sp did not associate with the agarose beads (Fig. 2.7). This indicated that VirB4s is able to bind ATP and that the change from Lys to Arg in the Walker A motif does not affect this association.

2.3.3 VirB4 interacts with VirB8 *in vitro*

Having established that VirB4 can bind ATP, but cannot hydrolyze it under our assay conditions, our attention turned to investigating the main focus of my thesis, VirB4-VirB8 interaction. To facilitate this analysis, bicistronic constructs were made by cloning DNA fragments encoding full length VirB8s or its periplasmic domain, VirB8sp, downstream of *strepII-virB4s*, under the control of the T7 promoter. This allowed for co-expression of Strep-tagged VirB4s with VirB8s or VirB8sp. Following protein induction in *E. coli* cells, the resulting cell extract was affinity purified using a streptavidin column. VirB8s and VirB8sp expressed from the bicistronic constructs were co-eluted with StrepII-VirB4s, whereas non-tagged VirB8sp alone (negative control) showed only very weak binding to the column (Fig. 2.8).

To examine whether co-elution reflected a stable interaction between VirB4s and VirB8s/VirB8sp, the affinity column enriched samples were separated using two approaches. First, blue native-PAGE was used to follow the migration patterns of the two proteins. When VirB4s was co-expressed with VirB8s/VirB8sp, it primarily migrated in high molecular weight form (>440 kDa), with a minor portion detected in an even larger molecular mass range, whereas StrepII-VirB4s alone was exclusively found at ~440 kDa (Fig. 2.9). VirB8s and VirB8sp produced from the bicistronic constructs also co-migrated with StrepII-VirB4s in the high molecular weight range, while StrepII-VirB8sp alone (the sample was diluted 20-fold to facilitate electrophoresis and immunoblotting) was detected

as a monomer of approximately 20 kDa. These results indicated that VirB4s formed a homomultimer, and that VirB8s/VirB8sp associated with VirB4s as part of a larger complex. We also examined the interaction between VirB4s and VirB8s/VirB8sp using gel filtration. The co-expressed samples were separated using a Superdex 200 gel filtration column. Consistent with the blue native-PAGE results, StrepII-VirB4s was co-eluted with VirB8s and VirB8sp in fractions 8 (440 kDa) and 9 (Fig. 2.10A). In contrast, StrepII-VirB4s alone was only detected in fraction 9, while StrepII-VirB8sp alone was eluted in fractions with a much lower molecular weight (the exact value was not determined as it was beyond the resolution of Superdex 200 column). As a final test of the stability of the StrepII-VirB4s-VirB8s complex, the sample was applied onto a Superdex 75 column, which is designed to separate and detect smaller proteins like VirB8s. If VirB8s/VirB8sp was dissociated from StrepII-VirB4s, both proteins should be readily detected in fractions with a small size (<67 kDa), while StrepII-VirB4s should be eluted in the void volume, as large proteins do not enter the pores of the matrix. Both StrepII-VirB4s and VirB8s were only found in the void elution volume, indicating that they formed a large and stable complex (Fig. 2.10B). Collectively, these experiments showed that StrepII-VirB4s self-associates and forms a stable complex with VirB8s, and that the VirB8s periplasmic domain is sufficient for establishing this interaction.

2.3.4 Identifying VirB4-VirB8 interaction sites in VirB8s

To systematically dissect the VirB4-BirB8 binding interface, we set about creating VirB8sp variants, with the goal of identifying residues required for interaction with VirB4. Thirteen different variants were constructed using the available crystal structure of VirB8sp as a guide: Q144R, L151R, K182E, T201A/Y, R230D, I112R, W119A, Y126E, D152R, M102R, Y105R and E214R (Fig. 2.11) (Paschos *et al*, 2006). The 13 residues were selected from regions expected to be critical for protein-protein interactions. Mutants E214R, M102R and Y105R were designed to introduce charges into the dimer interface, and were predicted to interfere with self-association. Variants R230D and T201Y/A carried mutations at conserved residues on the solvent-exposed β-sheet face. Residues K182, L151, and Q144 were located in a conserved groove flanked by β4 strand and α3 helix. A similar groove was also found in nuclear transport factor 2, where it functions as an interaction hotspot (Paschos *et al*, 2006; Terradot *et al*, 2005). Although the residues in this groove are not conserved, K182, L151, and Q144 were all predicted to participate in protein binding, according to the crystal structure (Paschos *et al*, 2006). Amino acids Y126, D152, and W119 were located in a patch of high sequence identity among VirB8 homologs and I112R was created because its equivalent mutation in VirB8a (R107P) disrupted VirB8a-VirB9a/VirB10a interactions in a yeast two-hybrid study (Kumar & Das, 2001). The purified VirB8sp mutants were incubated in the presence of StrepII-VirB4s before being subjected to blue native-PAGE. In accordance with previous observations, StrepII-VirB4s and wild-type VirB8sp co-migrated as a high molecular mass complex (440 kDa) (Fig. 2.12A). Most of the variants retained their ability to interact with StrepII-VirB4s, except for VirB8sp^{R230D}, which was detected exclusively in the lower molecular mass range (< 67 kDa). When not incubated with StrepII-VirB4s, all

mutants migrated in a low molecular form (< 67 kDa) (Fig. 2.12C). This experiment suggested that the residue R230 in VirB8s is critical for VirB4-VirB8 interaction.

2.3.5 Disrupting VirB4a-VirB8a interaction by making equivalent mutation in VirB8a

An important aim of this thesis is to evaluate whether blocking VirB4-VirB8 interaction results in any structural or functional defects of T4SS *in vivo*. Because all functional assays were set up using *A. tumefaciens*, we needed to create a VirB8a mutant equivalent to VirB8s^{R230D} and validate its influence on VirB4-VirB8 interaction before proceeding to any functional analyses. To first assess the behaviour of the VirB8a mutant, we used a bacterial two-hybrid (BTH) system. There are a number of different BTH assay options, but the one that we chose to work with was based on the reconstitution, in an *E. coli* adenylate cyclase (CyaA) deficient strain, of two parts of the CyaA catalytic domain: T18 (1-224aa) and T25 (225-339aa) (Karimova *et al*, 1998). When CyaA is split into T18 and T25, it does not catalyze the synthesis of cAMP. However, if two interacting proteins are genetically fused to the two CyaA fragments respectively, T18 and T25 will be brought into close spatial proximity and catalyze the synthesis of cAMP. cAMP in turn activates the cAMP receptor protein Crp, triggering the transcription of reporter genes such as *lacZ* (Karimova *et al*, 1998). Genes encoding VirB4a, VirB8a, or their truncated variants were cloned either downstream of T18/T25 in the case of C-terminal fusions, or upstream of T18/T25 to create N-terminal fusions. Reciprocal combinations of interacting partners were tested since based on our experience, some VirB proteins, when fused to different Cya fragments (T18 or T25) display differential binding affinity. Complementary constructs encoding VirB4a or its truncated versions (198-401aa, 329-650aa, 237-789aa), and VirB8a or its periplasmic domain were co-electroporated into an *E. coli cya* strain. Double transformants were streaked on plates containing X-gal or were inoculated into liquid culture and assessed for β-galactosidase activity. We expected that if the two proteins interacted, the colonies should appear blue on X-gal plates, and their interaction affinity could therefore be quantified using β-galactosidase assays. A strain co-expressing T18-VirB10s and T25-VirB8s (whose interaction was established previously using both biochemical approaches and BTH assays) was examined as a positive control. Meanwhile, constructs encoding T18/25 fusions with our subjects (VirB4a, VirB8a, or their truncated variants) were paired with plasmids encoding T18/T25 alone as negative controls to eliminate non-specific interactions. As expected, VirB8s interacted with VirB10s, while the negative controls showed no interaction. Regarding the experimental groups, all possible combinations of VirB4a/VirB4s (full length, 198-401aa, 329-650aa, 237-789aa) and VirB8a/VirB8s (full length, periplasmic domain) were tested for interaction at 3 different temperatures (26°C, 30°C, 37°C) and one pair showed a positive result: T18-VirB4 (198-401aa) + T25-VirB8ap (Fig. 2.13). This is consistent with a previous yeast two-hybrid study showing the same VirB4a fragment (198-401aa) as a VirB8a binding partner (Ward *et al*, 2002).

Using the structures of both VirB8s and VirB8a as a guide, we determined that Q223 in VirB8a corresponded to R230 in VirB8s corresponds (Bailey *et al*, 2006; Terradot *et al*, 2005) and consequently changed Q223 to either Glu or Lys using site directed mutagenesis as they are "safer" substitutions for Gln (Bordo & Argos, 1991; Shenoy & Visweswariah, 2003). The plasmids encoding T25-VirB8ap$^{Q223E/K}$ were tested in the BTH system, in parallel to the wild-type VirB8a as described above. Both mutants exhibited drastically reduced interaction levels, relative to the wild-type VirB8ap, as indicated by the lack of color formation on X-Gal plates and little β-galactosidase production (Fig. 2.13). This suggested that the equivalent mutation is capable of inhibiting the interaction between the *Agrobacterium* VirB4 and VirB8 proteins.

2.3.6 The impact of VirB4-VirB8 interaction on T4SS assembly and substrate translocation

To investigate the importance of VirB4-VirB8 interaction in *A. tumefaciens* T4SS, both wild-type VirB8a and the mutants VirB8a$^{Q223E/K}$ were tested for their ability to complement a *virB8a* deletion strain CB1008. Since the beginning of the *virB8a* gene overlaps the end of the upstream *virB7a* gene, and tumor induction by an *A. tumefaciens virB8a* deletion strain was only restored by co-expression of VirB7a and VirB8a *in trans* (Berger & Christie, 1994), *virB7a* and *virB8a* were co-amplified and cloned into a pTrc200 vector for complementation experiments. Mutations Q223E/K were created using site-directed mutagenesis on pTrc*virB7avirB8a*. Upon virulence induction by AS (a phenolic compound that induces the *virB* operon of *A. tumefaciens*) on AB minimal medium, the wild-type strain C58, the *virB8a* deletion mutant CB1008 and the mutant carrying pTrc*virB8a/virB8a*$^{Q223E/K}$ were studied in parallel using a series of functional assays. First, the cellular abundance of the T4SS components was checked using SDS-PAGE and western blotting (Fig. 2.14). As expected, VirB8a was not detected in CB1008 but was expressed at native levels in all other strains, implying that VirB8a$^{Q223E/K}$ were as stable as the wild-type. In the absence of VirB8a, the levels of VirB3a and VirB6a were diminished, while VirB5a and VirB4a were slightly reduced. VirB7a was also present at a lower level in CB1008 than in the wild-type, which may be attributed to the deletion of the end of the *virB7a* gene. VirB5a and VirB7a were restored to the same extent by the wild-type and mutant VirB8a proteins *in trans*, although the mutants VirB8a$^{Q223E/K}$ were less effective at complementing the defects in VirB3a and VirB6a, particularly VirB8a^{Q223K}, the production of which, however, was stronger in complementing VirB4a levels. This experiment suggested that the contact between VirB4a and VirB8a has a stabilizing effects on VirB3a, VirB6a, and possibly VirB4a as well (Yuan *et al*, 2005).

Apart from protein stability examination, we also conducted experiments aimed at evaluating the effect of VirB8 mutations on the assembly of extracellular T-pili. T-pili were sheared off the *Agrobacterium* cell surface and were subjected to immunoblotting to probe the presence of the major pilus component VirB2a and the minor component VirB5a as shown for the wild-type strain control (Fig. 2.15) (Schmidt-Eisenlohr *et al*, 1999). Neither VirB2a nor VirB5a was present in the pilus preparation of CB1008,

suggesting a defect in T-pilus assembly (Fig. 2.15). Wild-type levels of both proteins were detected upon expression of either the wild-type VirB8a or its mutant variants**,** indicating that these mutations did not affect the assembly of extracellular pili.

Finally, we compared the efficiency of T-DNA transfer by the wild-type and mutant *virB8a* expressing strains. When used to infect *Kalanchoë diagremontiana* leaves, virulence induced wild-type strain led to the formation of a small tumor at the wound site 4 weeks after infection, whereas no tumor was observed without virulence induction for the wild-type strain, or by virulence-induced CB1008 (Fig. 2.16A). Complementation with wild-type VirB8a gave rise to a large tumor, while the two point mutant exhibited opposing phenotypes: the Q223E mutant failed to induce tumor formation, suggesting it was unable to complement under these circumstances, while the Q223K mutant stimulated the development of a large tumor, suggesting that it was as functional as the wild-type protein in this instance. The different effects were unexpected since the change from Gln to Glu is theoretically milder than to Lys with regards to protein conformation. This result was, however, not successfully reproduced, as all subsequent attempts to grow the plants failed, leading to the development of an alternative strategy for assessing conjugative plasmid transfer mobilized by *A. tumefaciens* T4SS. T4SS, when produced in the donor strain, is capable of directing the conjugal transfer of IncQ group plasmids; at the same time, when T4SS is expressed in the recipient strain, it also increases the efficiency of *A. tumefaciens* to accept IncQ plasmids (Beijersbergen *et al*, 1992; Bohne *et al*, 1998). In this study, the activity of the wild-type strain, the *virB8* deletion mutant and the mutant containing pTrc*virB7avirB8a* or pTrc*virB7avirB8a*$^{Q223E/K}$ to serve as donors during conjugation of the IncQ plasmid pLS1 was tested as a means of assessing the effect of VirB8a mutations on substrate translocation. The donor strains carrying pLS1 (carbenicillin resistant) were co-cultivated individually with the recipient *A. tumefaciens* strain (erythromycin resistant, expressing a native T4SS for plasmid uptake) under virulence-inducing conditions. In the absence of VirB8a, plasmid transfer efficiency was decreased to zero (Fig. 2.16B). Complementation with wild-type VirB8a resulted in conjugation efficiency greater than that of the wild-type strain (the transfer efficiency of the wild-type was regarded as 100%), while the two mutants VirB8a^{Q223E} and VirB8a^{Q223K} restored transfer efficiency to 65% and 54% respectively. This contradicts our observations in the tumor induction assay, where VirB8a^{Q223K}-producing strain produced a larger tumor. The conjugation assay results are considered more convincing here as they are more readily reproducible and quantifiable. Although the mutant-expressing strains exhibited lower transfer rates than the wild-type complemented cells, the transfer function was not completely abolished. In order to create a more drastic change, the residue Q223 was further mutated to Ala; however, the resulting protein was unstable upon expression.

Taken together, these results suggested that either the VirB4-VirB8 interaction was dispensable for T4SS function, or that mutagenesis of a single amino acid was not sufficient to fully block VirB4a-VirB8a interaction *in vivo*. In addition to Q223, other

surface exposed residues on the same β-sheet may need to be mutated to obtain a more drastic phenotype.

3.4 DISCUSSION

Since previous research proposed that VirB4 contributes to T4SS biogenesis and functions as an NTPase in coordination with other components (Yuan *et al*, 2005), in the work presented here, I set out to purify *B. suis* VirB4 as a recombinant protein and to test its ability to hydrolyze ATP. While StrepII tagged VirB4s did not display any enzymatic activity in our experiment, it was able to bind ATP. VirB4-VirB8 interaction was demonstrated via co-migration or co-fractionation of the two proteins during blue native-PAGE and size exclusion chromatography. The VirB8 residue required for this interaction was identified to be R230, by screening VirB8s mutants using blue native-PAGE and the corresponding residue in VirB8a was subsequently mutagenized and the disrupting effect of the VirB8a mutations was validated using the BTH assay. Towards our ultimate goal of evaluating the importance of VirB4-VirB8 interaction on T4SS activity, the VirB8a mutants were subjected to *in vivo* analyses to examine their ability to complement a *virB8a* deletion mutant.

Although the function and oligomerization state of VirB4 family proteins have been addressed by multiple groups before (Dang & Christie, 1997; Dang *et al*, 1999; Rabel *et al*, 2003), little is currently known about their membrane topology, atomic structure and NTP hydrolyzing activity, with the exception of TrwK encoded by the IncW plasmid R388, and TraB encoded by the IncN plasmid pKM101 (Arechaga *et al*, 2008; Durand *et al*, 2010). This is largely attributed to the difficulty in working with VirB-like proteins due to their low protein solubility (Christie, 2004), as was observed here as well. While another group has opted to extract the insoluble *A. tumefaciens* VirB4 protein from inclusion bodies using denaturing conditions (8 M urea), we preferred to use native protein to circumvent the problems associated with protein refolding (Lilie *et al*, 1998; Shirasu *et al*, 1994)**.** Instead, we focused on optimizing conditions to improve VirB4 solubility under overexpressing conditions. Benzyl alcohol was reported to increase membrane fluidity in *E. coli* in the same way as high temperature, and like during a heat shock response, benzyl alcohol exposure leads to elevated expression of endogenous chaperones including GroEL-GroES and DnaK-DnaJ-GrpE (Shigapova *et al*, 2005). These heat shock proteins are thought to facilitate the isolation of soluble proteins by preventing protein aggregation and enhancing the folding of nascent proteins. The presence of GroEL in the purified StrepII-VirB4s preparation was unexpected, as benzyl alcohol has been successfully applied to increase soluble protein yield without contaminating the samples and association of VirB4 with heat shock proteins had not been reported either previously (Marco *et al*, 2005). This could be due to the massive production of chaperones, which may associate with target proteins and co-purify along with them. By adding an additional anion exchange chromatography step right before gel filtration, GroEL was successfully eliminated from the final preparation, as evidenced by immunoblotting. Although it has been reported before that incubating protein samples in

the presence of ATP and Mg^{2+} prior to ion exchange chromatography can effectively disrupt any complexes formed between GroEL and target proteins (Zahrl *et al*, 2005), I did not detect any difference when ATP and Mg^{2+} were included.

ATPase activity has been reported for VirB4 homologs (Arechaga *et al*, 2008; Durand *et al*, 2010); however, when our soluble, purified VirB4s sample was examined for its ability to hydrolyze ATP, it did not show any higher activity than the negative control, an active site mutant of VirB4s (StrepII-VirB4s^{K464R}). It is possible that the buffer conditions used were not optimal, as salt selection seems to be critical for ATPase activity. Chloride salts, which were used in our NTPase buffer, have been subsequently reported to be much less favorable than acetate salts for ATP hydrolysis by TraB and TrwK (Arechaga *et al*, 2008; Durand *et al*, 2010). We also can not exclude the possibility that ATP hydrolysis by VirB4s may require other components (VirB8), substrates (VirE2, single stranded DNA) or membrane association, as has been shown for the other two T4SS ATPases, VirB11 and VirD4, whose activities are stimulated by membrane association and signal/double stranded DNA binding respectively (Krause *et al*, 2000b; Rivas *et al*, 1997; Tato *et al*, 2005). These possibilities were not exploited due to low protein yield and poor overall stability of StrepII-VirB4s. I did note, however, that wild-type StrepII-VirB4s could readily bind ATP, and this binding was preserved in the active site mutant. This is consistent with previous literature showing that an equivalent active site mutation in another ATPase also retained nucleotide binding ability (Sung *et al*, 1988). This characteristic may be attributed to the positive charge of both Lys (wild-type) and Arg (mutant), which promotes interaction with the negatively charged phosphoryl groups of ATP (Sung *et al*, 1988). Given the ATPase activity reported for other VirB4 like proteins (Arechaga *et al*, 2008; Durand *et al*, 2010), and the defects in substrate translocation that result from the K464R mutation in the Walker A box of VirB4a, we predict that VirB4s may function in some capacity as an energy provider in the T4SS, despite the lack of ATPase activity observed in our study (Yuan *et al*, 2005). Moreover, the three C-terminal α-helices, which inhibit the ATPase activity of TrwK, are also conserved in both *A. tumefaciens* VirB4 and *B. suis* VirB4, suggesting that stimulation of their activities may require deletion of their C-termini (Pena *et al*, 2011).

In addition to potentially acting as an energy supplier during substrate translocation, VirB4 also mediates the physical assembly of T4SS. As a member of a complex multi-component system, VirB4 relies on its direct contacts with other components to fully perform its function. To facilitate the biochemical study of VirB4-VirB8 interaction, we took advantage of bicistronic constructs, which allow for the co-expression of two proteins at similar levels. Using both gel filtration and blue native-PAGE, we determined that VirB4s assembles into a high order oligomer, predicted to be a homohexamer on the basis of size. I was also able to determine that VirB8s forms a stable complex with VirB4 through sites in its periplasmic domain, which strongly suggests that at least part of VirB4s is localized in the periplasm. Since the periplasm is devoid of ATP (Wulfing & Pluckthun, 1994), and ATP is predicted to be required for VirB4 function in the cell, VirB4 is most likely oriented in a way that its Walker A (433-440aa) and Walker B (631-

636aa) boxes are positioned in the cytoplasm. Using the BTH assay, the region of VirB4a interacting with VirB8a was determined. The experiment was, however, complicated by the fact the T25 and T18 fragments needed to be fused to protein regions in the cytoplasm (Borloo *et al*, 2007), and there was no clear consensus on VirB4 membrane topology, making construct design a challenge. To simplify the situation, both the VirB8a periplasmic domain (expressed as a cytoplasmic protein) and different segments of VirB4a were examined, along with the full length versions of both proteins. Out of all the possibilities tested, only the VirB4a region extending from amino acids 198 to 401 gave a positive result in the BTH assay, when combined with VirB8ap. Notably, interaction between VirB8ap and another VirB4a fragment extending from amino acids 329 to 650 identified via a yeast two-hybrid (YTH) study was not observed (Draper *et al*, 2006). This could be explained by the different nature of the two hosts (*E. coli* bacteria versus the yeast *Saccharomyces*). In the BTH system, protein interactions may be established in the cytoplasm, periplasm or even the cytoplasmic membrane of *E. coli* cells (the only requirement being that the two Cya domain fragments, T18 and T25, are present in the cytoplasm); however, in the YTH system, both interacting partners have to be translocated into the nucleus, where the environment can be very different from that of their native host (Bruckner *et al*, 2009). Given the association between VirB4a (198-401aa) and the periplasmic domain of VirB8a, at least a portion of the 198-401aa region of VirB4 should be present in the periplasm. This is, however, contradictory to a model of VirB4 topology constructed on the basis of alkaline phosphatase fusion studies, amino acid hydrophobicity and "positive inside" rule, where residues 131-440 are predicted to be localized to the cytoplasm (Fig. 2.1), and suggests that there may be an additional periplasmic domain found between the TMII (108-131aa) and the Walker A box (433-440aa) (Dang & Christie, 1997; von Heijne, 1992).

Having identified a VirB4 interaction domain, I then set about systematically assessing its association with VirB8 by screening VirB8sp mutants using blue native gel electrophoresis, and R230 of VirB8sp was identified as an important residue for VirB4-VirB8 interaction. The R230D substitution also had a major functional consequence since the *B. suis virB8* deletion mutant expressing VirB8s^{R230D} was severely impaired for its intracellular survival inside macrophages (Paschos *et al*, 2006). R230 is located on a conserved solvent exposed β sheet face opposite to the VirB8s dimer interface (Terradot *et al*, 2005). With a versatile side group, the interaction between Arg and a binding partner may involve all three main protein-protein interaction mechanisms: hydrogen bonding, electrostatic interactions and/or hydrophobic interactions (Jones & Thornton, 1996). Since the change from Arg to Asp completely abolished VirB4s-VirB8s interaction, R230 may primarily promote association with VirB4s through interaction with a negatively charged residue on the VirB4s surface. The β4 strand where R230 is located is on one edge of the β sheet face; however, it lacks the protective features of a typical edge β strand like short length and bulge structure (Hoskins *et al*, 2006). Such surface exposed unprotected edge β strands have been shown to have higher propensity to constitute protein-protein interfaces, further underlining the significance of the β sheet face in protein binding (Hoskins *et al*, 2006). Indeed, this is supported by the fact that

mutation of another residue on the same β sheet, T201, abolished the interaction between VirB8s and VirB10s (Paschos *et al*, 2006), suggesting that this interface may coordinate interactions with multiple T4SS components. Other potential interaction hotspots like the groove area and the dimer interface are not important for VirB4-VirB8 interaction.

We had, at this point, set up a system for effectively probing protein interactions in the T4SS from a functional perspective. By probing VirB4-VirB8 mutant interaction in *A. tumefaciens* as opposed to *B. suis*, I was able to circumvent the requirement for biosafety level 3 conditions (needed for *B. suis* experimentation). The cellular abundance of T4SS components is usually regarded as an important indicator of T4SS integrity. In *A. tumefaciens*, deletion of *virB8a* resulted in reduced cellular levels of VirB3a and VirB6a and similarly, loss of VirB4a lead to barely detectable levels of VirB3a and VirB8a (Yuan *et al*, 2005); therefore, it was inferred that VirB4a may modulate VirB3a stability through its interaction with, and maybe also the stabilization of, VirB8a. The diminishment of VirB6a in the absence of VirB8a but not VirB4a could be due to a very small quantity of VirB8a in the *virB4a* mutant cells, which may not be detectable on immunoblots (Yuan *et al*, 2005). The fact that VirB8a^{Q223K} seemed more potent than VirB8a^{Q223E} in complementing the VirB4a levels suggested that a positive charge at this site may reinforce its interaction with VirB4a (consistent with the nature of R230 in VirB8s).

Loss of VirB8a impaired T-pilus formation, tumor induction and conjugation of pLS1 by *A. tumefaciens* (as plasmid donor). All of these functional defects could be rescued by expressing VirB8a *in trans*. It was anticipated that the mutants VirB8a$^{Q223E/K}$, which failed to interact with VirB4a in BTH system, would be unable to complement the *virB8a* mutant phenotype. However, surprisingly, both mutants were as efficient as the wild-type protein in complementing T-pilus assembly, indicating that the extracellular structure of the T4SS was not adversely affected by either VirB8a mutation. In the conjugation assay, neither of the mutations completely abolished the activity of T4SS in plasmid delivery, raising the possibility that either VirB4-VirB8 interaction is not essential for T4SS functions or the substitutions, which blocked interaction in the BTH system, did not prevent interaction in the native host *A. tumefaciens*. Given the drastic phenotype of *B. suis* cells expressing the equivalent mutant (Paschos *et al*, 2006), the lack of any dramatic phenotypic effect for the mutants in our study is more likely due to reduced, but not abolished, VirB4-VirB8 interaction. VirB8s R230 and VirB8a Q223 differ in that the former is positively charged while the latter is merely polar. Arg has a versatile side chain capable of establishing intermolecular contacts via electrostatic interactions (positive charge), hydrogen bonding (through the guanidinum group) and hydrophobic effects (through the methylene group). When the complementary site on the other molecule is negatively charged, electrostatic interactions should be a major player, and in this scenario, mutation of an Arg to a negatively charged Asp will make the two sites repel each other, leading to complete inhibition of the interaction. Gln, on the other hand, has a high propensity to hydrogen bond, as its amide group can both accept and donate two hydrogen bonds. Mutation of Gln into either Glu or Lys may still preserve the hydrogen bonding between interacting sites of VirB4 and VirB8. The inconsistency between the

BTH assay and the functional assays may be explained by the fact that the VirB4a fragment used to establish VirB4-VirB8 interaction in BTH assay only covers ~ ¼ of the entire protein and this peptide may fold differently from the full length protein, and thus may not reflect the true effect of each mutation in its native context. Having this in mind, Q223 was further mutagenized to Ala in order to decrease the polarity of this site; however, this mutation destabilized the protein. Amongst the residues of the surface exposed β sheet face, Q223 was the only one selected for examination and this is largely based on the loss of co-migration of StrepII-VirB4s and VirB8sp in blue native-PAGE. Other neighbouring conserved residues like Tyr222, Ser225 and Asp227 may also participate in VirB4-VirB8 interaction. Mutations of these residues are worth investigating further in the future and will be coupled with more in-depth studies, *e.g.* extracting the cell membrane bound T4SS and analyzing the co-localization defects of the VirB components

2.5 FIGURES AND TABLES:

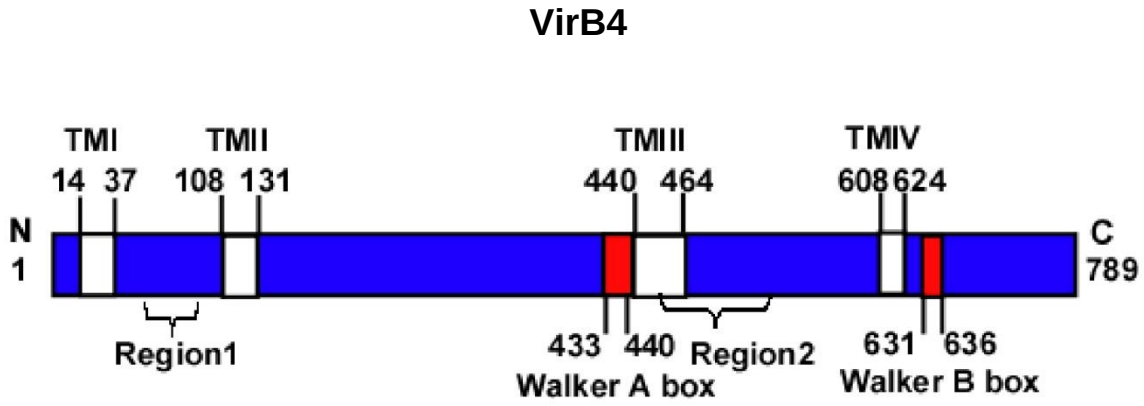

Fig. 2.1 Schematic representation of *A. tumefaciens* VirB4 (VirB4a). Studies using VirB4a and alkaline phosphatase fusion proteins (VirB4a::PhoA) revealed two periplasmic regions (Region 1: 58-84 aa; Region 2: 450-514 aa). TMI to TMIV indicate four potential transmembrane domains, based on the presence of a sufficient number of contiguous hydrophobic residues to span the cytoplasmic membrane, and on their proximities to junctions of active VirB4::PhoA fusion proteins. Walker A and Walker B boxes are shaded in red and are conserved motifs involved in NTP binding and hydrolysis. The model is based on work published by Dang & Christie (1997).

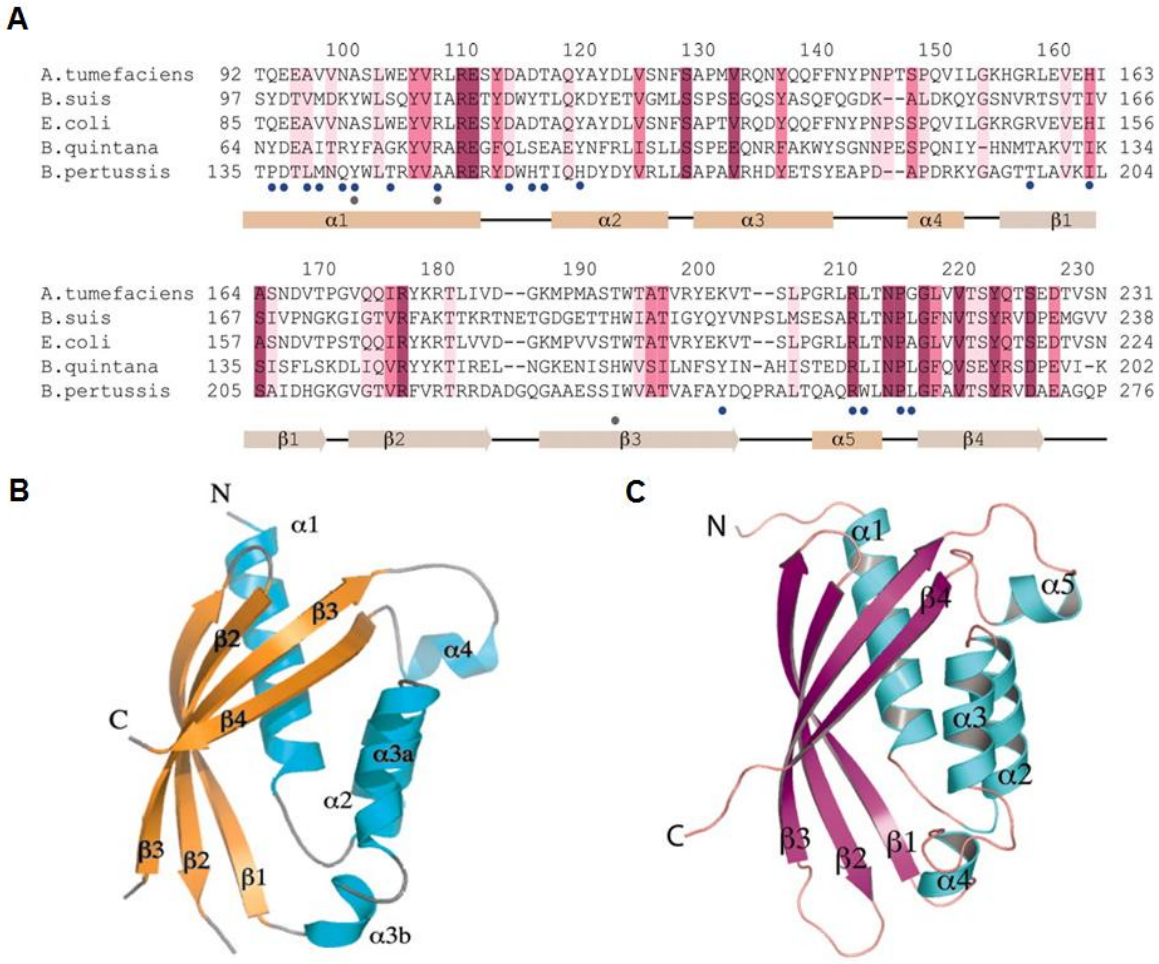

Fig.2.2 Multiple sequence alignment and crystal structures of VirB8 family proteins. (A) Multiple sequence alignment showing the conservation among VirB8 homologs. Conserved residues are shaded in pink. The higher the conservation level, the darker the color. The secondary structural elements are shown underneath the alignment. The blue dots indicate residues involved in dimerization while grey dots mark residues essential for activity. (B) and (C) are X-ray structures of *A. tumefaciens* VirB8 and *B. suis* VirB8 periplasmic domains respectively. Diagrams are adapted from Terradot *et al*. (2005) and Bailey *et al* (2006).

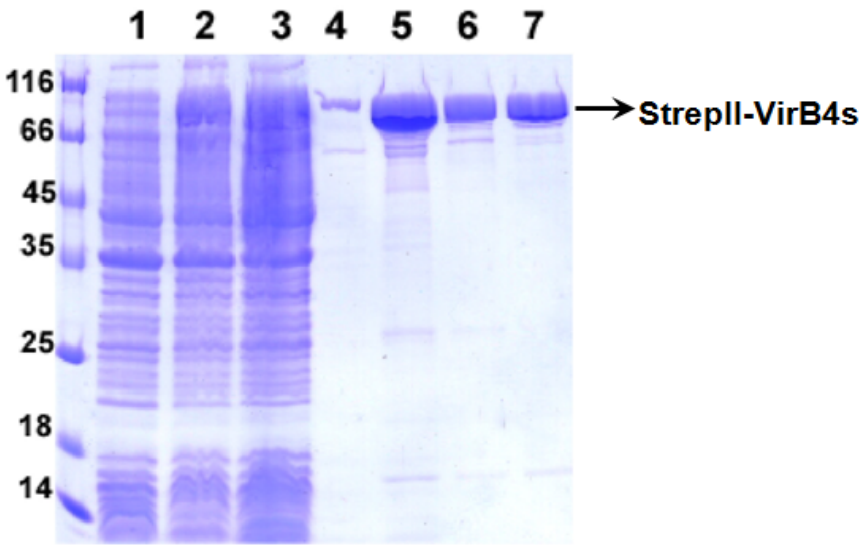

Fig. 2.3 Purification of StrepII-VirB4s. StrepII-VirB4s was overproduced in *E. coli* cells and was subjected to streptavidin affinity chromatography, enriched with ultrafiltration and purified using gel filtration chromatography. 1, cell extract before induction; 2, cell extract after induction; 3, soluble cell extract following cell lysis; 4, eluate from affinity column; 5, concentrated eluate from affinity column; 6, eluate from gel filtration; 7, final concentrated sample. Molecular weights of reference proteins are shown on the left.

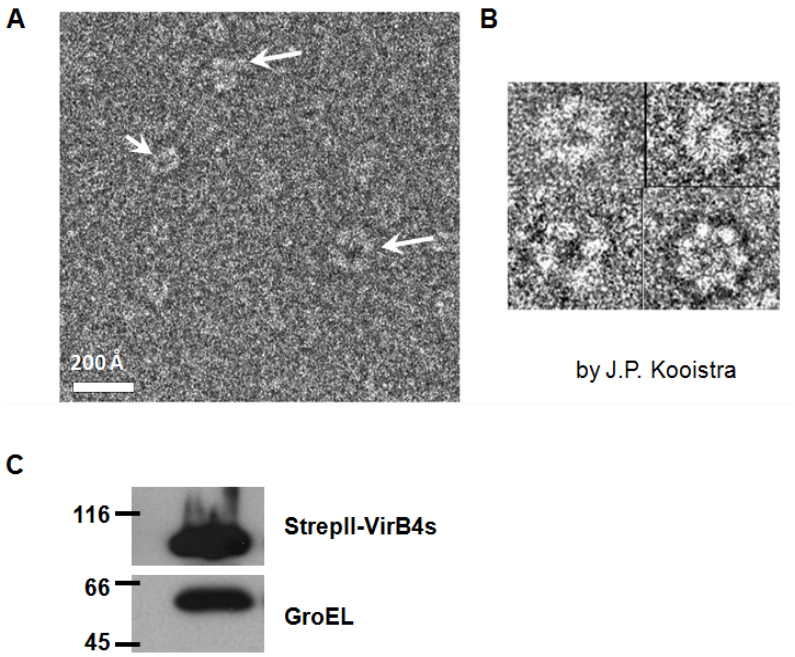

Fig. 2.4 Purified StrepII-VirB4s is contaminated by GroEL. (A) and (B) Electron micrographs of purified StrepII-VirB4s following negative staining. Microscopy was conducted by J.P. Kooistra. (C) Detection of GroEL in purified StrepII-VirB4s preparations using immunoblotting.

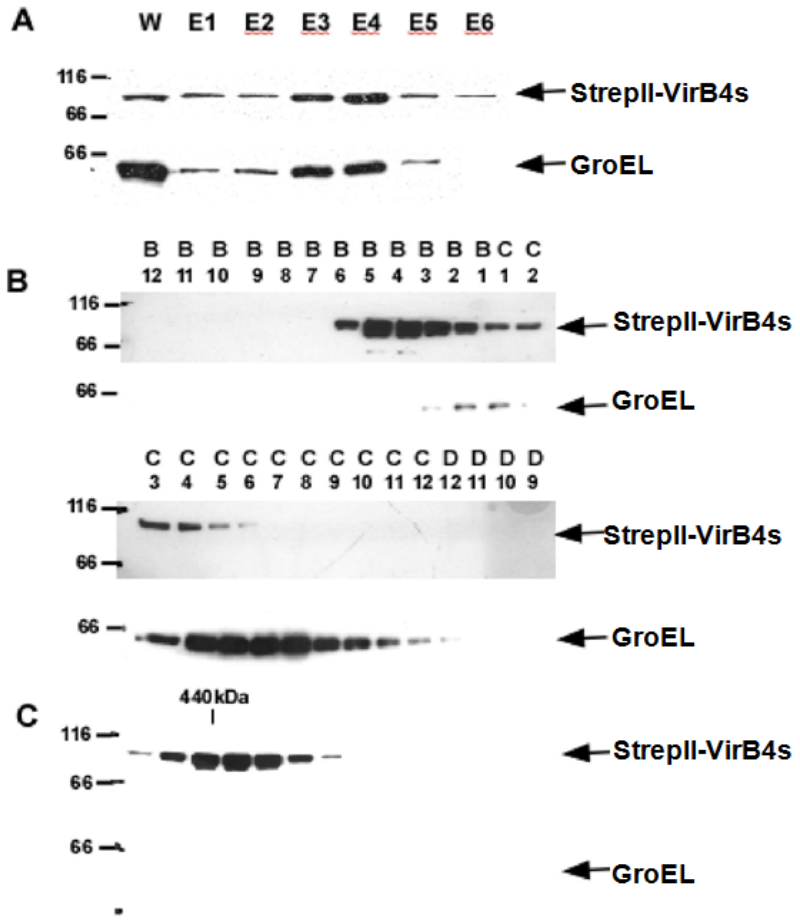

Fig. 2.5 Removal of contaminant GroEL from purified StrepII-VirB4s. The cell extract was subjected to streptavidin affinity chromatography, enriched using ultrafiltration, and purified using anion exchange chromatography and gel filtration chromatography. Elution fractions at each step were analyzed using SDS-PAGE and western blotting with anti-Strep and anti-GroEL antibodies. (A) Streptavidin affinity chromatography. W: the last washing fraction; E1-E6: elution fractions 1-6. (B) Anion exchange chromatography on a HiTrap Q Sepharose Fast Flow column. The numbers on top indicate the elution fractions. (C) Gel filtration chromatography on a Superdex 200 column. Molecular weights of reference proteins are shown on the left (for western blotting) and on the top (for the corresponding size of elution fractions).

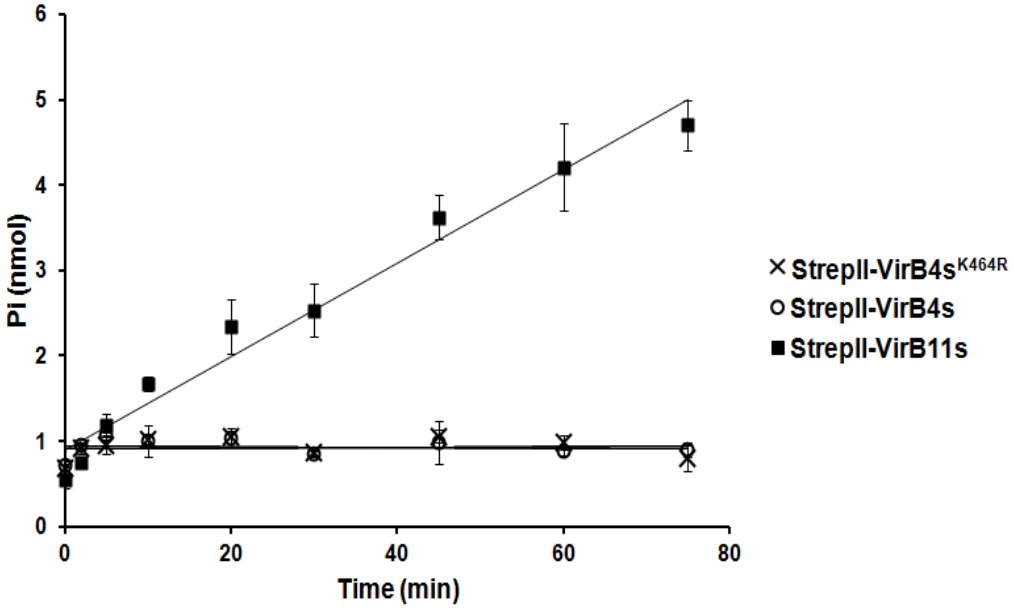

Fig.2.6 ATPase activities of StrepII-VirB4s and its active site variant StrepII-VirB4s^{K464R} using malachite green assay (progress curve of the reaction is shown). StrepII-VirB11s was used as a positive control. Fifty micromolar of each protein was used in the assay. The amount of the released phosphate upon ATP hydrolysis was followed for up to 75 min. The error bars designate standard deviations between four replicates.

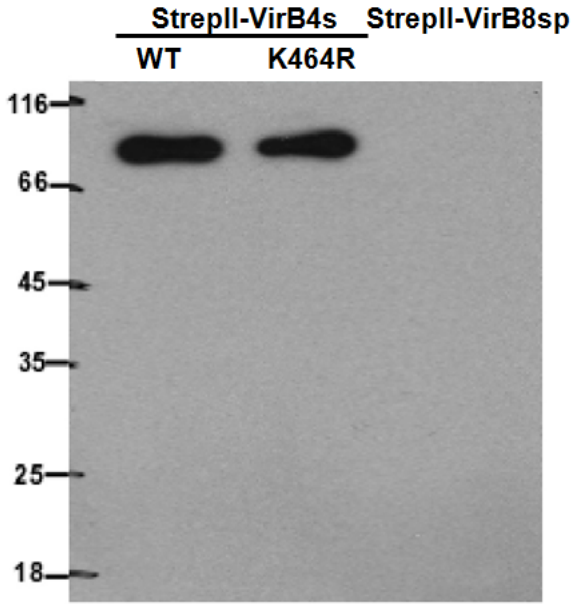

Fig. 2.7 ATP-binding activity of StrepII-VirB4s. StrepII-VirB4s and its active site mutant StrepII-VirB4S^{K464R}, and StrepII-VirB8sp (negative control) were incubated with ATP-agarose before the beads were washed and the bound proteins were eluted. The eluted samples were examined with SDS-PAGE and western blotting using anti-Strep antibodies. Molecular weights of reference proteins are shown on the left. The blot is representative of three experiments.

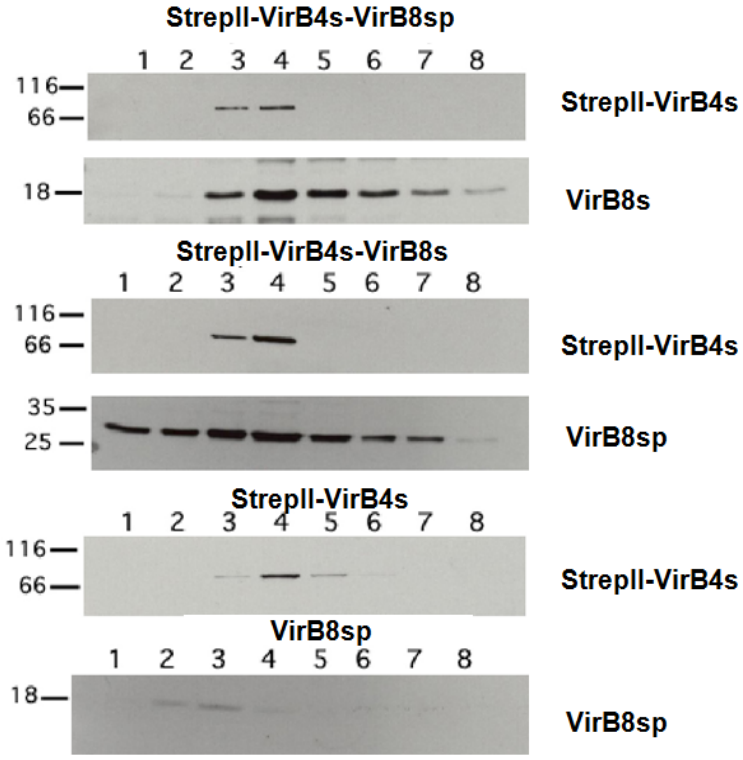

Fig.2.8 VirB4s and VirB8s/VirB8sp were co-eluted in affinity chromatography. Cell extract from *E. coli* strains overexpressing StrepII-VirB4s and VirB8s/VirB8sp from bicistronic constructs was passed through a Streptavidin Sepharose affinity column. Elution fractions (fraction number is shown above the blots) were analyzed using SDS-PAGE and western blotting. The two bottom blots show the elution profiles of StrepII-VirB4s alone (positive control) and VirB8sp alone (negative control) from affinity column. Molecular weights of reference proteins are shown on the left. The blots are representative of three experiments.

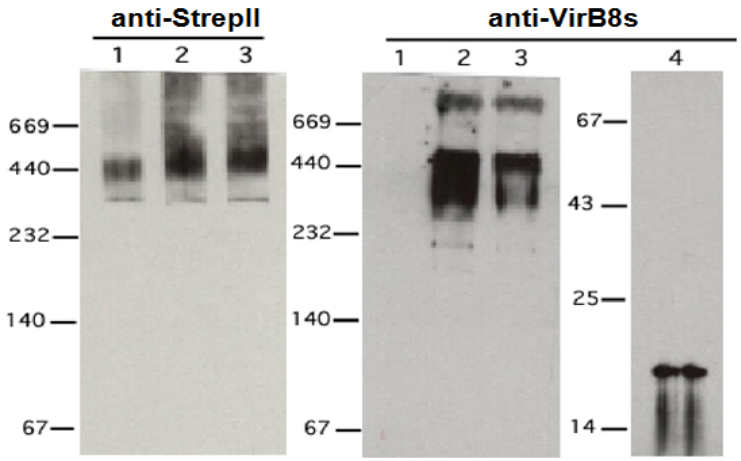

Fig. 2.9 Blue native-PAGE showing VirB4s-VirB8s interaction. Streptavidin enriched samples, StrepII-VirB4s (lanes 1), StrepII-VirB4s-VirB8s (lanes 2), StrepII-VirB4s-VirB8sp (lanes 3), and StrepII-VirB8sp (lane 4), were subjected to blue native gel electrophoresis followed by western blotting using both anti-StrepII and anti-VirB8s antibodies. Molecular weights of reference proteins are shown on the left. The blots are representative of three experiments.

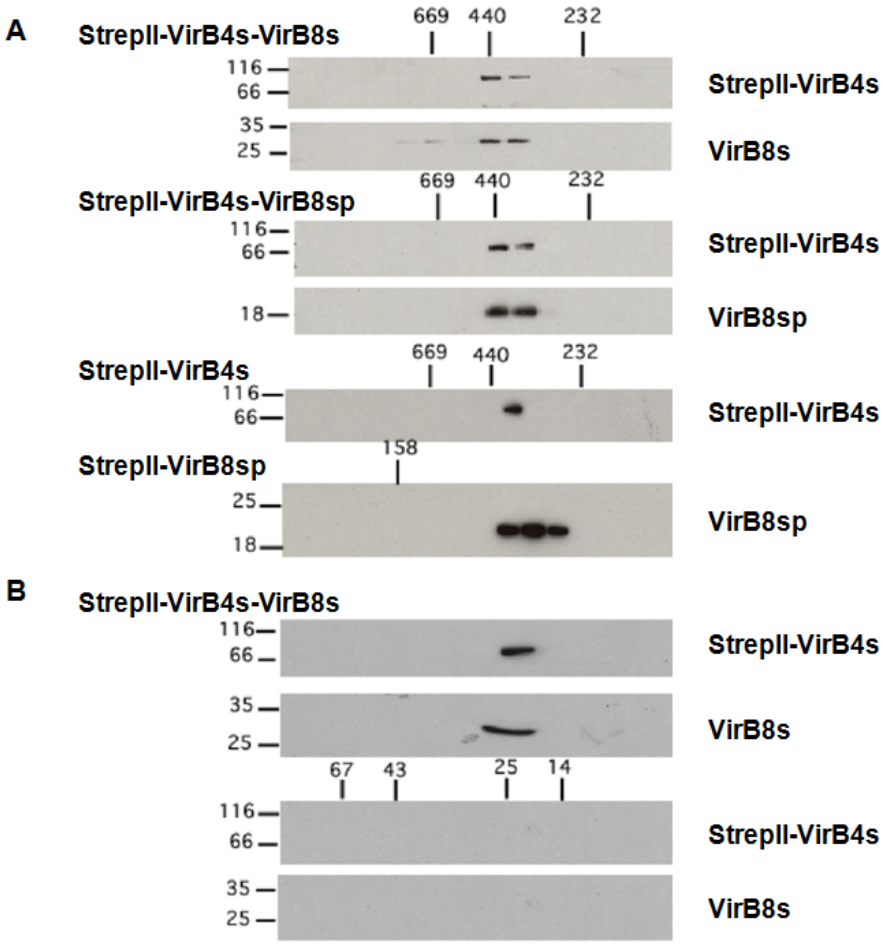

Fig. 2.10 Gel filtration analysis of VirB4s oligomer and its interaction with VirB8s. Affinity column enriched samples were loaded onto (A) Superdex 200 and (B) Superdex 75 for gel filtration chromatography. Eluted fractions were analyzed using SDS-PAGE and western blotting with anti-Strep and anti-VirB8s antibodies. Molecular weights of reference proteins are shown on the left (for western blotting) and on top (for corresponding size of elution fractions). The blots are representative of three experiments.

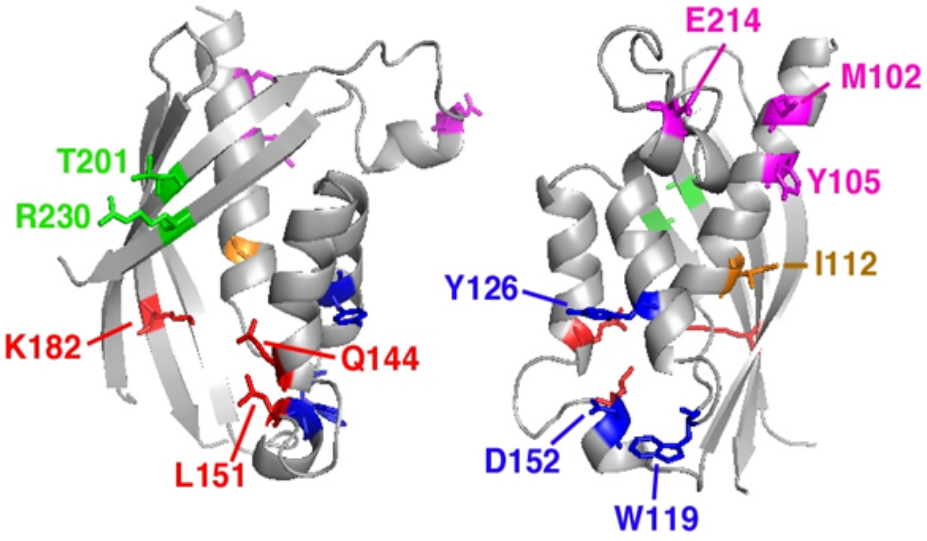

Fig. 2.11 Ribbon representation of the VirB8sp crystal structure in different orientations showing mutagenized amino acids. Dimer interface residues E214, M102, and Y105 are shown in magenta; active-site groove residues K182, L151, and Q144 are shown in red; β-sheet's solvent-exposed face residues R230 and T201 are shown in green; high identity patch Y126, D152, and W119 are shown in blue; and I112 is shown in brown. The diagram was adapted from Paschos *et al* (2006).

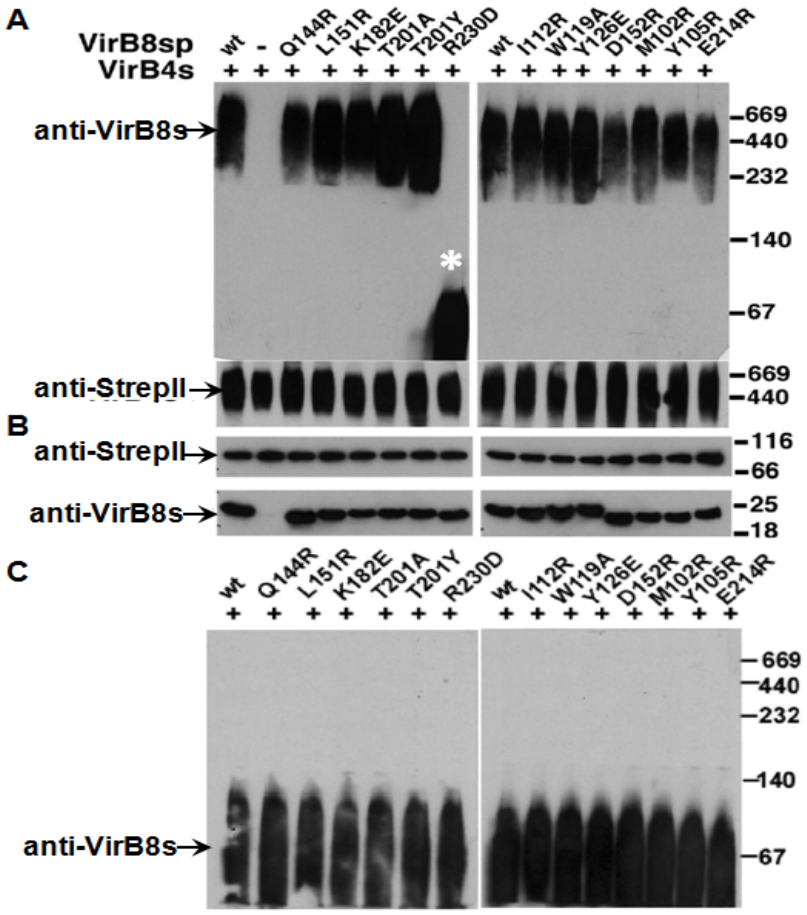

Fig. 2.12 Disruption of VirB4s-VirB8s interaction by a single mutation in VirB8sp. Purified VirB8sp variants were either incubated individually (C) or mixed with affinity column-enriched StrepII-VirB4s (A) before being resolved on a blue native gel, followed by western blotting using anti-StrepII and anti-VirB8s antibodies. The same samples were also analyzed using SDS-PAGE and western blotting to ensure similar protein loading in each lane (B). Molecular weights of reference proteins are shown on the right. The blots are representative of three experiments.

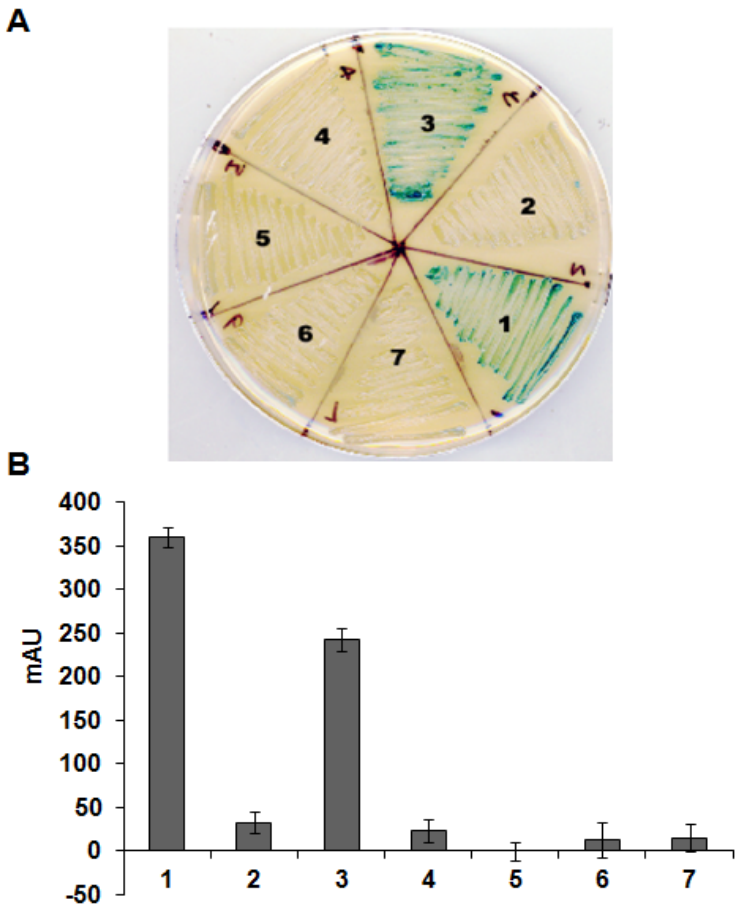

Fig. 2.13 Disruption of VirB4-VirB8 interaction using the BTH system. (A) X-Gal plate: blue color indicates positive interaction. (B) β-galactosidase assay. In both (A) and (B), the numbers indicate plasmids tested as follows: 1, pKT25*virB8s* + pUT18C*virB10s* (positive control); 2, pKT25 + pUT18C (vector alone negative control); 3, pKT25*virB8ap* + pUT18C*virB4a198-401*; 4, pKT25*virB8ap*^{Q223E} + pUT18C*virB4a198-401*; 5, pKT25*virB8ap*^{Q223K} + pUT18C*virB4a198-401*; 6, pKT25*virB8ap* + pUT18C (negative control); 7, pKT25 + pUT18C*virB4a198-401* (negative control). The error bars in (B) indicate standard deviations from three independent experiments.

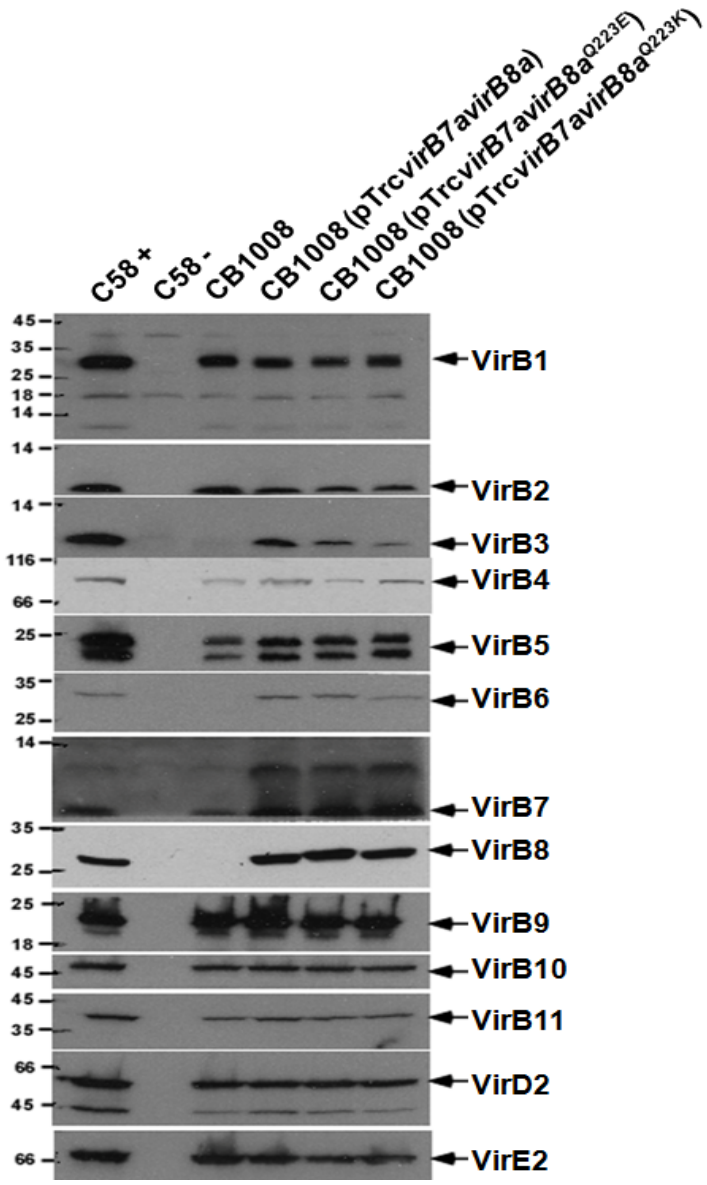

Fig. 2.14 Cellular levels of T4SS components in the wild-type strain C58, the *virB8a* deletion mutant strain CB1008, and the CB1008 strains expressing wild-type VirB8a and its variants. Samples were examined using SDS-PAGE and western blotting. "+" indicates that the C58 cells were induced with AS. "-" indicates that the C58 cells were not induced with AS. The rest of the strains were subjected to virulence induction before cell harvest. Molecular weights of reference proteins are shown on the left. The blots are representative of three experiments.

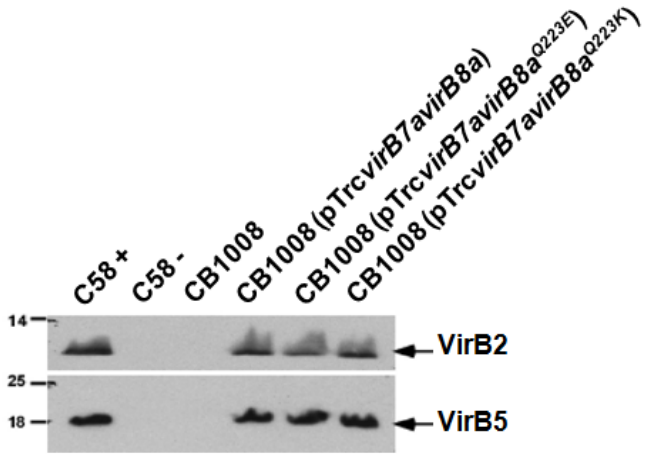

Fig. 2.15 Pilus preparation of the wild-type strain C58, the *virB8a* deletion mutant strain CB1008, and the CB1008 strain expressing wild-type VirB8a and its variants. Pili were sheared off of the cell surface and sedimented using ultracentrifugation. Samples were analyzed with SDS-PAGE and western blotting for the presence of the major component VirB2a and the minor component VirB5a. "+" indicates that the C58 cells were induced with AS. "-" indicates that the C58 cells were not induced with AS. The rest of the strains were subjected to virulence induction before sample harvest. Molecular weights of reference proteins are shown on the left. The blots are representative of three experiments.

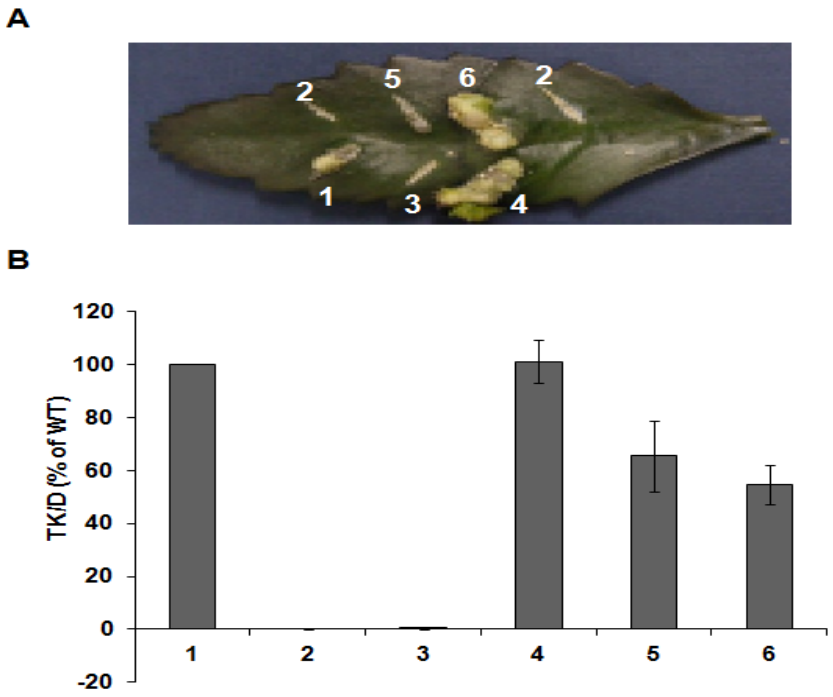

Fig. 2.16 Impact of VirB8a mutations on substrate translocation by *A. tumefaciens* T4SS. (A) Tumor induction elicited by T-DNA transfer from the *virB8a* deletion mutant expressing VirB8a variants. Virulence-induced *A. tumefaciens* strains were inoculated onto scratches on *Kalanchoe daigremontiana* leaves and tumors were observed after growth for 4 weeks. The strains used to infect plants were as follows: 1 wild-type strain C58 (virulence induced); 2 wild-type strain C58 (not induced); 3, *virB8a* deletion mutant CB1008 carrying the empty vector pTrc200; 4, CB1008 (pTrc*virB7avirB8a*); 5, CB1008 (pTrc*virB7avirB8a*^{Q223E}); 6, CB1008 (pTrc*virB7avirB8a*^{Q223K}). (B) Donor activity of the *virB8a* deletion mutant expressing VirB8 variants in the conjugation of the IncQ group plasmid pLS1. The donor strains used in conjugation were as follows: 1, wild-type strain A348 (virulence induced); 2, wild-type strain A348 (not induced); 3, *virB8a* deletion mutant PC1008 carrying the empty vector pTrc200; 4, PC1008 (pTrc*virB7avirB8a*); 5, PC1008 (pTrc*virB7avirB8a*^{Q223E}); 6, PC1008 (pTrc*virB7avirB8a*^{Q223K}). TK: the number of transconjugants; D: the number of donors; WT: wild-type. The ratio TK/D was calculated for each strain and their percentages, compared to the WT, were plotted in (B). Error bars indicate the standard deviations between four independent experiments.

Table 2.1 Bacterial strains used in the study

Strains	Genotype or description	References
Escherichia coli		
JM109	F' (*tra*D36 *pro*AB+ *lac*IqZΔM15) *rec*A1 *end*A1 *gyr*A96 *thi*-1 *hsd*R17 *rel*A1 *sup*E44 Δ(*lac-pro*AB)	Yanisch-Perron *et al*., 1985
GJ1158	*omt*T *hsd*S *gal dcm* Δ*mal*Ap510 *mal*P::(*pro*Up-T7 RNAP) *mal*Q::*lacZhyb*11 Δ(*zhf*-900::Tn10dTet)	Bhandari *&* Gowrishankar, 1997
BTH101	F^- *cya-99*, *araD139, galE15, galK16, rpsL1 (Str^r), hsdR2, mcrA1, mcrB1*	Karimova *et al*., 2005
TOP10	F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *deo*R *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G	Invitrogen
Agrobacterium. tumefaciens		
C58	Wild-type strain carrying the nopaline producing Ti plasmid, pTiC58.	van Larekeke *et al*. 1974
CB1008	C58 carrying an in frame deletion of *virB8* on pTiC58	Aly *et al*. 2008
A348	Wild-type strain carrying the octopine producing Ti plasmid, pTiA6NC.	Van Larebeke *et al*., 1974
PC1008	A348 carrying an in frame deletion of *virB8* on pTiA6NC	Berger & Christie, 1994
UIA143 pTiA6	A348 derivative, *recA*, ery^R	Bohne *et al*., 1998

Table 2.2 Plasmids used in this study

Plasmid	Description	References
pT7-7*strepII*	Carb^R, T7 promoter controlled expression vector carrying a 50 bp *Nco*I fragment encoding StrepII-Tag and factor Xα cleavage-site	Balsinger *et al*. 2004
pT7-7*strepII-virB4s*	pT7-7StrepII carrying a 2.5kb *Acc65*I/*EcoR*I *virB4* fragment amplified from *B. suis*	Yuan *et al*. 2005
pT7-7*strepII-virB4s-virB8s*	Bicistronic construct with a 720bp *EcoR*I/*Pst*I *virB8s* fragment cloned downstream of *virB4s* in pT7-7StrepII-VirB4s	Yuan *et al.* 2005
pT7-7*strepII-virB4s-virB8sp*	Bicistronic construct with a 492bp *EcoR*I/*Pst*I fragment encoding the periplasmic domain of VirB8s cloned downstream of *virB4s* in pT7-7StrepII-VirB4s	Yuan *et al.* 2005
pT7-7*strepII-virB8sp*	pT7-7StrepII carrying a 492bp *Acc65*I/*Pst*I fragment encoding VirB8s periplasmic domain	Terrado *et al.,* 2005
pKT25	Kan^R, derivative of pSU40 allowing fusion at the C-terminus of the T25 fragment of *B. pertussis* adenylate cyclase in BTH assay	Karimova *et al*. 2005
pKNT25	Kan^R, derivative of pSU40 allowing fusion at the N-terminus of the T25 fragment of *B. pertussis* adenylate cyclase in BTH assay	Karimova *et al*. 2005
pUT18C	Carb^R, derivative of pUC19 allowing fusion at the C-terminus of theT18 fragment of *B. pertussis* adenylate cyclase in BTH assay	Karimova *et al*. 2005
pUT18	Carb^R, derivative of pUC19 allowing fusion at the N-terminus of the T18 fragment of *B. pertussis* adenylate cyclase in BTH assay	Karimova *et al*. 2005
pUT18C*virB4s*	Carb^R, a 2.5kb *Xba*I/*Kpn*I *virB4s* fragment cloned downstream of the sequence encoding the T18 fragment	This study
pKT25*virB4s*	Kan^R, a 2.5kb *Xba*I/*Kpn*I *virB4s* fragment cloned downstream of the sequence encoding the T25 fragment	This study
pUT18*virB4s*	Carb^R, a 2.5kb *Xba*I/*Kpn*I *virB4s* fragment cloned upstream of the sequence encoding the T18 fragment	This study
pKNT25*virB4s*	Kan^R, a 2.5kb *Xba*I/*Kpn*I *virB4s* fragment cloned upstream of the sequence encoding the T25 fragment	This study
pUT18C*virB8s*	Carb^R, a 720bp *Xba*I/*Kpn*I *virB8s* fragment cloned downstream of the sequence encoding the T18 fragment	by Sivanesan, D.
pKT25*virB8s*	Kan^R, a 720bp *Xba*I/*Kpn*I *virB8s* fragment cloned downstream of the sequence encoding the T25 fragment	by Sivanesan, D.
pUT18CvirB8sp	Carb^R, a 492bp *Xba*I/*Kpn*I *virB8sp* fragment cloned downstream of the sequence encoding the T18 fragment	This study
pKT25*virB8sp*	Kan^R, a 492bp *Xba*I/*Kpn*I *virB8sp* fragment cloned downstream of the sequence encoding the T25 fragment	This study
pUT18*virB8sp*	Carb^R, a 492bp *Xba*I/*Kpn*I *virB8sp* fragment cloned upstream of the sequence encoding the T18 fragment	This study
pKNT25*virB8sp*	Kan^R, a 720bp *Xba*I/*Kpn*I *virB8sp* fragment cloned upstream of the sequence encoding the T25 fragment	This study

pUT18C*virB4a*	Carb^R, a 2.4kb *Xba*I/*Kpn*I fragment encoding pTiC58 VirB4 cloned downstream of the sequence encoding the T18 fragment	This study
pKT25*virB4a*	Kan^R, a 2.4kb *Xba*I/*Kpn*I fragment encoding pTiC58 VirB4 cloned downstream of the sequence encoding the T25 fragment	This study
pKT25*virB8ap*	Kan^R, a *Xba*I/*Kpn*I fragment encoding the periplasmic domain of pTiC58 VirB8 cloned downstream of the sequence encoding the T25 fragment	This study
pUT18C*virB8ap*	Carb^R, a *Xba*I/*Kpn*I fragment encoding the periplasmic domain of pTiC58 VirB8 cloned downstream of the sequence encoding the T18 fragment	This study
pUT18C*virB8ap*	Carb^R, a *Xba*I/*Kpn*I fragment encoding the periplasmic domain of pTiC58 VirB8 cloned downstream of the sequence encoding the T18 fragment	This study
pKT25*virB8ap*	Kan^R, a *Xba*I/*Kpn*I fragment encoding periplasmic domain of pTiC58 VirB8 cloned downstream of the sequence encoding the T25 fragment	This study
pTrc200	Str^R, spc^R, pVS1 origin, *lacI^q*, *trc* promotor expression vector	Schmidt-Eisenlohr *et al*., 1999
pTrc*virB7avirB8a*	Str^R, spc^R, a *EcoR*I/*BamH*I fragment encoding pTiC58 VirB7-VirB8 cloned downstream of the Trc promoter	This study
pfullB	Carb^R, carrying the full *virB* region from *A. tumefaciens* pTiC58	by C. Baron
pLS1	Carb^R, IncQ plasmid, used for VirB/D4 mediated conjugative transfer experiment	Stahl *et al.* 1998

Table 2.3 Growth media and buffers used in this study

Medium	Composition
LB	1% tryptone; 0.5% yeast extract; 0.5% NaCl
LBON	1% tryptone; 0.5% yeast extract
YEB	1% tryptone; 0.5% sucrose; 0.1%; 2mM $MgSO_4$
ABMM	1% glucose; 0.39% MES; 1 mM Na-K-phosphate pH 5.5; 1 x AB salts (20 x AB salts: 20 g NH_4Cl; 6 g $MgSO_4$ x $7H_2O$; 3 g KCl; 0.2 g $CaCl_2$; 50 mg $FeSO_4$ x $7H_2O$ per liter; pH 5.5)
Buffer	**Composition**
Buffer W (washing)	100 mM Tris/Cl pH 8.0; 1 mM EDTA; 150 mM NaCl; stored at 4°C
Buffer E (elution)	Buffer W with 2.5 mM D-desthiobiotin; stored at 4°C
10 × NTPase buffer	0.5 M Tris/Cl pH 7.5; 500 mM KCl; 20 mM $MgCl_2$; 10 mM DTT
Solution A	0.045% Malachite green hydrochloride
Solution B	4.2% Ammonium Molybdate in 4 M HC
Solution C	35% Tri-sodium citrate
Na-K-P buffer (0.5 M)	Make 500 ml 0.5 M KH_2PO_4; adjust to pH 5.5 by adding 0.5 M Na_2HPO_4
Blue native gel buffer (3 ×)	1.5 M 6-aminocaproic acid; 150 mM BisTris; adjust to pH 7.0 by adding what?; stored at 4°C
Blue native-PAGE anode buffer	100 mM BisTris; adjust to pH 7.0 by adding what?; stored at 4°C
Blue native-PAGE cathode buffer	100 mM tricine; 30 mM BisTris; adjust to pH 7.0 by adding what?; stored at 4°C
5% Coomassie blue	0.5 g Coomassie blue G 250; 0.656 g 6-aminocaproic acid
Membrane destaining buffer	25% methanol; 10% acetate; 65% H_2O
4 × Lower Tris	1.5 M Tris/Cl pH 8.8; 0.4% SDS
4 × Upper Tris	0.5 M Tris/Cl pH 6.8; 0.4% SDS
Staining buffer	30% methanol, 10% acetate; 0.1% Coomassie blue R250
Gel destaining buffer	20% methanol; 8% acetate
10 × Laemmli running buffer	0.25 M Tris; 1.92 M glycine; 1% SDS
2 × Laemmli sample buffer	0.125 M Tris/Cl; 4% SDS; 20% glycerol; 10% β-mercaptoethanol; a few grains of bromophenol blue
Blotting buffer	192 mM glycine; 25 mM Tris; 20% methanol
Schägger gel buffer 3 ×	3 M Tris/Cl; 0.3% SDS; adjust to pH 8.45
Schägger gel anode buffer 5 ×	1 M Tris/Cl; adjust to pH 8.9
Schägger gel cathode buffer 5 ×	0.5 M Tris/Cl; 0.5 M tricine; 5% SDS; adjust to pH 8.25
ONPG buffer (9 ml)	8.1 ml Z-buffer; 0.9 ml 8 mg/ml ONPG in Z-buffer; 11.25 µl 10% SDS; 26.64 µl β-mercaptoethanol (99.77%)
Z-buffer (500 ml)	8.05 g $Na_2HPO_4 \bullet 7H_2O$; 2.75 g NaH2PO4•H_2O; 0.375 g KCl; 0.123 g $MgSO_4 \bullet 7H_2O$; adjust to pH 7.0
Stop buffer	1 M Na_2CO_3

ACA buffer	750 mM 6-aminocaproic acid; 50 mM BisTris; adjust to pH 7.0 by adding NaOH

Table 2.4 Oligonucleotides used in the study.

Primer name	Sequence *
Inverse PCR primers	
VirB4sK464R5’	5'-TGGT**AGA**ACTGTCCTGATGAACTTCTGCCTCGC-3’
VirB4sK464R3’	3'-ACAGT**TCT**ACCAGCGCCGGACTGGCCGATGATG-3’
VirB8aQ223E5’	5’-TCCTAC**GAG**ACATCGGAAGATACCGTTTCG-3’
VirB8aQ223E3’	5’-CGATGT**CTC**GTAGGAGGTGACAACCAAGCC-3’
VirB8aQ223K5’	5’TCCTAC**AAG**ACATCGGAAGATACCGTTTCG-3’
VirB8aQ223K3’	5’-CGATGT**CTT**GTAGGAGGTGACAACCAAGCC-3’
VirB8aQ223A5’	5’-TCCTAC**GCG**ACATCGGAAGATACCGTTTCG-3’
VirB8aQ223A3’	5’-CGATGT**CGC**GTAGGAGGTGACAACCAAGCC-3’
VirB8aQ223R5’	5’-TCCTAC**GAC**ACATCGGAAGATACCGTTTCG-3’
VirB8aQ223R3’	5’-CGATGT**GTC**GTAGGAGGTGACAACCAAGCC-3’
VirB8aQ223D5’	5’-TCCTAC**CGG**ACATCGGAAGATACCGTTTCG-3’
VirB8aQ223D3’	5’-CGATGT**CCG**GTAGGAGGTGACAACCAAGCC-3’
Plasmid construction for BTH assay	
Ct-VirB4s5’	5’-GCGGTCTAGACGGGATGATGGGCGCTCAATCCAAATAC-3’
Ct-VirB4s3’	5’-GCGGGGTACCTCACCTTCCTGTTGATTTGGACGACGCAAT-3’
VirB4s3’-Nt	5’-GCGCGGTACCCCCCTTCCTGTTGATTTGGAC-3’
Ct-VirB8sp5’	5’-GCGGTCTAGACGGGCGCGTCAACGCACAGACAGGT-3’
Ct-VirB8sp3’	5’-GCGGGGTACCTCATTGCACCACTCCCAT-3’
VirB8sp3’-Nt	5’-GCGGGGTACCCCTTGCACCACTCCCATTTC-3’
Ct-VirB4a5’	5’-GCGGTCTAGACGGGATGCTCGGAGCAAGTGGCACA-3’
Ct-VirB4a3’	5’-GCGGGGTACCTCAATCTTTTGCCTCGTGGTAACG-3’
Ct-VirB4a(198-401)5’	5’-GCGGTCTAGACGGGCGATTCACGCCGGTTCCCGTC-3’
Ct-VirB4a(198-401)3’	5’-GCGGGGTACCTCAGTGACCTGAGCCGGATCCCTCGGG-3’
Ct-VirB4a(329-650)5’	5’-GCGGTCTAGACGGGGATCTCAATAATCTTGCAGAT-3’
Ct-VirB4a(329-650)3’	5’-GCGGGGTACCTCAGTCGACGACCGCCGCAAATTTAGG-3’
Ct-VirB4a(237-789)5’	5’-GCGGTCTAGACGGGTCTATTTACTCGTTTCGCGAA-3’
Ct-VirB8ap5’	5’-GCGGTCTAGACGGGACCATGGTGCCGCTGATCAGG-3’
Ct-VirB8ap3’	5’-GCGGGGTACCTCATGGTTCGCTGTGGCCTGC-3’
Plasmid construction for *in vivo* complementation	
VirB7VirB85’	5’-CGCGTGAATTCATGAAATATTGCCTG-3’
VirB7VirB83’	5’-CGCGTGGATCCTCATGGTTCGCTGTGGCC-3’

* Mutagenized codons are bolded. Restriction sites are underlined.

APPENDIX:

APPENDIX 1: FUNCTIONAL ANALYSIS OF A VIRB4-LIKE PROTEIN IN *ACTINOBACILLUS ACTINOMYCETEMCOMITANS*

A1.1 INTRODUCTION

As a Gram negative facultative anaerobe, *Actinobacillus actinomycetemcomitans* is the causative agent of localized juvenile periodontitis, severe adult periodontitis, and other diseases like thyroid and brain abscesses and urinary tract infections (Zambon *et al*, 1988). In order to identify *A. actinomycetemcomitans* antigens that induce T cell-mediated destructive immunity, a genomic library of *A. actinomycetemcomitans* was subjected to expression cloning screening using periodontitis-associated $CD4^+$ T cells as probes (Teng & Hu, 2003). A gene encoding a protein of 38.6 kDa was identified to be recognized by T-cells (Teng & Hu, 2003). The C-terminus of the protein (137-340aa) shares sequence similarity with VirB4 family proteins, CagE of the T4SS encoded by the *Helicobacter pylori cag* pathogenecity island (39% identity) and VirB4 of *A. tumefaciens* (27% identity), and thus has been designated 'CagE' (Teng & Hu, 2003). CagE is localized in the cytoplasm of *A. actinomycetemcomitans* and, when secreted to the extracellular milieu, targets the surface of host cells and elicits apoptosis of primary human epithelia, endothelia, osteoblasts, and T cells (Teng & Hu, 2003; Teng & Zhang, 2005). Truncation of the N-terminus of CagE drastically reduces the apoptosis level of target cells (Teng & Hu, 2003). Despite its homology to the VirB4 family proteins, CagE does not contain any Walker A or Walker B boxes, which are required for NTP hydrolysis by other VirB4-like proteins. Interestingly, its N-terminus contains a soluble lytic transglycosylase (SLT) domain (ESSYNPYAIAVVNDIPLAQQPKTLQ) (Teng & Hu, 2003), which is also found within the N-terminus of VirB1 of *A. tumefaciens* T4SS and is required for peptidoglycan degradation to enable T4SS assembly in the periplasmic space (Baron *et al*, 1997) (Fig. A1.1). To analyze the relationship between CagE and T4SS, the ability of CagE to complement *A. tumefaciens virB1a* and *virB4a* deletion strains was assessed.

A1.2 MATERIALS AND METHODS

A1.2.1 Bacterial strains, plasmids, oligonucleotides and culture conditions

Strains and plasmids used in this study are listed in Tables A1 and A2. Media, buffers and other materials are listed in Table A3. Oligonucleotides are listed in Table A4. For bacterial cultivation and *A. tumefaciens* virulence induction, please refer to Section 2.2.1 of Chapter 2. pVSBAD*virB4a* containing strains were cultured in ABMM supplemented with glycerol instead of glucose as a carbon source and 0.2% arabinose was used to

induce protein expression as described previously (Yuan *et al*, 2005). Antibiotics were added to maintain plasmids as described in Section 2.2.1 as well.

A1.2.2 Plasmid construction

Manipulation of DNA including plasmid extraction, PCR amplification, digestion, ligation and sequencing were conducted following standard procedures (Maniatis *et al*, 1982). Restriction enzymes were purchased from New England Biolabs and MBI Fermentas and *E. coli* JM109 was used as a cloning host. The complementing construct pTrc*cagE* was created by amplifying the *cagE* gene from pQE30*cagE*, followed by digestion with *Nco*I/*Sal*I and cloning into similarly digested pTrc200.

A1.2.3 SDS-PAGE and western blotting

Please refer to Section 2.2.4 in Chapter 2.

A1.2.4 T-pilus isolation

Please refer to Section 2.2.10 in Chapter 2.

A1.2.5 Conjugation assay

Please refer to Section 2.2.11 in Chapter 2.

A1.3 RESULTS AND DISCUSSION

A1.3.1 Complementation of the *virB4*a deletion strain by CagE

pTrc*cagE* was transformed into the *virB4a* deletion mutant, followed by analysis using functional assays. A previous study suggested that the levels of VirB3a and VirB8a were diminished in the absence of VirB4a and that these defects could be rescued by overexpressing both wild-type VirB4a and VirB4s from plasmids (Yuan *et al*, 2005). In order to determine whether CagE had a similar function to these proteins, cell extracts of the wild-type strain C58, the *virB4a* mutant CB1004 containing the empty plasmid pTrc200 (negative control), and the CB1004 strains carrying pVSBAD*virB4* or pTrc*cagE*, were analyzed using SDS-PAGE and western blotting. VirB4a produced from pVSBAD*virB4a* stabilized VirB3a and VirB8a as reported before, while the CagE producing strain showed no detectable VirB3a and a very small quantity of VirB8a (Fig. A1.2). VirB8a was also seen at a lower level in the sample obtained from the uninduced CB1004 (pTrc*cagE*), which could be caused by leakage from the *trc* promoter of pTrc vector (Oehler *et al*, 1990), since no VirB8a was detected in the plasmid alone-containing strain. I also examined T-pilus assembly: T-pili were isolated from the virulence-induced strains as described previously to test for the presence of the major T-pilus component VirB2a and the minor component VirB5a (Schmidt-Eisenlohr *et al*, 1999). Neither VirB2

nor VirB5 was detected in the strain producing CagE, while VirB4a production complemented T-pilus formation (Fig. A1.3). Finally, T4SS-mediated IncQ group plasmid pLS1 transfer between agrobacteria was employed to test substrate translocation. The donor cells (wild-type A348 strain and *virB4* deletion strains PC1004 containing pTrc*cagE*, pVSBAD*virB4*, or pTrC200) carrying pLS1 were mated with the recipient cells UIA143 (pTiA6) on ABMM plates to enable plasmid transfer. The production of wild-type VirB4a could partially restore native levels of plasmid transfer, while the CagE-producing strain behaved like the negative control (Fig. A1.4). These functional assays suggested that, despite their sequence similarity, CagE is not likely to act in the same way as VirB4a during T4SS assembly and substrate translocation. This is perhaps not surprising, since CagE is a cytoplasmic protein without any nucleotide-binding site, while VirB4 is associated with the inner membrane and is postulated to drive substrate translocation by hydrolyzing NTP.

A1.3.2 Complementation of the *virB1a* deletion strain by CagE

The construct expressing CagE was also transformed into the *virB1a* deletion mutant PC1001, followed by investigation using the same assays described in A1.3.1. Consistent with previous research, expression of both VirB1a and VirB1s restored T-pilus formation in the *virB1a* deletion mutant while CagE did not (Fig. A1.5). Donor activities of the strains carrying pLS1 in conjugation were also assessed. The conjugation efficiency of the *virB1a* deletion mutant was minimized and VirB1a and VirB1s expressed from plasmids restored conjugation levels to 98% and 53% respectively (Fig. A1.6). The CagE producing strain also exhibited a conjugation rate of around 95%. Collectively, CagE may substitute for VirB1 in conjugation (substrate transfer) but not in extracellular structure assembly. As discussed in Chapter 1, the N-terminal lytic transglycosylase domain of VirB1 facilitates T4SS biogenesis and macromolecule transfer by degrading the peptidoglycan cell wall (Christie *et al*, 2005; Hoppner *et al*, 2004). The active site Glu60 is conserved in CagE and this might be crucial for CagE's possible function as a cell wall degrading enzyme during conjugation. Since CagE only shows N-terminal homology to VirB1 family proteins, the lack of pilus formation might be due to the absence of the C-terminal domain of VirB1a, which ends up being processed to VirB1a* (refer to Chapter 1 for the background information on VirB1a*) and is indispensable for pilus assembly (Zupan *et al*, 2007).

The data obtained thus far are insufficient for us to draw any firm conclusions about the functional relatedness shared by CagE and T4SS components. Although CagE has some similar features to VirB1 and VirB4, the functional analyses do not fully support a model whereby CagE plays comparable roles to those of the two VirB proteins. Additionally, due to the poor quality of the CagE-specific antiserum (which even failed to react with purified CagE protein), we were unable to determine whether CagE was expressed or not. Within the 6 kb flanking region of the *cagE* gene, there are at least four other ORFs encoding proteins similar to components of the *H. pylori cag* pathogenicity island: CagC (similar to VirB2), CagD, CagH, and CagM, suggesting this locus may

comprise a *cag*-like pathogenicity island (Teng & Hu, 2003). Given the impact of CagE on eukaryotic cells, researchers are currently inclined to believe that CagE is a toxin delivered through a Cag-like secretion system and if this is true, further research is needed to characterize the adjacent *cag*-like genes of *A. actinomycetemcomitans*.

A1.4 FIGURES

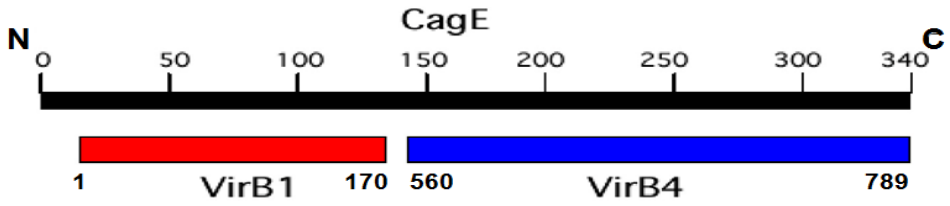

Fig. A1.1 A schematic model showing the homology between *A. actinomycetemcomitans* CagE and *A. tumefaciens* VirB1/ VirB4.

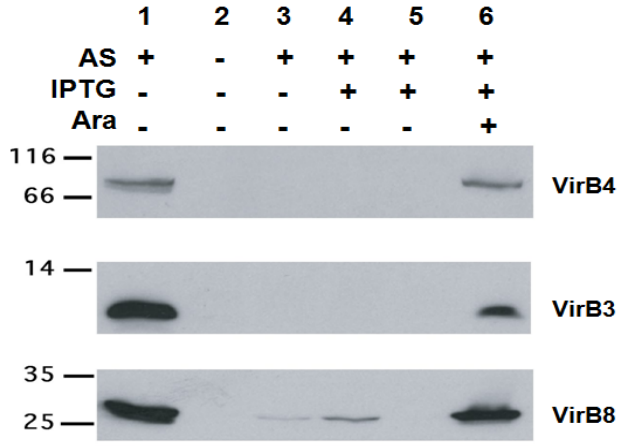

Fig. A1.2 Cellular levels of VirB3a and VirB8a in the *virB4a* deletion mutant carrying a CagE expressing construct. 1, the wild-type strain C58 (virulence induced); 2, C58 (not induced); 3, the *virB4a* deletion mutant CB1004 carrying pTrc*cagE* (CagE expression not induced); 4, CB1004 carrying pTrc*cagE* (CagE expression induced); 5, CB1004 carrying the empty vector pTrc200; 6, CB1004 carrying pBAD*virB4a* (VirB4a expression induced). AS: acetosyringone; IPTG: isopropyl-β-D-thiogalactopyranoside; Ara: arabinose. Molecular weights of reference proteins are shown on the left. Blots are representative of three independent experiments.

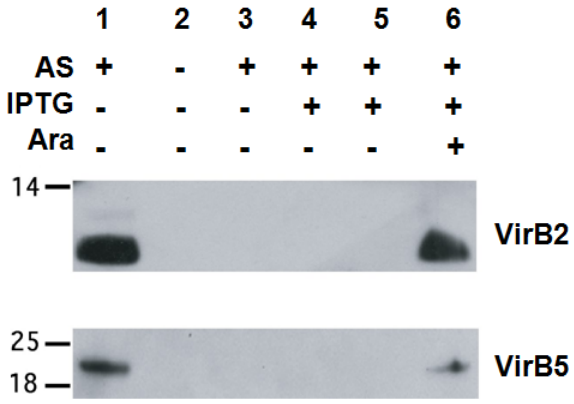

Fig. A1.3 T-pilus formation in the *virB4a* deletion mutant expressing CagE. 1, wild-type strain C58 (virulence induced); 2, C58 (not induced); 3, *virB4a* deletion mutant CB1004 carrying pTrc*cagE* (CagE expression not induced); 4, CB1004 carrying pTrc*cagE* (CagE expression induced); 5, CB1004 carrying the empty vector pTrc200; 6, CB1004 carrying pBAD*virB4a* (VirB4a expression induced). AS: acetosyringone; IPTG: isopropyl-β-D-thiogalactopyranoside; Ara: arabinose. Molecular weights of reference proteins are shown on the left. Blots are representative of three independent experiments.

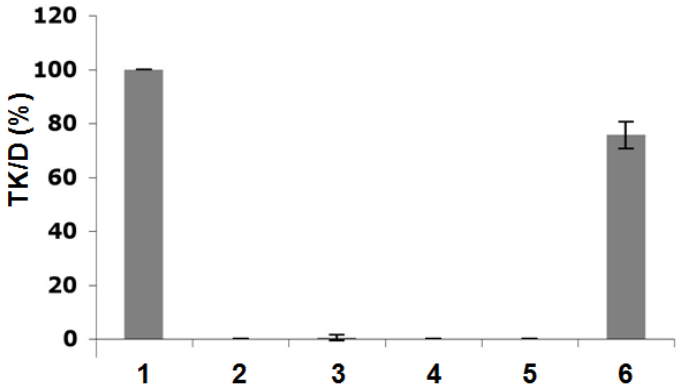

Fig. A1.4 Donor activity of the *virB4a* deletion mutant expressing CagE in the conjugative transfer of the IncQ group plasmid pLS1. Donor strains tested were: 1, wild-type strain A348 (virulence induced); 2, wild-type strain A348 (not induced); 3,*virB4a* deletion mutant PC1004 carrying pTrc*cagE* (CagE expression not induced); 4, PC1004 carrying pTrc*cagE* (CagE expression induced); 5, PC1004 carrying the empty vector pTrc200; 6, PC1004 carrying pBAD*virB4a* (VirB4a expression induced). TK/D = number of transconjugants / total number of donors. The results of strains 2-6 were expressed as the percentages of their TK/D vs. that of the virulence induced wild-type strain. Error bars indicate the standard deviations from three independent experiments.

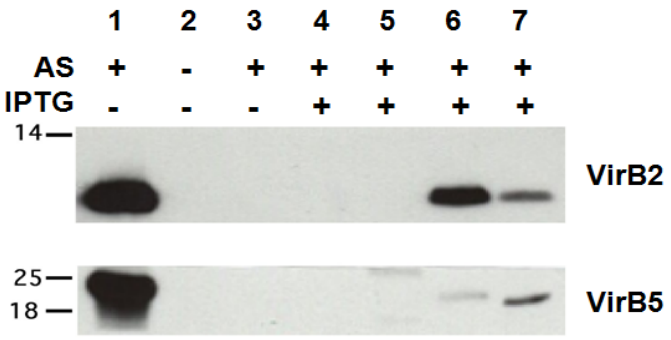

Fig. A1.5 T-pilus formation in the *virB1a* deletion mutant expressing CagE. 1, wild-type strain C58 (virulence induced); 2, C58 (not induced); 3, *virB1a* deletion mutant PC1001 carrying pTrc*cagE* (CagE expression not induced); 4, PC1001 carrying pTrc*cagE* (CagE expression induced); 5, PC1001 carrying the empty vector pTrc200; 6, PC1001 carrying pTrc*virB1a* (VirB1a expression induced); 7, PC1001 carrying pTrc*virB1s* (VirB1s expression induced). AS: acetosyringone; IPTG: isopropyl-β-D-thiogalactopyranoside. Molecular weights of reference proteins are shown on the left. Blots are representative of three independent experiments.

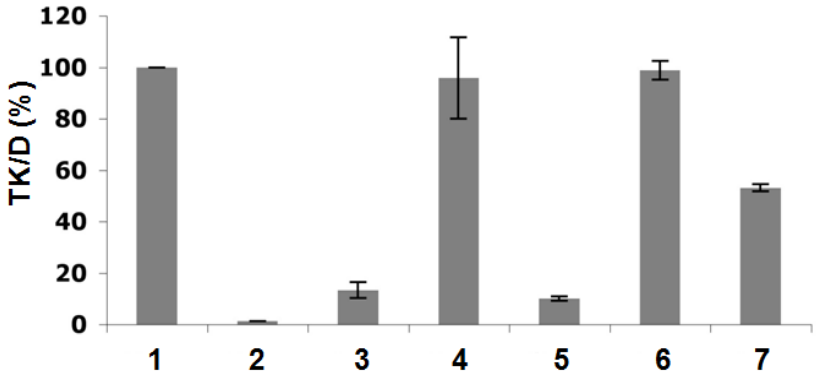

Fig. A1.6 Donor activity of the *virB1a* deletion mutant expressing CagE in the conjugative transfer of the IncQ group plasmid pLS1. Donor strains tested were: 1, wild-type strain A348 (virulence induced); 2, wild-type strain A348 (not induced); 3, *virB1a* deletion mutant PC1001 carrying pTrc*cagE* (CagE expression not induced); 4, PC1001 carrying pTrc*cagE* (CagE expression induced); 5, PC1001 carrying the empty vector pTrc200; 6, PC1001 carrying pTrc*virB1a* (VirB1a expression induced); 7, PC1001 carrying pTrc*virB1s* (VirB1s expression induced). TK/D = the number of transconjugants / the total number of donors. The results of strains 2-7 were expressed as the percentages of their TK/D vs. that of the virulence induced wild-type strain. Error bars indicate the standard deviations from three independent experiments.

APPENDIX 2: CO-FRACTIONATION OF VIRB PROTEINS IN THE *A. TUMEFACIENS VIRB3* AND *VIRB8* DELETION STRAINS

A2.1 INTRODUCTION

As described in Chapter 1, the membrane bound T4SS of *A. tumefaciens* can be extracted using the mild detergent dodecyl-β-D-maltopyranoside (DDM) (Krall *et al*, 2002). Similar procedures have been successfully applied to analyze the protein complex of mitochondrial respiratory chain (Cruciat *et al*, 2000). Due to its mild nature, DDM does not disrupt the integrity of VirB/D4 complexes; instead, it promotes the formation of soluble micelles containing proteins, phospholipids, and detergent (Cruciat *et al*, 2000). Adaptation of this technique for investigations into our system allowed us to detect two subassemblies of the VirB/D4 complex upon separation using gel filtration chromatography and blue native-PAGE (Krall *et al*, 2002). The small complex mainly comprises the pilus components VirB2a and VirB5a, and the high molecular mass complex contains the core components VirB6a through VirB10a. VirB6a, VirB7a and VirB8a were also found to associate with the small complex (Krall *et al*, 2002; Yuan *et al*, 2005). However, *virB4a* deletion perturbed the association of VirB6a and VirB7a with the high molecular weight complex and the co-fractionation of VirB2a and VirB5a in the small complex (Yuan *et al*, 2005). In the following section, the *virB3a* and *virB8a* deletion strains were characterized using DDM extraction coupled with blue native-PAGE and gel filtration chromatography, to evaluate the importance of VirB3a and VirB8a in T4SS subcomplex assembly.

A2.2 MATERIALS AND METHODS

A2.2.1 Bacterial strains, plasmids, oligonucleotides and culture conditions

Strains and plasmids used in this study are listed in Table A1 and A2. Media, buffers and other materials are listed in Table A3. Oligonucleotides are listed in Table A4. For bacteria cultivation, please refer to Section 2.2.1 in Chapter 2.

A2.2.2 SDS-PAGE and western blotting

Please refer to Section 2.2.4 in Chapter 2.

A2.2.3 Cell membrane isolation and VirB complex extraction

Agrobacteria were grown on 15 cm ABMM plates and virulence induction was performed as described in Section 2.2.1. Cells were washed off plates using 50 mM Na-K-P buffer and then collected by centrifugation at 10,000 × g at 4°C for 1 h. Bacteria from 16 plates were resuspended in 20 ml Na-K-P buffer before 3 passages through a French press cell (AMINCO) at 18,000 psi in the presence of 0.5 mM PMSF. Cell debris

was removed by centrifuging at 12,000 × g at 4°C for 1 h. The membrane fraction was isolated by ultracentrifuging the supernatant in a Ti-90 rotor and an Optima L-90K preparative ultracentrifuge (Beckman Coulter) at 136,720 × g at 4°C for 2 h. The membrane pellets were resuspended in 1 ml ACA (aminocaproic acid) buffer. The total protein concentration of each sample was adjusted to 10 mg/ml. DDM was added in the protein sample to get a final concentration of 2% and the mixtures were incubated at 4°C for 2 h with mild shaking. Ultracentrifugation was performed again at 136,720 × g at 4°C for 2 h to separate DDM-soluble and –insoluble proteins.

A2.2.4 Blue native-PAGE

Please refer to Section 2.2.8 in Chapter 2 for a detailed procedure of blue native-PAGE. The subsequent immunoblotting was performed using anti-VirB2a and anti-VirB5a antisera (Schägger & Jagow, 1991).

A2.2.5 Gel filtration chromatography

Five hundred microliters of DDM-solubilized sample was loaded on a Superdex 200 10/300 GL column (GE Healthcare). Chromatography was performed as specified in Section 2.2.8 in Chapter 2. The elution fractions were subjected to SDS-PAGE and western blotting using VirB protein-specific antibodies.

A2.3 RESULTS AND DISCUSSION

Before the co-fractionation of VirB proteins was examined, their cellular abundance in the *virB3a* deletion mutant CB1003 and the *virB8a* deletion mutant CB1008 strains were analyzed using SDS-PAGE and western blotting (Fig. A2.1). Deletion of *virB8a* or *virB3a* did not affect the accumulation of most VirB components, although there were a few exceptions in both cases. In the *virB8a* deletion strain CB1008, VirB3a and VirB6a were not detected, suggesting that accumulation of both proteins relies on VirB8a. As VirB3a and VirB8a were both diminished in a *virB4a* deletion mutant, it is possible that VirB4a stabilizes VirB3a through VirB8a (Yuan *et al*, 2005). In the *virB3a* mutant strain (CB1003), VirB6a was drastically decreased, while VirB9a was also detected at a lower level, indicating that VirB6a stabilization also partly depends on VirB3a. Blue native-PAGE of the DDM-extracted samples revealed that VirB2a and VirB5a co-fractionated in a low molecular weight complex of about 100 kDa in both deletion mutants (like in the wild-type), although the signals were weaker than in the wild-type strain (Fig. A2.2). In gel filtration experiments, proteins were eluted in ACA buffer containing 0.03% DDM, a concentration higher than the critical micellar concentration so as to prevent decreased protein solubility following detergent dilution during elution. For the wild-type strain, VirB4a, VirB9a, VirB10a and VirB11a were co-eluted in the fractions with a molecular mass of over 440 kDa, while VirB2a, VirB3a and VirB5a were mainly detected in fractions having a lower molecular weight of ~200 kDa (Fig. A2.3). VirB6 and VirB8a were associated with both subcomplexes. VirE2 (relaxase excising T-DNA) and VirD2

(single-stranded DNA binding protein) were detected in a wider range of fractions. In the *virB8a* mutant (CB1008) strain, although VirB3a, VirB6a and VirB8a were absent, VirB4a, VirB9a, VirB10a and VirB11a were still present in the high molecular mass range, and VirB2a and VirB5a were still co-eluted in the lower molecular mass range (Fig. A2.4A). The *virB3a* mutant strain exhibited very similar co-fractionation pattern (Fig. A2.4B).

Taken together, deletion of neither *virB8a* nor *virB3a* changed the membrane association and co-fractionation of other components compared to the wild-type strain. However, functional studies (Chapter 2, section 2.3.6) showed that VirB8 is indispensable for T-pilus formation and plasmid conjugation. It therefore appears that, instead of being directly involved in subcomplex formation, VirB8 may be more important for incorporating the VirB2a-VirB5a complex into the T-pilus, and the subsequent substrate translocation.

A2.4 FIGURES

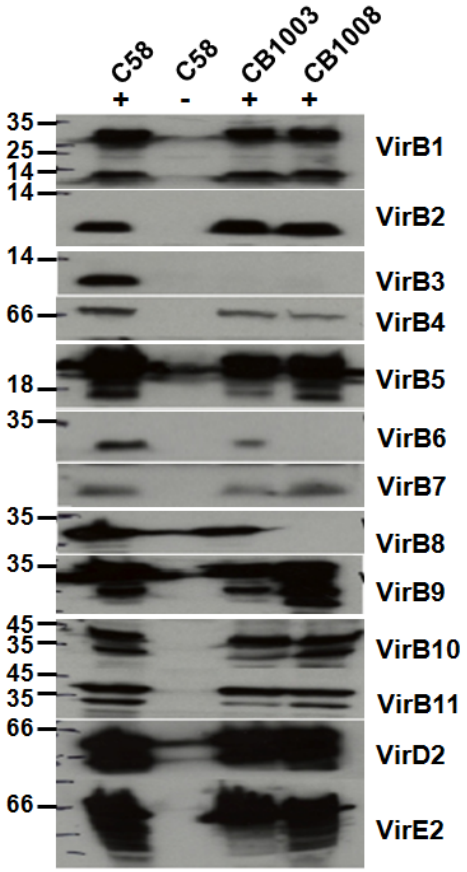

Fig. A2.1 Cellular levels of T4SS components in *virB3a* and *virB8a* deletion strains. C58: wild-type strain; CB1003: *virB3a* deletion mutant; CB1008: *virB8a* deletion mutant. "+": virulence induced; "-": virulence not induced. Molecular weights of reference proteins are shown on the left. Blots are representative of three independent experiments.

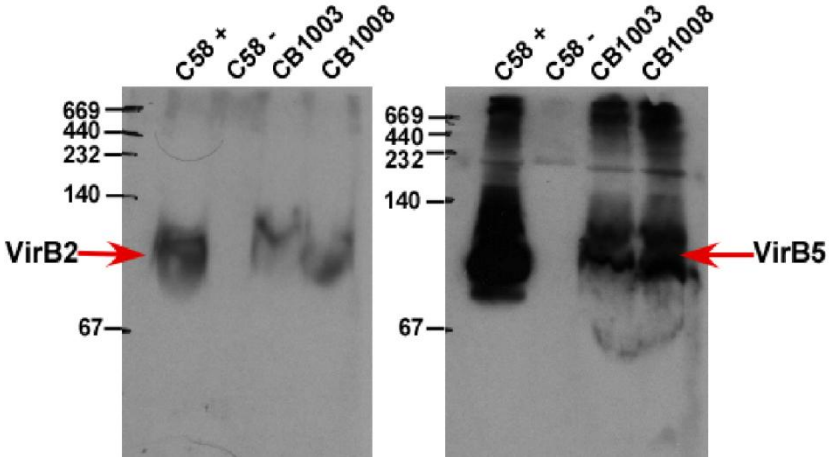

Fig. A2.2 Blue native-PAGE of the DDM-extracted membrane proteins from the *virB*3a and *virB8a* deletion strains. C58: wild-type strain; CB1003: *virB3a* deletion mutant; CB1008: *virB8a* deletion mutant. "+": virulence induced; "-": virulence not induced. Molecular weights of reference proteins are shown on the left. Blots are representative of three independent experiments.

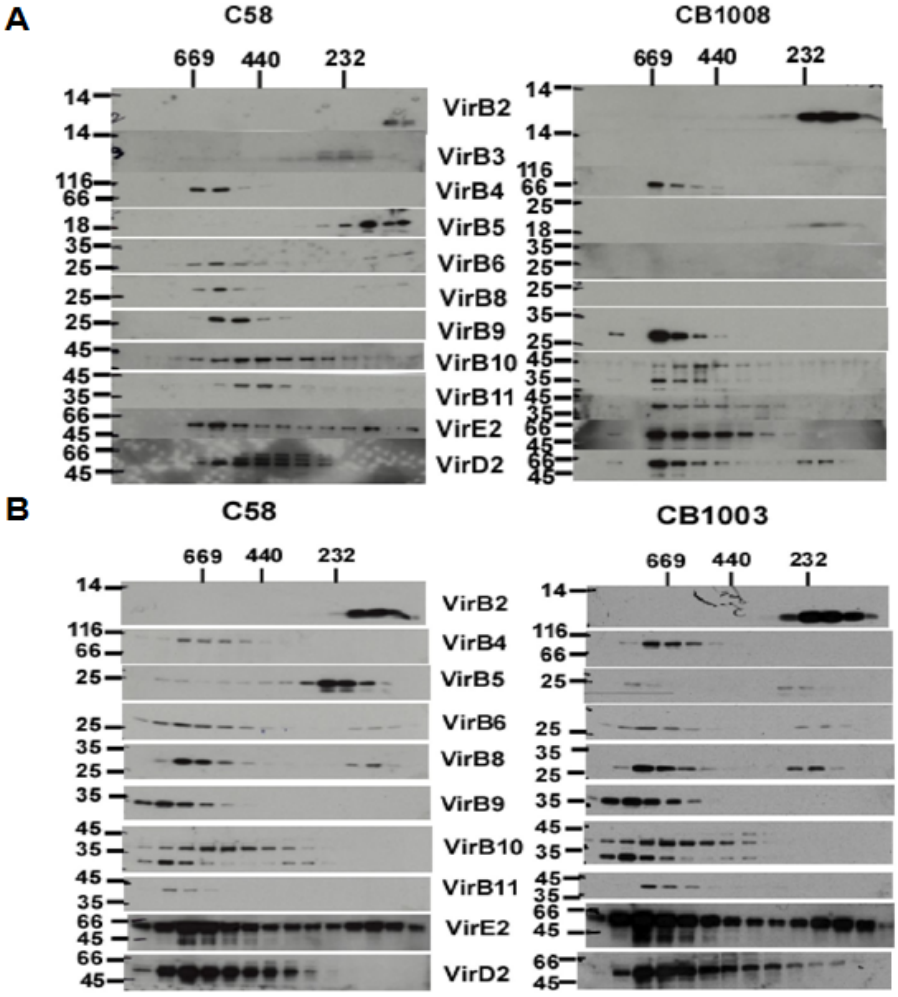

Fig. A2.3 Gel filtration of DDM-extracted membrane proteins from the *virB3a* and the *virB8a* deletion strains. C58: wild-type strain; CB1003: *virB3a* deletion mutant; CB1008: *virB8a* deletion mutant. Molecular weights of reference proteins are shown on the left (for western blots) and on the top (for gel filtration). Blots are representative of three independent experiments.

APPENDIX 3: IDENTIFICATION OF THE VIRB8/VIRB10-VIRB10 INTERACTION SITES

A3.1 INTRODUCTION

VirB10 is a bitopic inner membrane protein with a short N-terminal cytoplasmic domain (Das & Xie, 1998). It was initially reported to associate with the outer membrane protein VirB9 in yeast two-hybrid assays and this was later confirmed using co-immunoprecipitation, enzyme-linked immunosorbent assay (ELISA), bacterial two-hybrid assay (BTH) and surface plasmon resonance (SPR) (Beaupré *et al*, 1997; Das & Xie, 2000; Sivanesan *et al*, 2010). A recent study revealed the assembly of a core cylindrical structure from the homologs of VirB7, VirB9 and VirB10 encoded by the conjugative plasmid pKM101 (Fronzes *et al*, 2009b). The core complex is divided into an inner membrane layer (I) composed of the N-termini of the VirB9 and VirB10 homologs, and an outer membrane layer (O) consisting of the VirB7 homolog and the C-terminal domains of the VirB9 and VirB10 homologs (Fronzes *et al*, 2009b). Additionally, VirB10 homologs from several conjugation T4SSs also interact with the inner membrane anchored VirD4-like proteins (Gilmour *et al*, 2003; Llosa *et al*, 2003).

VirB10 shares similar domain arrangement with the *E. coli* energy coupling protein TonB, in that they both have a biotopic membrane topology and a poly-proline region adjacent to the transmembrane (TM) domain, conferring the periplasm domain of the protein with an elongated structure (Cascales & Christie, 2004a). TonB and VirB10 also both possess several protein-protein interaction motifs, including a coiled coil region near their TM domains, and highly conserved hydrophobic residues in their C-termini (Cascales & Christie, 2004a). In *E. coli*, TonB transduces the proton motive force by shuttling between the two membranes and changing conformation to provide energy for importing iron-siderophore complexes and vitamin B12 across the outer membrane (Larsen *et al*, 1999; Letain & Postle, 1997). Given that the *virB10a* deletion in *A. tumefaciens* arrests the transfer of T-DNA from the inner membrane core components VirB6a and VirB8a to the outer membrane associated VirB2a and VirB9a without direct contact with T-DNA, VirB10a is hypothesized to function as an energy sensor like TonB (Cascales & Christie, 2004a; Cascales & Christie, 2004b). Consistent with this notion, ATP hydrolysis by VirB11a and VirD4a induce conformational changes in VirB10a, as manifested by a change in the protease susceptibility of VirB10 (Cascales & Christie, 2004a). This structural transition is also necessary for the interaction between VirB10a and the outer membrane VirB7a-VirB9a heterodimer (Cascales & Christie, 2004a). These results not only unequivocally substantiated a role for VirB10 as an energy sensor, but also underlined the importance of the contacts between VirB10 and other T4SS components.

A direct contact between VirB8 and VirB10 was shown by both yeast two-hybrid assay and *in vitro* crosslinking experiments (Das & Xie, 2000; Paschos *et al*, 2006; Ward *et al*, 2002). Disruption of VirB8-VirB10 interaction *in vivo* by mutagenizing the amino acid

T201 of VirB8s (*B. suis* VirB8) resulted in a remarkable decrease in the survival rate of *B. suis* upon entry into macrophages and significantly impaired the ability of T4SS to enhance plasmid uptake by an *A. tumefaciens* recipient strains (Paschos *et al*, 2006; Sivanesan *et al*, 2010). However, the precise interaction interface in VirB10 had never been addressed. The crystal structure of the periplasmic C-terminus of *H. pylori* ComB10 (VirB10 homolog) revealed that this protein crystallizes as a dimer and that each monomer comprises an extensively modified β-barrel, with an α-helical antenna projecting from one side (Terradot *et al*, 2005) (Fig. A3.1A, B). A prominent feature of this structure is a groove formed between the β4 strand and the β6a/b strands. Residues within this region interact with the residues in the α2 and α3 helices of another monomer to form a substantial crystal-packing interface (Fig. A3.1C). In addition to the α2 and α3 helices, the α1 helix also protrudes from one side of the β-barrel. In the crystal structure of another VirB10 homolog, TraF, encoded by the Inc plasmid pKM101, the α1 extends laterally to make contacts with the β barrel of another monomer (Chandran *et al*, 2009). Therefore, this region was also regarded as a potential protein-protein interaction hotspot. However, these possible binding sites may not reflect the functional state of the protein, since the two subunits of the dimer are unusually arranged in opposite orientation in the crystal (Fig. A3.1C).

In order to elucidate the VirB8/VirB10-VirB10 binding interfaces, bioinformatic methods were implemented to predict surface-exposed interaction residues. VirB10s variants (*B. suis* VirB10), with single mutations at these sites, were created using site-directed mutagenesis. As a complementary approach, random mutagenesis was also conducted. The mutants were subsequently screened using BTH assays to identify VirB8/VirB10-VirB10 interaction sites. As *A. tumefaciens* T4SS components are less amenable than those of *B. suis* to *in vitro* work, all experiments in this section were conducted using *B. suis* homologs.

A3.2 MATERIALS AND METHODS

A3.2.1 Bacterial strains, plasmids, oligonucleotides and culture conditions

Strains and plasmids used in this study are listed in Table A1 and A2. Media, buffers and other materials are listed in Table A3. Oligonucleotides are listed in Table A4. Bacteria were cultivated as described in Section 2.2.1 in Chapter 2.

A3.2.2 Site directed mutagenesis and random mutagenesis

Manipulation of DNA including plasmid extraction, PCR amplification, digestion, ligation and sequencing were conducted following standard procedures (Maniatis *et al*, 1982). Restriction enzymes were purchased from New England Biolabs and MBI Fermentas and *E. coli* JM109 was used as a cloning host. Site-directed mutagenesis of the *virB10s* gene was conducted on the backbone of pUT18C*virB10s* as described before (Shenoy & Visweswariah, 2003). Primers carrying the desired mutations were used to

conduct inverse PCR in Mastercycler Gradient (Eppendorf). The PCR cycle was: 95°C 5 min; 60°C 30 s, 72°C 4 min (+10 s), 9 °C 30 s for 30 cycles; 72°C 7 min. The parental strands in 50 μl reaction were digested using 2 μl DpnI (10 U/μl). DNA was purified using a MinElute Gel Extraction Kit (QIAGEN), followed by electroporation (Eppendorf Electroporator 2510) into JM109 competent cells. The plasmids were extracted using QIAprep Spin Miniprep Kit (QIAGEN) and sent for sequencing.

Random mutagenesis was performed using a GeneMorph II EZClone kit (Stratagene). The DNA fragment encoding the periplasmic domain of VirB10s was amplified using error prone PCR in the presence of Mn^{2+} (Wilson & Keefe, 2001). In 50 μl reaction, 500 ng of template was mixed with 125 ng of each primer. The PCR cycle was: 95°C for 2 min; 95°C for 30 s, 55°C for 30 s, 72°C for 1 min 5 s for 30 cycles; 72°C for 10 min. All parameters were optimized to ensure a low mutation rate. The mutated PCR products were then used as megaprimers to amplify the whole plasmid of pUT18C*virB10s* via inverse PCR as described above, followed by digestion with *Dpn*I to remove wild-type parental DNA. The products were subsequently transformed into *E. coli* XL10-Gold competent cells provided in the kit using heat shock and were screened using BTH assays.

A3.2.3 SDS-PAGE and western blotting

Please refer to Section 2.2.4 in Chapter 2.

A3.2.4 Bacterial two-hybrid assay

BTH assay was performed as described in Section 2.2.9 in Chapter 2. Clones encoding the VirB10s mutants (created via site-directed mutagenesis) were transformed into BTH101 cells containing pKT25 plasmids expressing wild-type VirB8s and VirB10s respectively. Regarding the mutants generated using random mutagenesis, the ones showing drastic changes for either VirB8s-VirB10s or VirB10s-VirB10s interaction, with respect to the wild-type, were sent for sequencing. Cell extracts were prepared from the cultures of BTH101 (pKT25*virB8s*, pUT18C*virB10s* mutants) to check VirB10s and VirB8s protein levels using SDS-PAGE and western blotting.

A3.3 RESULTS AND DISCUSSION

A3.3.1 Disruption of VirB8s/VirB10s-VirB10s interactions by random mutagenesis

Interaction site characterization was initiated by conducting random mutagenesis of the *virB10s* gene. Mutations were generated in the fragment encoding the VirB10s periplasmic domain via error prone PCR. All parameters, including template concentration, primer concentration, and cycle number, were optimized to ensure a low mutation rate. Up to 220 clones were obtained and each was screened in the BTH system

using β-galactosidase assays (Karimova *et al*, 2000). Those showing specific reduction in VirB8s or VirB10s binding were sent for sequencing. It was found that mutant S315G (positioned within the β6b strand) was defective for both VirB8s-VirB10s and VirB10s-VirB10s interactions, although its effect on the previous was more drastic than on the latter (Fig. A3.2). Double mutant N258K (β3b strand) and Y299C (α1 helix) was specifically attenuated for VirB8s-VirB10s interaction whereas a double mutant carrying changes G253C (β3b strand) and K172R (region of undefined structure) only showed reduction in VirB10s-VirB10s interaction. The changes in the double mutants were further examined individually using site-directed mutagenesis (please see Section A3.3.3). The stability of these mutant proteins was also checked in parallel with other mutants created using site-directed mutagenesis.

A3.3.2 Multiple sequence alignment of VirB10s homologs and prediction of potential protein interaction sites

The ComB T4SS encoded by *H. pylori* directs DNA import from the extracellular environment (Hofreuter *et al*, 2001). The crystal structure of ComB10 provided us with valuable clues about potential interaction sites (Terradot *et al*, 2005). Structural similarities shared between this protein and VirB10s were supported by the results of *in silico* investigations using the Phyre online server, which searches the protein databases Protein Data Bank and Structural Classification of Proteins for homologous proteins with known three dimensional structures (Kelley & Sternberg, 2009). A model of the VirB10s structure was subsequently constructed using the ComB10 structure as a template (Kelley & Sternberg, 2009) (Fig. A3.1D). The overall shape of VirB10s is very similar to that of ComB10 except for two major differences: the β1c, β6b, and β7a strands in the back of the β-barrel and the β strands at the bottom were not conserved in VirB10s. The VirB10s sequence was also entered as a query to the Consurf online server, which searches for VirB10s homologous proteins using Position-Specific Iterative Basic Local Alignment Search Tool (PSI-BLAST) (Altschul *et al*, 1990; Glaser *et al*, 2003), to generate a multiple sequence alignment of these protein sequences using the default program MUSCLE (Edgar, 2004). This alignment and the structural model were used to predict functionally important regions of VirB10s using Consurf (Glaser *et al*, 2003). As shown in Fig. A3.3, the model of VirB10s can be visualized with the conservation scores color-coded onto its surface. In general, evolutionarily conserved residues exposed on the surface are usually functionally important, whereas highly conserved residues hidden within the protein core are more likely to be required for maintaining protein conformation (Glaser *et al*, 2003). Following these criteria, 38 surface exposed conserved residues were selected as target sites for our analysis. These residues are evenly distributed over the molecular structure model and can be grouped into 5 regions based on their localization: top antenna, groove area, α1 helix, bottom area and back of the β barrel. This, by in large, correlates with the interaction hotspots predicted for ComB10, since the amino acids in the top antenna and the groove are involved in crystal packing interface of ComB10, and α1 helix, as another projection on the protein surface, is also supposed to be a very interesting feature of the structure, as it contacts the β-barrel of two

other VirB10 monomers in the complex formed by a VirB10 homolog (Terradot *et al*, 2005, Chandran *et al*, 2009). The 38 residues were substituted with other amino acids and were evaluated for their effect on protein interactions using the BTH system as described below.

A3.3.3 Disruption of VirB8s/VirB10s-VirB10s interactions by site directed mutagenesis

The 38 selected residues were subjected to site-directed mutagenesis in the same fashion as described previously. The mutants were subsequently screened for VirB8s-VirB10s and VirB10s-VirB10s interactions in the BTH system using β-galactosidase assays. Samples were also taken from the VirB8s-VirB10s cultures to check protein stability of the mutant VirB10s proteins. The mutants were analyzed in four groups based on generation order (the data have to be presented this way because they need to be shown along with the controls of each batch, but all mutants are summarized according to their localization in the end). In the first group (Fig. A3.4A), I266D, T316A and S315G were compromised for both interactions, but this was likely due to decreased protein stability as determined by immunoblotting (Fig. A3.4A). Of the stable mutant proteins, T319A and N334K seemed to disrupt VirB8s-VirB10s interaction. In the second group (Fig. A3.4B), most mutant proteins appeared stable, and were expressed at similar levels. Substitution of Y236 with an Arg specifically reduced the VirB8s-VirB10s interaction, while Y236F affected both interactions. Changing the same residue into a Cys had a destabilizing effect. Mutation of T363 to a Val diminished both interactions and moderately reduced VirB10s protein level, while replacing this residue with a Lys only affected VirB8s-VirB10s interaction. In the third group (Fig. A3.4C), R307E, E318R and G253C/S mutant proteins were barely detectable by immunoblotting, however, of these three, only G253C/S exhibited a major decrease in interactions while the others had only a very subtle effect. Amongst the 11 mutants in the final group (Fig. A3.4D), only L243D displayed reduced interactions with both VirB8s and VirB10s; however, this could be explained by its lower cellular abundance.

A mutant is only considered to disrupt protein interactions if it is expressed at abundant levels and at the same time shows defects in protein interactions. To summarize the data according to the location of the different mutations (Table 2.1), changes in the top antenna residues and the α1 helix did not appear to affect protein stability and at the same time, were not important for interactions either, with the only exception being N334, where N334K mutant exhibited a minor reduction in VirB8s-VirB10s binding. For the amino acids in the back, T363 was particularly interesting, as different substitutions of this residue had differential effects: the change to Val led to decreased VirB10s protein level, while a Lys substitution had no effect on protein stability, indicating that the polar nature of this residue is essential for protein conformation. T319, L320 and N258 from the groove area specifically affected VirB8s-VirB10s interaction. At the bottom, Y236 was, like T363, differentially affected by different mutations. These results suggested that the crystal packing interface (residues in the top antenna interacting with the groove

residues in another monomer) as published before (Terradot *et al*, 2005), may not reflect the actual dimer interface. The top antenna may instead contribute to making contacts with the outer membrane associated VirB7-VirB9 complex, given the role for VirB10 as a bridging agent between the two membranes. It is interesting that some mutant proteins were scarcely visible on immunoblots, but displayed surprisingly high interactions levels. This may be explained by VirB10 mutant misfolding, which results in exposure of hydrophobic regions on the protein surface, leading to non-specific interactions and corresponding false positive signals. Therefore, the VirB10s mutants need further analysis using circular dichroism (CD) spectroscopy to ensure proper protein folding, and the interaction sites identified using the BTH study need further validation using *in vitro* binding assays such as pull down assays, blue native-PAGE, or co-immunoprecipitation, before embarking on *in vivo* investigations.

A3.4 FIGURES AND TABLES

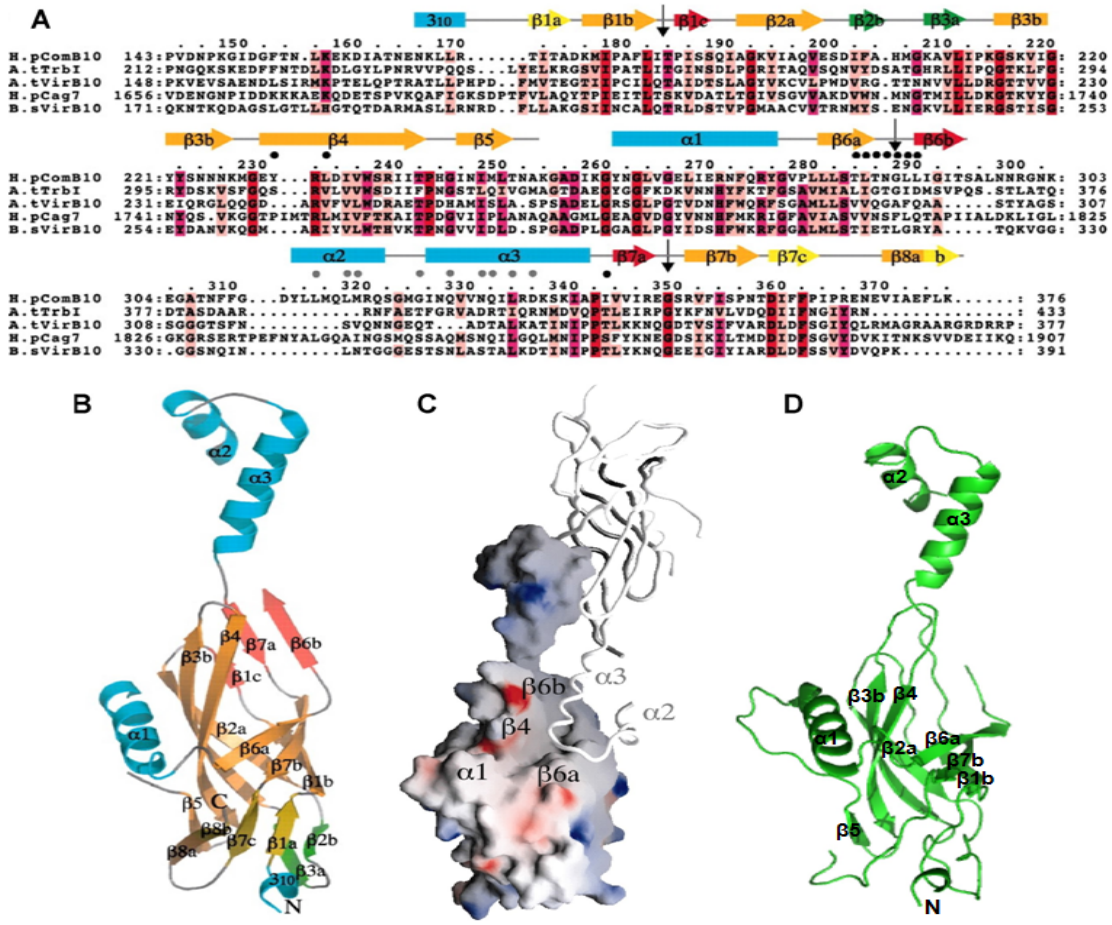

Fig. A3.1 Multiple sequence alignment and structures of VirB10 family proteins. (A) Multiple sequence alignment of the periplasmic domains of VirB10 proteins. Conserved residues are shaded in pink. The higher the conservation is, the darker the color is. The secondary structural elements are shown right above the alignment. The black and grey dots indicate residues in the crystal packing interface (The residues marked with black dots in one monomer interact the residues marked with grey dots in the other monomer). (B) X-ray structure of a monomer of *H. pylori* ComB10. (C) Dimer of ComB10 showing the crystal packing interface. One monomer is shown as surface representation of charge potential, while the other one is shown as a ribbon. The diagrams (A), (B) and (C) are adapted from Terradot *et al* (2005). (D) Structural model of the periplasmic domain of *B. suis* VirB10 generated using Phyre (Kelley & Sternberg, 2009).

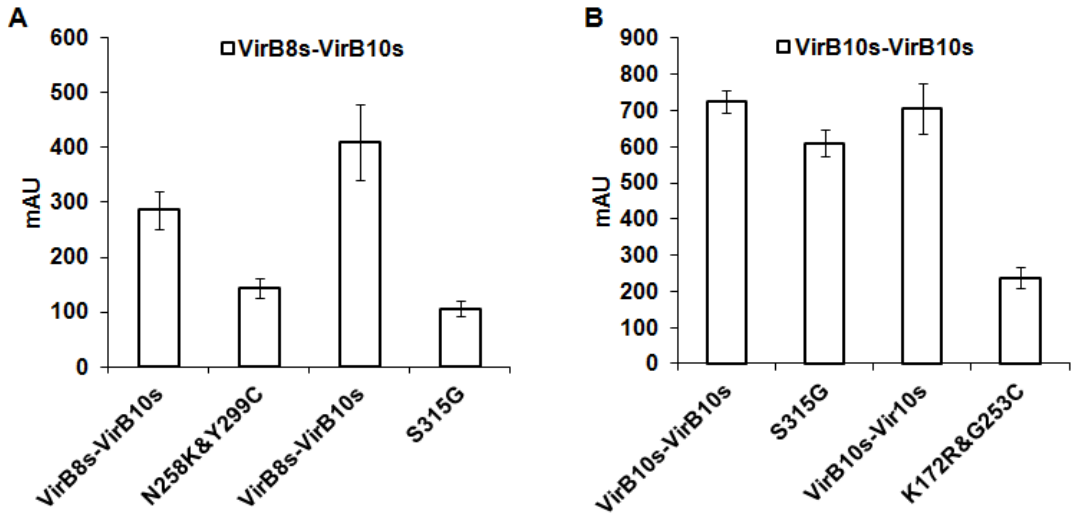

Fig. A3.2 BTH assays showing the interactions between wild-type VirB8s (A)/VirB10s (B) and VirB10s mutants generated via random mutagenesis. The VirB10s mutant constructs were generated by random mutagenesis, and were characterized in different batches; therefore, wild-type controls from each batch were shown in parallel to the mutants. Error bars indicate standard deviation from three independent experiments with duplicates for each.

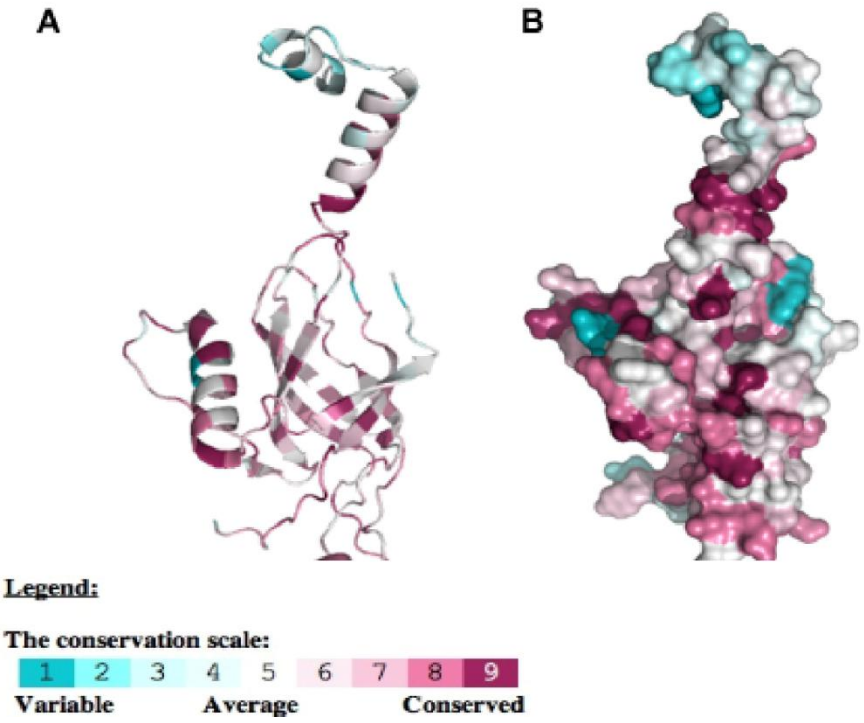

Fig. A3.3 Predicted functionally important sites of VirB10sp using the Consurf web server (Glaser *et al*, 2003). Ribbon (A) and surface (B) models of *B. suis* VirB10 with color-coded conservation scores. Models were generated using PyMol (The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC).

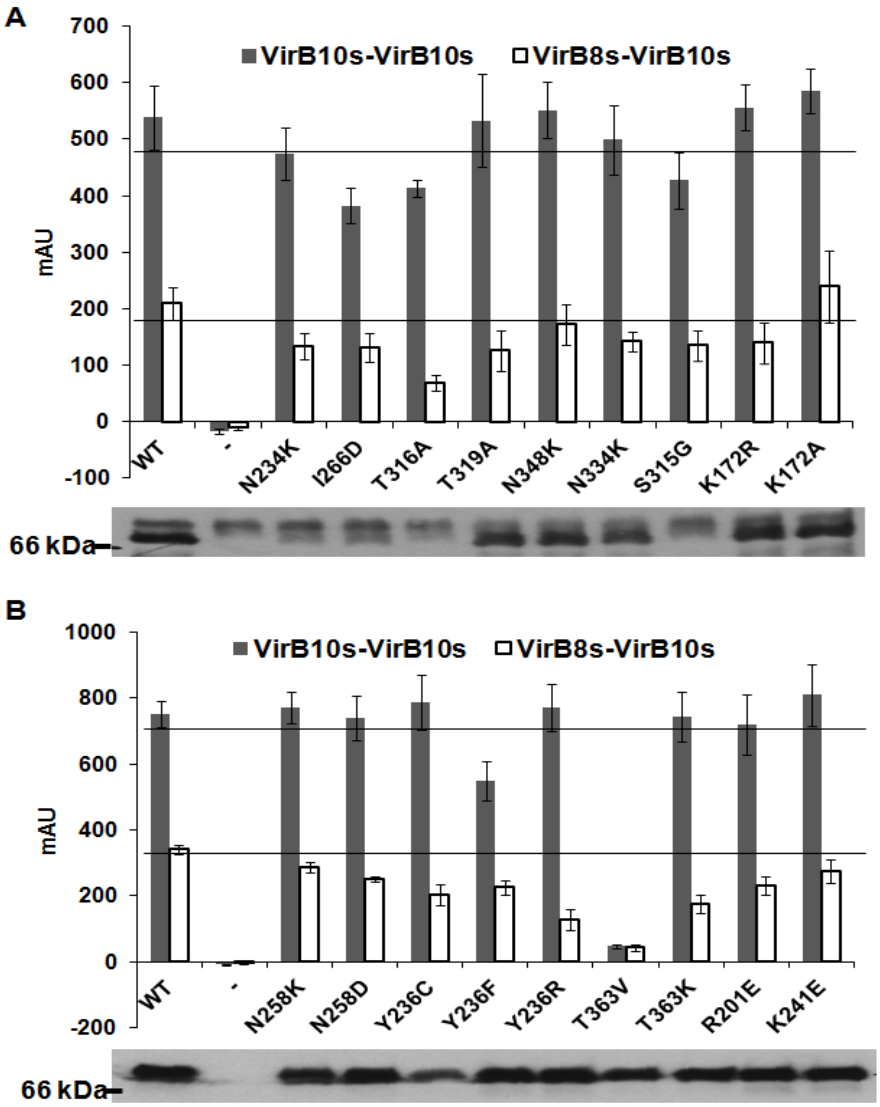

Fig. A3.4 BTH assays showing the interactions between wild-type VirB8s/VirB10s and VirB10s mutants obtained via site-directed mutagenesis. The mutants were examined based on their order of generation, so they were divided into four groups (A), (B), (C), and (D), and are presented here along with appropriate controls of each batch. The immunoblot underneath each panel shows the cellular levels of the VirB10s mutants. The loading order is the same as the order of the mutants in the bar graphs above. The black lines crossing the bars are aligned with the lower end of the error bars for the wild-type VirB10s interactions, so that any reduction appears more obvious. Error bars indicate the standard deviations from three independent experiments with duplicates for each.

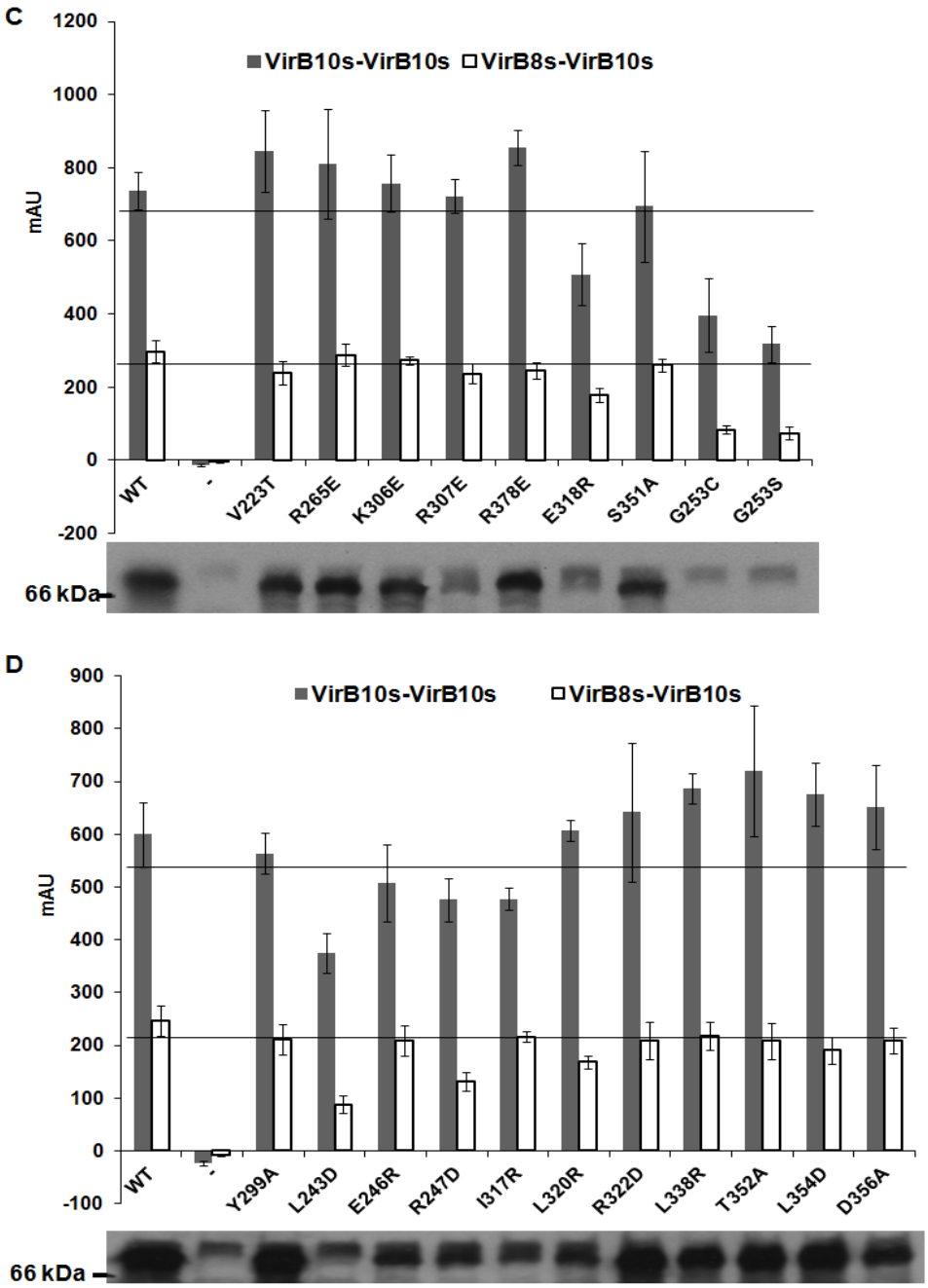

Fig. A3.4 BTH assays showing the interactions between wild-type VirB8s/VirB10s and VirB10s mutants obtained via site-directed mutagenesis. Please refer to the previous page for the legend.

Table A3.1 A summary of VirB10s mutants used in BTH assay. The mutants displaying abundant cellular protein levels and reduced VirB8/VirB10-VirB10 interaction levels are underlined.

Mutant	VirB10s level	VirB8s level	VirB8s-VirB10s	VirB10s-VirB10s
		Top antenna		
N334K	++	++	↓	/
L338R	+++	+++	/	/
N348K	+++	+++	/	/
S351A	+++	+++	/	/
T352A	+++	+++	/	/
L354D	+++	+++	/	/
D356A	+++	+++	/	/
		α1 helix		
Y299A	+++	+++	/	/
K306E	+++	+++	/	/
R307E	+++	+++	/	/
		Bottom loop area		
R201E	+++	+++	↓	/
N234K	+	+++	↓	/
Y236C	+	+++	↓	/
Y236F	+++	+++	↓	↓
Y236R	+++	+++	↓	/
K241E	+++	+++	↓	/
L243D	-	+	↓	↓
E246R	++	+++	/	/
R247D	+	+	↓	↓
		Groove area		
N258K	+++	+++	↓	/
N258D	+++	+++	↓	/
R265E	+++	+++	/	/
I266D	+	+++	↓	↓
S315G	+/-	+++	↓	↓
T316A	-	+	↓	↓
I317R	+	+	/	↓
E318R	-	+++	↓	↓
T319A	+++	+++	↓	/
L320R	+++	+++	↓	/
R322D	+++	+++	/	/
R378E	+++	+++	/	/
		Back		
V223T	+++	+++	/	/
G253C	-	+++	↓	↓
G253S	-	+++	↓	↓
T363K	+++	+++	↓	/
T363V	++	++	↓	↓

[Protein level] "+++": wild-type level; "++": slightly lower; "+": very low; "+/-": hardly detectable; "-": not detectable

[Interaction level] "/": no change; "↓": reduced

Table A1 Bacterial strains used in the study

Strains	Genotype or description	References
Escherichia coli		
JM109	F' (*tra*D36 *pro*AB+ *lac*IqZΔM15) *rec*A1 *end*A1 *gyr*A96 *thi*-1 *hsd*R17 *rel*A1 *sup*E44 Δ(*lac-pro*AB)	Yanisch-Perron *et al*., 1985
GJ1158	*omt*T *hsd*S *gal dcm* Δ*mal*Ap510 *mal*P::(*pro*Up-T7 RNAP) *mal*Q::*lacZhyb*11 Δ(*zhf*-900::Tn10dTet)	Bhandari *&* Gowrishankar, 1997
BTH101	F^- *cya-99*, *araD139, galE15, galK16, rpsL1 (Str^r), hsdR2, mcrA1, mcrB1*	Karimova *et al*., 2005
TOP10	F- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *deo*R *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G	Invitrogen
XL1—Gold	Tet^r Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1 gvrA96 relA1 lac* Hte [F' *proAB lacI^{q}ZΔM15* Tn*10* (Tet^r) Amy Cam^r]	Stratagene
Agrobacterium. tumefaciens		
C58	Wild-type strain carrying the nopaline producing Ti plasmid, pTiC58.	van Larekeke *et al*. 1974
CB1008	C58 carrying an in frame deletion of *virB8* on pTiC58	Aly *et al*. 2008
CB1004	C58 carrying an in frame deletion of *virB4* on pTiC58	Yuan *et al*. 2005
CB1003	C58 carrying an in frame deletion of *virB3* on pTiC58	Aly, K.A.
A348	Wild-type strain carrying the octopine producing Ti plasmid, pTiA6NC.	Van Larebeke *et al*., 1974
PC1004	A348 carrying an in frame deletion of *virB4* on pTiA6NC	Berger & Christie, 1994
PC1001	A348 carrying an in frame deletion of *virB1* on pTiA6NC	Berger & Christie, 1994
UIA143 pTiA6	A348 derivative, *recA*, ery^R	Bohne *et al*., 1998

Table A2 Plasmids used in this study

Plasmid	Description	References
pQE30*cagE*	Amp^R, a 1 kb *BamH*I/*Pst*I fragment encoding *A. actinomycetemcomitans* CagE under the control of T5 promoter	Teng *et al.,* 2003
pTrc200	Str^R, spc^R, pVS1 origin, $lacI^q$, *trc* promotor expression vector	Schmidt-Eisenlohr *et al*., 1999
pTrc*cagE*	Str^R, spc^R, a 1 kb *Nco*I/*Sal*I fragment encoding *A. actinomycetemcomitans* CagE cloned downstream of Trc promoter	This study
pVSBAD*virB4*	Str^R, spc^R, pVSBAD carrying a 2.4 kb *Acc65*I/*Pst*I *virB4* fragment from *A. tumefaciens* C58	Yuan *et al*., 2005
pTrc*virB1*	Str^R, spc^R, a 738 bp *Nco*I/*HinD*III fragment encoding *A. tumefaciens* VirB1 cloned downstream of Trc promoter	Höppner *et al*., 2004
pTrc*virB1s*	Str^R, spc^R, a 717 bp *Nco*I/*HinD*III fragment encoding *B. suis* VirB1 cloned downstream of Trc promoter	Höppner *et al*., 2004
pTrc*virB3*	Str^R, spc^R, an *EcoR*I/*BamH*I fragment encoding pTiC58 VirB3 cloned downstream of Trc promoter	This study
pKT25	Kan^R, derivative of pSU40 allowing fusion at the C-terminus of the T25 fragment of *B. pertussis* adenylate cyclase in BTH assay	Karimova *et al*. 2005
pKNT25	Kan^R, derivative of pSU40 allowing fusion at the N-terminus of the T25 fragment of *B. pertussis* adenylate cyclase in BTH assay	Karimova *et al*. 2005
pUT18C	$Carb^R$, derivative of pUC19 allowing fusion at the C-terminus of theT18 fragment of *B. pertussis* adenylate cyclase in BTH assay	Karimova *et al*. 2005
pUT18	$Carb^R$, derivative of pUC19 allowing fusion at the N-terminus of the T18 fragment of *B. pertussis* adenylate cyclase in BTH assay	Karimova *et al*. 2005
pLS1	$Carb^R$, IncQ plasmid, used for VirB/D4 mediated conjugative transfer experiment	Stahl *et al.* 1998
pUT18C*virB10s*	$Carb^R$, a 1.2 kb *Xba*I/*Kpn*I *virB10s* fragment cloned downstream of T18 fragment	Sivanesan, D.
pKT25*virB10s*	Kan^R, a 1.2 kb *Xba*I/*Kpn*I *virB10s* fragment cloned upstream of T25 fragment	Sivanesan, D.

Table A3 Growth media and buffers used in this study

Medium	Composition
LB	1% tryptone; 0.5% yeast extract; 0.5% NaCl
LBON	1% tryptone; 0.5% yeast extract
YEB	1% tryptone; 0.5% sucrose; 0.1%; 2mM $MgSO_4$
ABMM	1% glucose; 0.39% MES; 1 mM Na-K-phosphate pH 5.5; 1 x AB salts (20 x AB salts: 20 g NH_4Cl; 6 g $MgSO_4$ x $7H_2O$; 3 g KCl; 0.2 g $CaCl_2$; 50 mg $FeSO_4$ x $7H_2O$ per liter; pH 5.5)
Buffer	**Composition**
Na-K-P buffer (0.5 M)	Make 500 ml 0.5 M KH_2PO_4; adjust to pH 5.5 by adding 0.5 M Na_2HPO_4
Blue native gel buffer (3 ×)	1.5 M 6-aminocaproic acid; 150 mM BisTris; adjust to pH 7.0; stored at 4°C
Blue native-PAGE anode buffer	100 mM BisTris; adjust to pH 7.0; stored at 4°C
Blue native-PAGE cathode buffer	100 mM Tricine; 30 mM BisTris; adjust to pH 7.0; stored at 4°C
5% Coomassie Blue (10 ml)	0.5 g Coomassie blue G 250; 0.656 g 6-Aminocaproic acid
Membrane destaining buffer	25% methanol; 10% acetate; 65% H_2O
4 × Lower Tris	1.5 M Tris/Cl pH 8.8; 0.4% SDS
4 × Upper Tris	0.5 M Tris/Cl pH 6.8; 0.4% SDS
10 × Laemmli running buffer	0.25 M Tris; 1.92 M Glycine; 1% SDS
2 × Laemmli sample buffer	0.125 M Tris/Cl (pH 6.8); 4% SDS; 20% glycerol; 10% β-mercaptoethanol; a few grains of bromphenol blue
Blotting buffer	192 mM glycine; 25 mM Tris; 20% methanol
Schägger gel buffer 3 ×	3 M Tris/Cl; 0.3% SDS; adjust to pH 8.45
Schägger gel anode buffer 5 ×	1 M Tris/Cl; adjust to pH 8.9
Schägger gel cathode buffer 5 ×	0.5 M Tris/Cl; 0.5 M Tricine; 5% SDS; adjust to pH 8.25
ONPG buffer (9 ml)	8.1 ml Z-buffer; 0.9 ml 8 mg/ml ONPG in Z-buffer; 11.25 µl 10% SDS; 26.64 µl β-mercaptoethanol (99.77%)
Z-buffer (500 ml)	8.05 g $Na_2HPO_4 \bullet 7H_2O$; 2.75 g $NaH_2PO_4 \bullet H_2O$; 0.375 g KCl; 0.123 g $MgSO_4 \bullet 7H_2O$; adjust to pH 7.0
Stop buffer	1 M Na_2CO_3
ACA buffer	750 mM 6-aminocaproic acid; 50 mM BisTris; adjust to pH 7.0

Table A4 Oligonucleotides used in the study

Primer name	Sequence
Inverse PCR primers	
VirB10spN258K5’	5’-GATGCC**AAG**GTAAAGCAGGGCATGGCT-3’
VirB10spN258K3’	5’-CTTTAC**CTT**GGCATCATATTCACCCGA-3’
VirB10spN258D5’	5’-GATGCC**GAC**GTAAAGCAGGGCATGGCT-3’
VirB10spN258D3’	5’-CTTTAC**GTC**GGCATCATATTCACCCGA-3’
VirB10spY236C5’	5’-AACATG**TGT**AGCGATAACGGCAAGGTG-3’
VirB10spY236C3’	5’-ATCGCT**ACA**CATGTTGCGTGTGACCAC-3’
VirB10spY236F5’	5’-AACATG**TTT**AGCGATAACGGCAAGGTG-3’
VirB10spY236F3’	5’-ATCGCT**AAA**CATGTTGCGTGTGACCAC-3’
VirB10spY236R5’	5’-AACATG**CGT**AGCGATAACGGCAAGGTG-3’
VirB10spY236R3’	5’-ATCGCT**ACG**CATGTTGCGTGTGACCAC-3’
VirB10spT363V5’	5’-CCGCCG**GTA**CTGTACAAGAACCAGGGC-3’
VirB10spT363K5’	5’-CCGCCG**AAA**CTGTACAAGAACCAGGGC-3’
VirB10spT363K3’	5’-GTACAG**TTT**CGGCGGAATGTTGATCGT-3’
VirB10spR201E5’	5’-CGCAAC**GAA**GATTTCCTGCTCGCGAAG-3’
VirB10spR201E3’	5’-GAAATC**TTC**GTTGCGCAGAAGGCTCGC-3’
VirB10spK241E5’	5’-AACGGC**GAG**GTGTTGCTGATTGAGCGC-3’
VirB10spK241E3’	5’-CAACAC**CTC**GCCGTTATCGCTATACAT-3’
VirB10spV223T5’	5’-TCGACG**ACG**CCGGGCATGGCTGCCTGC-3’
VirB10spV223T3’	5’-GCCCGG**CGT**CGTCGAATCCAGACGGGT-3’
VirB10spR265E5’	5’-ATGGCT**GAA**ATTTATGTCCTGTGGACG-3’
VirB10spR265E3’	5’-ATAAAT**TTC**AGCCATGCCCTGCTTTAC-3’
VirB10spK306E5’	5’-TTCTGG**GAG**CGCTTTGGCGGCGCCTTG-3’
VirB10spK306E3’	5’-AAAGCG**CTC**CCAGAAGTGGGAGTCGAT-3’
VirB10spR307E5’	5’-TGGAAG**GAA**TTTGGCGGCGCCTTGATG-3’
VirB10spR307E3’	5’-GCCAAA**TTC**CTTCCAGAAGTGGGAGTC-3’
VirB10spR378E5’	5’-ATCGCC**GAA**GACCTAGATTTTTCGAGT-3’
VirB10spR378E3’	5’-TAGGTC**TTC**GGCGATATAGATGCCGAT-3’
VirB10spY299A5’	5’-CCCGGC**GCC**ATCGACTCCCACTTCTGG-3’
VirB10spY299A3’	5’-GTCGAT**GGC**GCCGGGCAAGCCTGCCCC-3’
VirB10spN234K5’	5’-ACACGC**GCC**ATGTATAGCGATAACGGC-3’
VirB10spN234K3’	5’-ATACAT**GGC**GCGTGTGACCACGCAGGC-3’
VirB10spL234D5’	5’-AAGGTG**GAT**CTGATTGAGCGCGGTTCA-3’
VirB10spL234D3’	5’-AATCAG**ATC**CAACACCTTGCCGTTATC-3’
VirB10spE246R5’	5’-CTGATT**AGG**CGCGGTTCAACCATCTCG-3’
VirB10spE246R3’	5’-ACCGCG**CCT**AATCAGCAACACCTTGCC-3’
VirB10spR247D5’	5’-ATTGAG**GAC**GGTTCAACCATCTCGGGT-3’
VirB10spR247D3’	5’-TGAACC**GTC**CTCAATCAGCAACACCTT-3’
VirB10spI266D5’	5’-GCTCGC**GAT**TATGTCCTGTGGACGCGC-3’
VirB10spI266D3’	5’-GACATA**ATC**GCGAGCCATGCCCTGCTT-3’

VirB10spT316A5’	5’-TTGAGC**GCG**ATCGAGACCCTCGGCCGC-3’
VirB10spT316A3’	5’-CTCGAT**CGC**GCTCAACATCAAGGCGCC-3’
VirB10spI317R5’	5’-AGCACG**AGA**GAGACCCTCGGCCGCTAT-3’
VirB10spI317R3’	5’-GGTCTC**TCT**CGTGCTCAACATCAAGGC-3’
VirB10spE318R5’	5’-ACGATC**AGG**ACCCTCGGCCGCTATGCA-3’
VirB10spE318R3’	5’-GAGGGT**CCT**GATCGTGCTCAACATCAA-3’
VirB10spT319A5’	5’-ATCGAG**GCC**CTCGGCCGCTATGCAACC-3’
VirB10spT319A3’	5’-GCCGAG**GGC**CTCGATCGTGCTCAACAT-3’
VirB10spL320R5’	5’-GAGACC**CGC**GGCCGCTATGCAACCCAG-3’
VirB10spL320R3’	5’-GCGGCC**GCG**GGTCTCGATCGTGCTCAA-3’
VirB10spR322D5’	5’-CTCGGC**GAC**TATGCAACCCAGAAGGTC-3’
VirB10spR322D3’	5’-TGCATA**GTC**GCCGAGGGTCTCGATCGT-3’
VirB10spL338R5’	5’-ATCAAC**CGC**AATACCGGCGGAGGTGAA-3’
VirB10spL338R3’	5’-GGTATT**GCG**GTTGATCTGATTTGAACC-3’
VirB10spN348K5’	5’-ACGAGC**AAG**CTGGCTTCAACTGCCTTG-3’
VirB10spN348K3’	5’-AGCCAG**CTT**GCTCGTCGATTCACCTCC-3’
VirB10spS351R5’	5’-CTGGCT**GCA**ACTGCCTTGAAGGATACG-3’
VirB10spS351R3’	5’-GGCAGT**TGC**AGCCAGGTTGCTCGTCGA-3’
VirB10spT352A5’	5’-GCTTCA**GCT**GCCTTGAAGGATACGATC-3’
VirB10spT352A3’	5’-CAAGGC**AGC**TGAAGCCAGGTTGCTCGT-3’
VirB10spL354D5’	5’-ACTGCC**GAT**AAGGATACGATCAACATT-3’
VirB10spL354D3’	5’-ATCCTT**ATC**GGCAGTTGAAGCCAGGTT-3’
VirB10spD356A5’	5’-TTGAAG**GCT**ACGATCAACATTCCGCCG-3’
VirB10spD356A3’	5’-GATCGT**AGC**CTTCAAGGCAGTTGAAGC-3’
VirB10spN334K5’	5’-GGTTCA**AAG**CAGATCAACCTCAATACC-3’
VirB10spN334K3’	5’-GATCTG**CTT**TGAACCCCCGCCGCCGAC-3’
VirB10spS315G5’	5’-ATGTTG**GGC**ACGATCGAGACCCTCGGC-3’
VirB10spS315G3’	5’-GATCGT**GCC**CAACATCAAGGCGCCGCC-3’
VirB10spK172R5’	5’-AGTCAA**AGG**AACACCAAGCAGGATGCT-3’
VirB10spK172R3’	5’-GGTGTT**CCT**TTGACTGTCGAGCAGGGC-3’
VirB10spK172A5’	5’-AGTCAA**GCG**AACACCAAGCAGGATGCT-3’
VirB10spK172A3’	5’-GGTGTT**CGC**TTGACTGTCGAGCAGGGC-3’
VirB10spG253C5’	5’-ATCTCG**TGT**GAATATGATGCCAACGTA-3’
VirB10spG253C3’	5’-ATATTC**ACA**CGAGATGGTTGAACCGCG-3’
VirB10spG253S5’	5’-ATCTCG**AGT**GAATATGATGCCAACGTA-3’
VirB10spG253S3’	5’-ATATTC**ACT**CGAGATGGTTGAACCGCG-3’
Random mutagenesis	
VirB10sp5’	5’-ATTCAAGCCCTGCTCGACAGT-3’
VirB10sp3’	5’-GATGAATTCGAGCTCGGTACC-3’
Plasmid construction for *in vivo* complementation	
VirB35’	5’-CGCGGAATTCATGAATGATCGTCTGG-3’
VirB33’	5’-CGCGGGATCCTTACGCCATTCCTCTC-3’
CagETrc5’	5’-CGCTGCCATGGTCCCTGAAATTTTATTAGC-3’

CagETrc3’	5’-CGCTGGTCGACTTAAACGACCTTTAAACATTTTTTTA-3’

*Mutagenized codons are bolded. Restriction sites are underlined.

CHAPTER 3:
SUMMARY AND FUTURE DIRECTIONS

Large scale genome sequencing has unveiled a plethora of novel protein sequences; however, protein function cannot always be inferred from sequence. As a consequence, understanding protein interactions and the dynamic behavior of these proteins in living organisms have become a major challenge in the post genomic era. As countless cellular activities depend on the interplay between proteins, the functions of these proteins can often be revealed through the investigation of their interacting partners (Guan & Kiss-Toth, 2008). Proteins participating in the same cellular event frequently comprise distinct components in a signaling pathway or assemble into a multi-component complex. Secretion systems encoded by microorganisms are one such example of membrane bound complexes, and are dedicated to the translocation of a wide range of macromolecules. Type IV secretion systems (T4SSs) serve as an outstanding model for studying dynamic protein interactions and their overall contribution to bacterial fitness.

VirB4, as the most widely conserved component of T4SSs, and as a putative energy supplier, plays a central role in T4SS, and understanding its activity will go a long way towards unravelling the mechanism underlying T4SS assembly and function. In Chapter 2, I set out to address the interplay between VirB4 and VirB8 using *in vitro* biochemical approaches, including affinity chromatography, blue native-PAGE and size exclusion chromatography. This was coupled with a genetic strategy to further probe the observed interactions: using a structure guided sequence alignment and amino acid conservation analysis, VirB8 residues predicted to be important for VirB4-VirB8 interaction were mutagenized. The resulting proteins were then characterized using blue native-PAGE, where R230 of VirB8s was identified as an essential site. However, *in vivo* functional assays revealed that the equivalent VirB8a mutants, when used to complement an *A. tumefaciens virB8* deletion mutant, had only a very subtle effect. As discussed in Chapter 2, this may be due to weak, residual interactions between VirB4 and mutant VirB8 proteins in their natural biological context. In moving forward, adjacent amino acids in the same solvent accessible region of VirB8 could be mutagenized to completely block VirB4-VirB8 binding. At the same time, since interaction site elucidation has been partly hindered by the lack of knowledge regarding VirB4 membrane topology, in the future, the membrane association by VirB4 should be re-evaluated using improved reporter assays – for example, by constructing GFP fusions (as cytoplasmic reporter) and alkaline phosphatase fusions (as periplasmic reporter) at the same time (Rapp *et al*, 2004). Fluorescence protease protection assay, which has been widely employed to investigate protein membrane topology in eukaryotic systems, may also be adapted to Gram negative bacteria (Lorenz *et al*, 2006).

Although no ATP hydrolyzing activity was detected for purified Strep-tagged VirB4s using a malachite green assay, this activity is worth further investigation, since low yield and instability of StrepII-VirB4s had prevented us from exploring more conditions,

including varying the pH, the type of nucleotides (GTP and ATP), and salt conditions. Altering the type of salt would be a priority, given that recent work on the VirB4 family protein TrwK, encoded by plasmid R388, found that acetate salts, compared to chloride salts, greatly stimulated the ATPase activity of a TrwK (Arechaga *et al*, 2008). It is also possible that VirB4s ATPase activity may be stimulated in the presence of another molecule, and so it could also be tested in the presence of its interacting partners like VirB8 or T4SS translocation substrates like single stranded DNA, VirD2, and VirE2 at different ratios. Additionally, phospholipid liposomes or detergent-phospholipid micelles could be added to reactions to simulate membrane association, as VirB4 is a putative membrane protein and its activity may require membrane insertion (Krause *et al*, 2000b; Lambert *et al*, 1998). *In silico* work predicted the presence of three α-helices in the C-termini of Vir4 family proteins (Pena *et al*, 2011). Interestingly, this C-terminal domain has been shown to have autoinhibitory effect on the ATPase activity of TrwK (Pena *et al*, 2011). Therefore, C-terminally truncated variants of Vir4s should be subjected to analysis, along with the full length protein.

The insolubility and instability of recombinant VirB4s has been a major obstacle in its biochemical/biophysical characterization, as most analyses require a stable, concentrated protein sample (*e.g.* 0.5-1 M for structural studies). Optimizing induction conditions and switching overexpression hosts only had marginal effects. Stimulating chaperone production by adding chemicals shortly prior to induction of StrepII-VirB4s did not significantly alleviate the problem either, but at the same time introduced sample contamination by GroEL. Although the chaperone contaminants could be removed by adding additional purification steps, these will further reduce protein yield. Consequently, in the future, we may consider extracting StrepII-VirB4s from inclusion bodies, where a large quantity of the recombinant protein exists in an almost homogeneous form (Fahnert *et al*, 2004). Traditional inclusion body solubilization involves extraction with a high concentration (6-8 M) of protein denaturant, such as urea or guanidinium chloride. This is typically followed by gradual removal of denaturants via dialysis. If the protein contains disulfide bonds, the dialysis buffer is usually supplemented with a redox system such as a mixture of reduced and oxidized glutathione, to provide an appropriate environment to enable formation and reshuffling of disulfide bonds (Clark, 1998; Lilie *et al*, 1998; Vallejo & Rinas, 2004). A major drawback to this solubilization approach is that high concentration of denaturants leads to the loss of native protein secondary structures, which are usually preserved in inclusion bodies, and may cause random coil formation and concomitant exposure of hydrophobic surfaces, increasing the propensity for aggregation (Dill & Shortle, 1991; Oberg *et al*, 1994; Przybycien *et al*, 1994). This may be avoided by lowering the urea concentration to 2 M and adjusting the buffer pH to a value distant from the pI of StrepII-VirB4s (predicted to be ~8.25) (Singh *et al*, 2005; Singh *et al*, 2012). Refolding of the solubilized protein may be performed in a pulse manner to maximize recovery since aggregation only occurs between partially folded intermediates, while fully folded proteins do not interact with newly added unfolded proteins (Fischer *et al*, 1992). As VirB4 family proteins contain multiple cysteine residues (9 in *A. tumefaciens* VirB4 and 10 in *B. suis* Vir4), indicating that disulfide bonds may be

important for their tertiary structures upon membrane insertion (disulfide bonds in prokaryotes are usually formed in the periplasm), a redox system should also be added to the refolding buffer. Finally, the sample may be purified using ion exchange and size exclusion chromatography (Singh *et al*, 2012). Depending on the subsequent assays to be performed, the sample may also be subjected to concentration using ultrafiltration columns. In order to aid proper refolding and to prevent precipitation, charged amino acids (L-Arg and L-Glu) could be added at 1:1 ratio to a concentration of 50 mM to all buffers (Golovanov *et al*, 2004). These two amino acids have been demonstrated to dramatically reduce protein aggregation and prevent proteolytic degradation in many instances (Guttman *et al*, 2010; Moriscot *et al*, 2011; Thakur *et al*, 2007). Although the exact mechanism is unknown, it has been suggested that L-Agr and L-Glu associate with the surface of proteins (without altering the protein properties), reducing surface "stickiness" between protein molecules and masking protease recognition sites (Golovanov *et al*, 2004).

In addition to its predicted function as an NTPase, VirB4 also shows homology to a T4SS-associated virulence factor encoded by *A. actinomycetemcomitans*, CagE (Teng & Hu, 2003). The sequence similarity shared by these two proteins lies mainly in the C-terminus of each protein, a region that encompasses the three α-helices proposed to inhibit the nucleotide hydrolyzing activity of another VirB4 family protein (Pena *et al*, 2011). In Appendix 1, the relationship between VirB4 and *A. actinomycetemcomitans* CagE was examined via heterologous complementation of an *A. tumefaciens virB4* deletion mutant with a CagE- expressing construct. Functional studies revealed that CagE could not substitute for VirB4 during complex assembly and substrate translocation. However, a clear immunoblot showing expression of CagE in the *virB4* mutant strain is needed before any definitive conclusions can be drawn here.

In Appendix 2, *A. tumefaciens virB3* and *virB8* deletion mutants were examined for their ability to assemble T4SS. Membrane-bound complexes were solubilized and analyzed using blue native-PAGE and gel filtration chromatography to test the subcomplex formation between the remaining VirB proteins. Both strains showed a co-fractionation profile similar to that of the wild-type strain, indicating that these two proteins were not required for the assembly of the T4SS apparatus. However, given that both *virB3* and *virB8* are indispensable for T-pilus biogenesis (Lai *et al*, 2000), their corresponding protein products may play an important role during the transfer of the T-pilin VirB2 from the inner membrane to the T-pilus. VirB4 is also known to mediate the dislocation of VirB2 from the inner membrane (Kerr & Christie, 2010), and thus it may be worth testing whether this requires the participation of VirB8, through VirB4-VirB8 interactions.

While characterizing deletion mutants can provide useful information about protein function, the precise role of any T4SS component still relies largely on its interaction with other components. By examining binary interactions between different VirB proteins, we were able to separate the complex 12-component system into simpler interaction

components. Besides VirB4, VirB8 also interacts with VirB10. In Appendix 3, VirB10s variants carrying mutations at potential interaction sites were created and screened using the BTH system. Residues required for both VirB8-VirB10 interaction and VirB10 self-oligomerization were identified. Future work in this area should be focused on validating these mutants using *in vitro* protein binding assays such as co-fractionation or co-elution in blue native-PAGE, pull down experiments, and gel filtration chromatography. Purified VirB10 variants should also be analyzed using CD spectroscopy, to ensure the overall protein shape is maintained (Kelly & Price, 2000). By following a similar experimental strategy to that implemented for examining the VirB4-VrB8 interaction, a *virB10* deletion mutant could be transformed with *virB10* mutants-encoding constructs, and the biological consequences of these mutations could be evaluated using *A. tumefaciens* functional assays such as T-pilus assembly, conjugation and tumor induction, and macrophage infection by *B. suis* (Paschos *et al*, 2006).

A major concern associated with our mutant study is that although the conformation of *B. suis* VirB mutants can be monitored using CD spectroscopy, their *A. tumefaciens* counterparts could not be readily assessed using the same approach since *A. tumefaciens* VirB proteins are less amenable to biochemical work. The decreased protein interaction level observed for a mutant in the BTH system may be caused by conformational changes in the protein, and not a specific disruption of protein interaction, which could explain the very subtle effect of the VirB8a mutations *in vivo*. In order to verify the specificity of protein interaction residues, compensatory changes in their binding partner may be created (Campbell *et al*, 2011; Richardson *et al*, 1999). For instance, protein A could be subjected to random mutagenesis, followed by screening using the BTH assay to identify mutants capable of restoring its interactions with the mutants of its binding partner, B, (which are known to disrupt interaction). The compensatory A mutant, along with its corresponding B mutant could be subsequently co-expressed in native host *A. tumefaciens*. If the mutation can restore T4SS activity, it will provide some support for the specificity of the interaction sites identified; however, this is all contingent on whether we could observe a drastic phenotype caused by the B mutant itself.

Single amino acid changes, in some instances, are not sufficient to fully block protein interactions *in vivo*, and as such, substitutions of multiple residues on the same interface have been proposed. However, this may increase the possibility of protein destabilization. An alternative approach could involve the use of peptide aptamers. Peptide aptamers are oligopeptides displayed on a protein scaffold, *e.g.* thioredoxin A (TrxA) (usually from the active site of TrxA so that TrxA enzymatic activity is destroyed) in a conformationally constrained form (Hoppe-Seyler *et al*, 2004). Such aptamers have been successfully utilized to probe minimal protein interaction domains and to specifically block protein interactions in both eukaryotic and prokaryotic systems (Loregian & Palu, 2005; Paradis-Bleau *et al*, 2005; Woodman *et al*, 2005). Compared to linear peptides, a constrained form renders peptide aptamers more resistant to proteolysis and more potent in terms of target binding (Hoppe-Seyler *et al*, 2004). Based on the structures or structure models of VirB proteins, oligopeptides may be designed to mimic the surface exposed regions

predicted to be important for interactions, *e.g.* the β4 strand of VirB8s where R230 (a VirB4-VirB8 interaction site identified in blue native-PAGE) is located. When fused to a scaffold protein and overexpressed, the ability of these peptide aptamers to inhibit VirB protein interactions could then be assessed using both biochemical experiments and BTH assays. Effective aptamers should be expressed with N-terminal signal peptides (for export into the periplasm, as many of the T4SS protein interactions take place in the periplasm) in wild-type *A. tumefaciens* or *B. suis* strains to determine their impact on T4SS function.

The research presented in this portion of my thesis has provided us with important information on T4SS component interactions and has provided clear direction for future work. Our long term goal is to not only fully understand the precise role for VirB4 in the process of T4SS assembly and substrate translocation, but also to identify novel targets for anti-virulence/pathogenicity treatment. Similar to the mechanism by which peptide aptamers act, many drugs exert therapeutic effects by competing for binding sites on a target protein with its endogenous interaction partners (Baines & Colas, 2006; Walensky *et al*, 2004). As T4SSs are broadly utilized by pathogens to transmit antibiotic resistance genes and to translocate virulence factors into host cells, established VirB interactions that are essential for T4SS function could be targeted using peptide aptamers that specifically block these interactions. The aptamers could then serve as a guide for small molecule drug discovery. Compared to traditional treatments that usually kill the bacteria, new drugs designed in this way are advantageous in that they do not perturb the development of other species in the human microbiome and, more importantly, they effectively disarm pathogens without killing them, and thus exert less selective pressure on these pathogens, potentially delaying or decreasing the occurrence of mutations – and accompanying resistance - in pathogenic virulence-related genes (Alksne, 2002; Baron & Coombes, 2007; Rasko & Sperandio, 2010).

REFERENCES:

Abdallah AM, Pittius NCGv, Champion PAD, Cox J, Luirink J, Vandenbroucke-Grauls CMJE, Appelmelk BJ, Bitter W (2007) Type VII secretion - *Mycobacteria* show the way. *Nat Rev Microbiol* **5:** 883-891

Alksne LE (2002) Virulence as a target for antimicrobial chemotherapy. *Expert Opin Investig Drugs* **11:** 1149-1159

Allardet-Servent A, Michaux-Charachon S, Jumas-Bilak E, Karayan L, Ramuz M (1993) Presence of one linear and one circular chromosome in the *Agrobacterium tumefaciens* C58 genome. *J Bacteriol* **175:** 7869-7874

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215:** 403-410

Aly KA, Baron C (2007) The VirB5 protein localizes to the T-pilus tips in *Agrobacterium tumefaciens*. *Microbiology* **153:** 3766-3775

Arechaga I, Pena A, Zunzunegui S, del Carmen Fernandez-Alonso M, Rivas G, de la Cruz F (2008) ATPase activity and oligomeric state of TrwK, the VirB4 homologue of the plasmid R388 type IV secretion system. *J Bacteriol* **190:** 5472-5479

Arnold R, Jehl A, Rattei T (2010) Targeting effectors: the molecular recognition of Type III secreted proteins. *Microbes Infect* **12:** 346-358

Atherton JC, Blaser MJ (2009) Coadaptation of *Helicobacter pylori* and humans: ancient history, modern implications. *J Clin Invest* **119:** 2475-2487

Atumakuri K, Cascales E, Christie PJ (2004) Energetic components VirD4, VirB11 and VirB4 mediate early DNA transfer reactions required for bacterial type IV secretion. *Mol Microbiol* **54**

Backert S, Fronzes R, Waksman G (2008) VirB2 and VirB5 proteins: specialized adhesins in bacterial type-IV secretion systems? *Trends Microbiol* **16:** 409-413

Backert S, Meyer TF (2006) Type IV secretion systems and their effectors in bacteiral pathogenesis. *Curr Opin Microbiol* **9:** 207-217

Backert S, Tegtmeyer N, Selbach M (2010) The versatility of *Helicobacter pylori* CagA effector protein functions: The master key hypothesis. *Helicobacter* **15:** 163-176

Bailey S, Ward D, Middleton R, Grossmann JG, Zambryski PC (2006) *Agrobacterium tumefaciens* VirB8 structure reveals potential protein-protein interactions sites. *Proc Natl Acad Sci USA* **103:** 2582-2587

Baines IC, Colas P (2006) Peptide aptamers as guides for small-molecule drug discovery. *Drug Discov Today* **11:** 334-341

Baron C (2005) From bioremediation to biowarfare: on the impact and mechanism of type IV secretion systems. *FEMS Microbiol Lett* **253:** 163-170

Baron C, Coombes B (2007) Targeting bacterial secretion systems: benefits of disarmament in the microcosm. *Infect Disord Drug Targets* **7:** 19-27

Baron C, Llosa M, Zhou S, Zambryski PC (1997) VirB1, a component of the T-complex transfer machinery of *Agrobacterium tumefaciens*, is processed to a C-terminal secreted product, VirB1. *J Bacteriol* **179:** 1203-1210

Baron C, O'Callaghan D, Lanka. E (2002) Bacterial secrets of secretion: EuroConference on the biology of type IV secretion processes. *Mol Microbiol* **43:** 1359-1365

Baron C, Thorstenson YR, Zambryski PC (1997b) The lipoprotein VirB7 interacts with VirB9 in the membranes of *Agrobacterium tumefaciens*. *J Bacteriol* **179:** 1211-1218

Basler M, Pilhofer M, Henderson GP, Jensen GJ, Mekalanos JJ (2012) Type VI secretion requires a dynamic contractile phage tail-like structure. *Nature* [Epub ahead of print]

Beaupré CE, Bohne J, Dale EM, Binns AN (1997) Interactions between VirB9 and VirB10 membrane proteins involved in movement of DNA from *Agrobacterium tumefaciens* into plant cells. *J Bacteriol* **179:** 78-89

Bechtluft P, van Leeuwen RG, Tyreman M, Tomkiewicz D, Nouwen N, Tepper HL, Driessen AJ, Tans SJ (2007) Direct observation of chaperone-induced changes in a protein folding pathway. *Science* **318:** 1458-1461

Beijersbergen A, Dulk-Ras AD, Schilperoort RA, Hooykaas PJ (1992) Conjugative Transfer by the Virulence System of *Agrobacterium tumefaciens*. *Science* **256:** 1324-1327

Beijersbergen A, Smith SJ, Hooykaas PJ (1994) Localization and topology of VirB proteins of *Agrobacterium tumefaciens. Plasmid* **32:** 212-218

Berger BR, Christie PJ (1994) Genetic complementation analysis of the *Agrobacterium tumefaciens virB* operon: virB2 through *virB11* are essential virulence genes. *J Bacteriol* **176:** 3646-3660

Berrow NS, Bussow K, Coutard B, Diprose J, Ekberg M, Folkers GE, Levy N, Lieu V, Owens RJ, Peleg Y, Pinaglia C, Quevillon-Cheruel S, Salim L, Scheich C, Vincentelli R, Busso D (2006) Recombinant protein expression and solubility screening in *Escherichia coli*: a comparative study. *Acta Crystallogr D Biol Crystallogr* **62:** 1218-1226

Bhandari P, Gowrishankar J (1997) An *Escherichia coli* host strain useful for efficient overproduction of cloned gene products with NaCl as the inducer. *J Bacteriol* **179:** 4403-4406

Bohne J, Yim A, Binns AN (1998) The Ti plasmid increases the efficiency of *Agrobacterium tumefaciens* as a recipient in *virB*-mediated conjugal transfer of an IncQ plasmid. *Proc Natl Acad Sci USA* **95:** 7057-7062

Bondos SE, Bicknell A (2003) Detection and prevention of protein aggregation before, during, and after purification. *Anal Biochem* **316:** 223-231

Bordo D, Argos P (1991) Suggestions for "safe" residue substitutions in site-directed mutagenesis. *J Mol Biol* **217:** 721-729

Borloo J, De Smet L, Vergauwen B, Van Beeumen JJ, Devreese B (2007) A beta-galactosidase-based bacterial two-hybrid system to assess protein-protein interactions in the correct cellular environment. *J Proteome Res* **6:** 2587-2595

Boschiroli ML, Ouahrani-Bettache S, Foulongne V, Michaux-Charachon S, Bourg G, Allardet-Servent A, Cazevielle C, Lavigne J, Liautard JP, Ramuz M, O'Callaghan D (2002a) Type IV secretion and *Brucella* virulence. *Veterinary Microbiology* **90:** 341-348

Boschiroli ML, Quahrani-Bettache S, Foulongne V, Michaux-Charachon S, Bourg G, Allardet-Servent A, Cazevieille C, Liautard JP, Ramuz M, O'Callaghan D (2002b) The *Brucella suis virB* operon is induced intracellularly in macrophages. *Proc Natl Acad Sci USA* **99:** 1544-1549

Bruckner A, Polge C, Lentze N, Auerbach D, Schlattner U (2009) Yeast two-hybrid, a powerful tool for systems biology. *Int J Mol Sci* **10:** 2763-2788

Brundage L, Hendrick JP, Schiebel E, Driessen AJ, Wickner W (1990) The purified *E. coli* integral membrane protein SecY/E is sufficient for reconstitution of SecA-dependent precursor protein translocation. *Cell* **62:** 649-657

Campbell RN, Westhorpe F, Reece RJ (2011) Isolation of compensatory inhibitor domain mutants to novel activation domain variants using the split-ubiquitin screen. *Yeast* **28:** 569-578

Cascales E (2008) The type VI secretion toolkit. *EMBO Reports* **9:** 735-741

Cascales E, Christie PJ (2003) The versatile bacterial type IV secretion systems. *Nat Rev Microbiol* **1:** 137-149

Cascales E, Christie PJ (2004a) *Agrobacterium* VirB10, an ATP energy sensor required for type IV secretion. *Proc Natl Acad Sci USA* **101:** 17228-17233

Cascales E, Christie PJ (2004b) Definition of a bacterial type IV secretion pathway for a DNA substrate. *Science* **304:** 1170-1173

Chandran V, Fronzes R, Duquerroy S, Cronin N, Navaza J, Waksman G (2009) Structure of the outer membrane complex of a type IV secretion system. *Nature* **462:** 1011-1015

Chen LS, Chen YC, Wood DW, Nester EW (2002) A new type IV secretion system promotes conjugal transfer in *Agrobacterium tumefaciens*. *J Bacteriol* **184:** 4838-4845

Chilton MD, Currier TC, Farrand SK, Bendich AJ, Gordon MP, Nester EW (1974) *Agrobacterium tumefaciens* DNA and PS8 bacteriophage DNA not detected in crown gall tumors. *Proc Natl Acad Sci U S A* **71:** 3672-3676

Christie PJ (2004) Type IV secretion: the *Agrobacterium* VirB/D4 and related conjugation systems. *Biochim Biophys Acta* **1694:** 219-234

Christie PJ, Atmakuri K, Krishnamoorthy V, Jakubowski S, Cascales E (2005) Biogenesis, architecture, and function of bacterial type IV secretion systems. *Annu Rev Microbiol* **59:** 451-485

Christie PJ, Vogel JP (2000) Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells. *Trends Microbiol* **8:** 354-360

Clark EDB (1998) Refolding of recombinant proteins. *Curr Opin Biotechnol* **9:** 157-163

Coers J, Monahan C, Roy CR (1999) Modulation of phagosome biogenesis by *Legionella pneumophila* creates an organelle permissive for intracellular growth. *Nat Cell Biol* **1:** 451-453

Cogan EB, Birrell GB, Griffith OH (1999) A robotic-based automated assay for inorganic and organic phosphates. *Anal Biochem* **271:** 29-35

Cooksey DA, Moore LW (1980) Biological control of crown gall with fungal and bacterial antagonists. *Phytopathology* **70:** 506-509

Cornelis GR, Wolf-Watz H (1997) The *Yersinia* Yop virulon: a bacterial system for subverting eukaryotic cells. *Mol Microbiol* **23:** 861-867

Costerton JW, Ingram JM, Cheng KJ (1974) Structure and function of the cell envelope of Gram-negative bacteria. *Bacteriol Rev* **38:** 87-110

Cruciat CM, Brunner S, Baumann F, Neupert W, Stuart RA (2000) The cytochrome bc1 and cytochrome c oxidase complexes associate to form a single supracomplex in yeast mitochondria. *J Biol Chem* **275:** 18093-18098

d'Enfert C, Chapon C, Pugsley AP (1987) Export and secretion of the lipoprotein pullulanase by *Klebsiella pneumoniae*. *Mol Microbiol* **1:** 107-116

Dang TA, Christie PJ (1997) The VirB4 ATPase of *Agrobacterium tumefaciens* is a cytoplasmic membrane protein exposed at the periplasmic surface. *J Bacteriol* **179:** 453-462

Dang TA, Zhou XR, Graf B, Christie PJ (1999) Dimerization of the *Agrobacterium tumefaciens* VirB4 ATPase and the effect of ATP-binding cassette mutations on the assembly and function of the T-DNA transporter. *Mol Microbiol* **32:** 1239-1253

Das A, Xie Y (1998) Construction of transposon Tn3phoA: its application in defining the membrane topology of the *Agrobacterium tumefaciens* DNA transfer proteins. *Mol Microbiol* **27:** 405-414

Das A, Xie YH (2000) The *Agrobacterium* T-DNA transport pore proteins VirB8, VirB9 and VirB10 interact with one another. *J Bacteriol* **182:** 758-763

de Jong MF, Sun YH, den Hartigh AB, van Dijl JM, Tsolis RM (2008) Identification of VceA and VceC, two members of the VjbR regulon that are translocated into macrophages by the *Brucella* type IV secretion system. *Mol. Microbiol* **70:**1378-1396

Dill KA, Shortle D (1991) Denatured states of proteins. *Annu Rev Biochem* **60:** 795-825

Draper O, Middleton R, Doucleff M, Zambryski PC (2006) Topology of the VirB4 C terminus in the *Agrobacterium tumefaciens* VirB/D4 type IV secretion system. *J Biol Chem* **281:** 37628-37635

du Plessis DJ, Nouwen N, Driessen AJ (2011) The Sec translocase. *Biochim Biophys Acta* **1808:** 851-865

Durand E, Oomen C, Waksman G (2010) Biochemical dissection of the ATPase TraB, the VirB4 homologue of the *Escherichia coli* pKM101 conjugation machinery. *J Bacteriol* **192:** 2315-2323

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high through put. *Nucleic Acids Res* **32:** 1792-1797

Eisenbrandt R, Lai EM, Lurz R, Kado CI, Lanka E (1999) Conjugative pili of IncP plasmids and the Ti plasmid T pilus are composed of cyclic subunits. *J Biol Chem* **274:** 22548-22555

Fahnert B, Lilie H, Neubauer P (2004) Inclusion bodies: formation and utilisation. *Adv Biochem Eng Biotechnol* **89:** 93-142

Filloux A (2004) The underlying mechanisms of type II protein secretion. *Biochim Biophys Acta* **1694:** 163-179

Fischer B, Perry B, Sumner I, Goodenough P (1992) A novel sequential procedure to enhance the renaturation of recombinant protein from *Escherichia coli* inclusion bodies. *Protein Eng* **5:** 593-596

Fortune SM, Jaeger A, Sarracino DA, Chase MR, Sassetti CM, Sherman DR, Bloom BR, Rubin EJ (2005) Mutually dependent secretion of proteins required for mycobacterial virulence. *Proc Natl Acad Sci U S A* **102:** 10676-10681

Franz B, Kempf VA (2011) Adhesion and host cell modulation: critical pathogenicity determinants of *Bartonella henselae*. *Parasit Vectors* **4:** 54

Fronzes R, Christie PJ, Waksman G (2009a) The structural biology of type IV secretion systems. *Nat Rev Microbiol* **7:** 703-714

Fronzes R, Schafer E, Wang L, Saibil HR, Orlova EV, Waksman G (2009b) Structure of a type IV secretion system core complex. *Science* **323:** 266-268

Fry DC, Kuby SA, Mildvan AS (1986) ATP-binding site of adenylate kinase: Mechanistic implications of its homology with ras-encoded p21, Fl-ATPase, and other nucleotide-binding proteins. *Proc Natl Acad Sci USA* **83:** 907-911

Fullner KJ, Stephens KM, Nester EW (1994) An essential virulence protein of *Agrobacterium tumefaciens*, VirB4, requires an intact mononucleotide binding domain to function in transfer of T-DNA. *Mol Gen Genet* **245:** 704-715

Gentschev I, Dietrich G, Goebel W (2002) The *E. coli* α-hemolysin secretion system and its use in vaccine development. *Trends Microbiol* **10:** 39-45

Gilmour MW, Gunton JE, Lawley TD, Taylor DE (2003) Interaction between the IncHI1 plasmid R27 coupling protein and type IV secretion system: TraG associates with the coiled-coil mating pair formation protein TrhB. *Mol Microbiol* **49:** 105-116

Glaser F, Pupko T, Paz I, Bell RE, Bechor-Shental D, Martz E, Ben-Tal N (2003) ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information *Bioinformatics* **19:** 163-164

Gold VA, Duong F, Collinson I (2007) Structure and function of the bacterial Sec translocon. *Mol Membr Biol* **24:** 387-394

Golovanov AP, Hautbergue GM, Wilson SA, Lian LY (2004) A simple method for improving protein solubility and long-term stability. *J Am Chem Soc* **126:** 8933-8939

Grahn AM, Haase J, Bamford DH, Lanka E (2000) Components of the RP4 conjugative transfer apparatus form an envelope structure bridging inner and outer membranes of donor cells: implications for related macromolecule transport systems. *J Bacteriol* **182:** 1564-1574

Grohmann E, Muth G, Espinosa M (2003) Conjugative plasmid transfer in Gram-positive bacteria. *Microbiol Mol Biol Rev* **67:** 277-301

Guan H, Kiss-Toth E (2008) Advanced technologies for studies on protein interactomes. *Adv Biochem Eng Biotechnol* **110:** 1-24

Guttman M, Prieto JH, Handel TM, Domaille PJ, Komives EA (2010) Structure of the minimal interface between ApoE and LRP. *J Mol Biol* **398:** 306-319

Hanson PI, Whiteheart SW (2005) AAA+ proteins: have engine, will work. *Nat Rev Mol Cell Biol* **6:** 519-529

Hapfelmeier S, Domke N, Zambryski PC, Baron C (2000) VirB6 is required for stabilization of VirB5, VirB3 and formation of VirB7 homodimers in *Agrobacterium tumefaciens*. *J Bacteriol* **182:** 4505-4511

Hare S, Bayliss R, Baron C, Waksman G (2006) A large domain swap in the VirB11 ATPase of *Brucella suis* leaves the hexameric assembly intact. *J Mol Biol* **360:** 56-66

Harlow E, Lane D (1988) Antibodies: A Laboratory Manual, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Hartl FU, Lecker S, Schiebel E, Hendrick JP, Wickner W (1990) The binding cascade of SecB to SecA to SecY/E mediates preprotein targeting to the *E. coli* plasma membrane. *Cell* **63:** 269-279

Haystead CM, Gregory P, Sturgill TW, Haystead TA (1993) γ-phosphate-linked ATP-sepharose for the affinity purification of protein kinases. Rapid purification to homogeneity of skeletal muscle mitogen-activated protein kinase kinase. *Eur J Biochem* **214:** 459-467

Hegde RS, Bernstein HD (2006) The surprising complexity of signal sequences. *Trends Biochem Sci* **31:** 563-571

Henderson IR, Nataro JP (2001) Virulence functions of autotransporter proteins. *Infect Immun* **69:** 1231-1243

Henderson IR, Navarro-Garcia F, Desvaux M, Fernandez RC, Ala'Aldeen D (2004) Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev* **68:** 692-744

Henderson IR, Navarro-Garcia F, Nataro JP (1998) The great escape: structure and function of the autotransporter proteins. *Trends Microbiol* **6:** 370-378

Hewlett EL (ed) (2007) *Bordetella* toxins. In *Bordetella:* Molecular Microbiology.: Horizon Bioscience, 97-118pp

Hofreuter D, Odenbreit S, Haas R (2001) Natural transforamtion competence in *Helicobacter pylori* is mediated by the basic components of a type IV secretion system. *Mol Microbiol* **41:** 379-391

Holland IB, Schmitt L, Young J (2005) Type 1 protein secretion in bacteria, the ABC-transporter dependent pathway (review). *Mol Membr Biol* **22:** 29-39

Hoppe-Seyler F, Crnkovic-Mertens I, Tomai E, Butz K (2004) Peptide aptamers: specific inhibitors of protein function. *Curr Mol Med* **4:** 529-538

Hoppner C, Liu Z, Domke N, Binns AN, Baron C (2004) VirB1 orthologs from *Brucella suis* and pKM101 complement defects of the lytic transglycosylase required for efficient type IV secretion from *Agrobacterium tumefaciens*. *J Bacteriol* **186:** 1415-1422

Hormaeche I, Alkorta I, Moro F, Valpuesta JM, Goni FM, De La Cruz F (2002) Purification and properties of TrwB, a hexameric, ATP-binding integral membrane protein essential for R388 plasmid conjugation. *J Biol Chem* **277:** 46456-46462

Hoskins J, Lovell S, Blundell TL (2006) An algorithm for predicting protein-protein interaction sites: Abnormally exposed amino acid residues and secondary structure elements. *Protein Sci* **15:** 1017-1029

Hubber A, Roy CR (2010) Modulation of host cell function by *Legionella pneumophila* type IV effectors. *Annu Rev Cell Dev Biol* **26:** 261-283

Jacob-Dubuisson F, Locht C, Antoine R (2001) Two-partner secretion in Gram-negative bacteria: a thrifty, specific pathway for large virulence proteins. *Mol Microbiol* **40:** 306-313

Jakubowski SJ, Krishnamoorthy V, Cascales E, Christie PJ (2004) *Agrobacterium tumefaciens* VirB6 domains direct the ordered export of a DNA substrate through a type IV secretion System. *J Mol Biol* **341:** 961-977

Jakubowski SJ, Krishnamoorthy V, Christie PJ (2003) *Agrobacterium tumefaciens* VirB6 protein participates in formation of VirB7 and VirB9 complexes required for type IV secretion. *J Bacteriol* **185:** 2867-2878

Johnson TL, Abendroth J, Hol WG, Sandkvist M (2006) Type II secretion: from structure to function. *FEMS Microbiol Lett* **255:** 175-186

Jones AL, Shirasu K, Kado CI (1994) The product of the *virB4* gene of Agrobacterium *tumefaciens* promotes accumulation of VirB3 protein. *J Bacteriol* **176:** 5255-5261

Jones S, Thornton JM (1996) Principles of protein-protein interactions. *Proc Natl Acad Sci U S A* **93:** 13-20

Judd PK, Kumar RB, Das A (2005) Spatial location and requirements for the assembly of the *Agrobacterium tumefaciens* type IV secretion apparatus. *Proc Natl Acad Sci USA* **102:** 11498-11503

Kalkum M, Eisenbrandt R, Lurz R, Lanka E (2002) Tying rings for sex. *Trends Microbiol* **10:** 382-387

Karimova G, Pidoux J, Anges U, Ladant D (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc Natl Acad Sci USA* **95:** 5752-5756

Karimova G, Ullmann A, Ladant D (2000) A bacterial two-hybrid system that exploits a cAMP signaling cascade in *Escherichia coli*. *Methods in enzymology* **328:** 59-73

Kelley LA, Sternberg MJE (2009) Protein structure prediction on the web: a case study using the Phyre server. *Nature Protocol* **4:** 363-371

Kelly SM, Price NC (2000) The use of circular dichroism in the investigation of protein structure and function. *Curr Protein Pept Sci* **1:** 349-384

Keret D, Giladi M, Kletter Y, Wientroub S (1998) Cat-scratch disease osteomyelitis from a dog scratch. *J Bone Joint Surg Br* **80:** 766-767

Kerr JE, Christie PJ (2010) Evidence for VirB4-mediated dislocation of membrane-integrated VirB2 pilin during biogenesis of the *Agrobacterium* VirB/VirD4 type IV secretion system. *J Bacteriol* **192:** 4923-4934

Krall L, Wiedemann U, Unisin G, Weiss S, Domke N, Baron C (2002a) Detergent extraction identifies different VirB protein subassemblies of the type IV secretion machinery in the membranes of *Agrobacterium tumefaciens*. *Proc Natl Acad Sci USA* **99:** 11405-11410

Krause S, Barcena M, Pansegrau W, Lurz R, Carazo JM, Lanka E (2000a) Sequence-related protein export NTPases encoded by the conjugative transfer region of RP4 and by the cag pathogenicity island of *Helicobacter pylori* share similar hexameric ring structures. *Proc Natl Acad Sci U S A* **97:** 3067-3072

Krause S, Pansegrau W, Lurz R, Cruz Fdl, Lanka E (2000b) Enzymology of type IV macromolecule secretion systems: The conjugative transfer regions of plasmids RP4 and R388 and the cag pathogenicity island of *Helicobacter pylori* encode structurally and functionally related nucleoside triphosphate hydrolases. *J Bacteriol* **182:** 2751-2770

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

Kuldau GA, Devos G, Owen J, McCaffrey G, Zambryski PC (1990) The *virB* operon of *Agrobacterium tumefaciens* pTiC58 encodes 11 open reading frames. *Mol Gen Genet* **221:** 256-266

Kumar RB, Das A (2001) Functional Analysis of the *Agrobacterium tumefaciens* T-DNA Transport Pore Protein VirB8. *J Bacteriol* **183:** 3636-3641

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227:** 680-685

Lai EM, Chesnokova O, Banta LM, Kado CI (2000) Genetic and environmental factors affecting T-pilin export and T-pilus biogenesis in relation to flagellation of *Agrobacterium tumefaciens*. *J Bacteriol* **182:** 3705-3716

Lai EM, Kado CI (1998) Processed VirB2 is the major subunit of the promiscuous pilus of *Agrobacterium tumefaciens*. *J Bacteriol* **180:** 2711-2717

Lai EM, Kado CI (2002) The T-pilus of *Agrobacterium tumefaciens*. *Trends Microbiol* **8:** 361-369

Lambert O, Levy D, Ranck JL, Leblanc G, Rigaud JL (1998) A new "gel-like" phase in dodecyl maltoside-lipid mixtures: implications in solubilization and reconstitution studies. *Biophys J* **74:** 918-930

Larsen RA, Thomas MG, Postle K (1999) Protonmotive force, ExbB and ligand-bound FepA drive conformational changes in TonB. *Mol Microbiol* **31:** 1809-1824

Leiman PG, Basler M, Ramagopal UA, Bonanno JB, Sauder JM, Pukatzki S, Burley SK, Almo SC, Mekalanos JJ (2009) Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. *Proc. Natl Acad. Sci. USA* **106**: 4154–4159

Letain TE, Postle K (1997) TonB protein appears to transduce energy by shuttling between the cytoplasmic membrane and the outer membrane in *Escherichia coli*. *Mol Microbiol* **24:** 271-283

Lewis KN, Liao R, Guinn KM, Hickey MJ, Smith S, Behr MA, Sherman DR (2003) Deletion of RD1 from M*ycobacterium tuberculosis* mimics bacille Calmette-Guerin attenuation. *J Infect Dis* **187:** 117-123

Li PL, Everhart DM, Farrand SK (1998) Genetic and sequence analysis of the pTiC58 *trb* locus, encoding a mating-pair formation system related to members of the type IV secretion family. *J Bacteriol* **180:** 6164-6172

Lilie H, Schwarz E, Rudolph R (1998) Advances in refolding of proteins produced in *E. coli*. *Curr Opin Biotechnol* **9:** 497-501

Llosa M, Zunzunegui S, Cruz Fdl (2003) Conjugative coupling proteins interact with cognate and heterologous VirB10-like proteins while exhibiting specificity for cognate relaxosomes. *Proc Natl Acad Sci USA* **100:** 10465-10470

Llosa M, Zupan J, Baron C, Zambryski PC (2000) The N- and C-Terminal Portions of the *Agrobacterium* VirB1 Protein Independently Enhance Tumorigenesis. *J Bacteriol* **182:** 3437-3445

Loregian A, Palu G (2005) Disruption of protein-protein interactions: towards new targets for chemotherapy. *J Cell Physiol* **204:** 750-762

Lorenz H, Hailey DW, Lippincott-Schwartz J (2006) Fluorescence protease protection of GFP chimeras to reveal protein topology and subcellular localization. *Nat Methods* **3:** 205-210

Lybarger SR, Sandkvist M (2004) A hitchhiker's guide to type IV secretion *Science* **304:** 1122-1123

Maniatis TA, Fritsch EF, Sambrook J (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, NY: Cold Srping Harbor Laboratory.

Marco A, Vigh L, Diamant S, Goloubinoff P (2005) Native folding of aggregation-prone recombinant proteins in *Escherichia coli* by osmolytes, plasmid- or benzyl alcohol–overexpressed molecular chaperones. *Cell Stress Chaperones* **10:** 329-339

Mattoo S, Cherry JD (2005) Molecular pathogenesis, epidemiology, and clinical manifestations of respiratory infections due to *Bordetella pertussis* and other *Bordetella* subspecies. *Clin Microbiol Rev* **18:** 326-382

Middleton R, Sjölander K, Krishnamurthy N, Foley J, Zambryski PC (2005) Predicted hexameric structure of the *Agrobacterium* VirB4 C teriminus suggests VirB4 acts as a docking site during type IV secretion. *Proc Natl Acad Sci USA* **102:** 1685-1690

Moriscot C, Gribaldo S, Jault JM, Krupovic M, Arnaud J, Jamin M, Schoehn G, Forterre P, Weissenhorn W, Renesto P (2011) Crenarchaeal CdvA forms double-helical filaments containing DNA and interacts with ESCRT-III-like CdvB. *PLoS One* **6:** e21921

Mossey P, Hudacek A, Das A (2010) *Agrobacterium tumefaciens* type IV secretion protein VirB3 is an inner membrane protein and requires VirB4, VirB7, and VirB8 for stabilization. *J Bacteriol* **192:** 2830-2838

Mougous JD, Cuff ME, Raunser S, Shen A, Zhou M, Gifford CA, Goodman AL, Joachimiak G, Ordonez CL, Lory S, Walz T, Joachimiak A, Mekalanos JJ (2006) A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. *Science* **312:** 1526-1530

Muller M, Klosgen RB (2005) The Tat pathway in bacteria and chloroplasts (review). *Mol Membr Biol* **22:** 113-121

Naroeni A, Jouy N, Ouahrani-Bettache S, Liautard JP, Porte F (2001) *Brucella suis*-impaired specific recognition of phagosomes by lysosomes due to phagosomal membrane modifications. *Infect Immun*. **69**:486-93

Natale P, Bruser T, Driessen AJ (2008) Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane--distinct translocases and mechanisms. *Biochim Biophys Acta* **1778:** 1735-1756

Neuhoff V, Stamm R, Eibl H (1985) Clear background and highly sensitive protein staining with Coomassie Blue dyes in polyacrylamide gels: a systematic analysis. *Electrophoresis* **6:** 427-448

O'Callaghan D, Cazevieille C, Allardet-Servent A, Boschiroli ML, Bourg G, Foulongne V, Frutus P, Kulakov Y, Ramuz M (1999) A homologue of the *Agrobacterium tumefaciens* VirB and *Bordetella pertussis* Ptl type IV secretion systems is essential for intracellular survival of *Brucella suis*. *Mol Microbiol* **33:** 1210-1220

Oberg K, Chrunyk BA, Wetzel R, Fink AL (1994) Native like secondary structure in interleukin-1 β inclusion bodies by attenuated total reflectance FTIR. *Biochemistry* **33:** 2628-2634

Oehler S, Eismann ER, Krämer H, Müller-Hill B (1990) The three operators of the *lac* operon cooperate in repression. *EMBO J* **9:** 973-979

Pappas G, Panagopoulou P, Christou L, Akritidis N (2006) *Brucella* as a biological weapon. *Cell Mol Life Sci* **63:** 2229-2236

Paradis-Bleau C, Sanschagrin F, Levesque RC (2005) Peptide inhibitors of the essential cell division protein FtsA. *Protein Eng Des Sel* **18:** 85-91

Parsons DA, Heffron F (2005) *sciS*, an icmF homolog in *Salmonella enterica* serovar Typhimurium, limits intracellular replication and decreases virulence. *Infect Immun* **73:** 4338-4345

Paschos A, Patey G, Sivanesan D, Gao C, Bayliss R, Waksman G, O'Callaghan D, Baron C (2006) Dimerization and interactions of *Brucella suis* VirB8 with VirB4 and VirB10 are required for its biological activity. *Proc Natl Acad Sci U S A* **103:** 7252-7257

Peek RM, Jr., Crabtree JE (2006) *Helicobacter infection* and gastric neoplasia. *J Pathol* **208:** 233-248

Pell LG, Kanelis V, Donaldson LW, Howell PL, Davidson AR (2009) The phage lambda major tail protein structure reveals a common evolution for long-tailed phages and the type VI bacterial secretion system. *Proc. Natl Acad. Sci. USA* **106**: 4160–4165

Pena A, Ripoll-Rozada J, Zunzunegui S, Cabezon E, de la Cruz F, Arechaga I (2011) Autoinhibitory regulation of TrwK, an essential VirB4 ATPase in type IV secretion systems. *J Biol Chem* **286:** 17376-17382

Pettersson J, Nordfelth R, Dubinina E, Bergman T, Gustafsson M, Magnusson KE, Wolf-Watz H (1996) Modulation of virulence factor expression by pathogen target cell contact. *Science* **273:** 1231-1233

Pizarro-Cerdá J, Moreno E, Gorvel JP (2000) Invasion and intracellular trafficking of *Brucella abortus* in nonphagocytic cells. *Microbes Infect*. **2**:829-35

Plaut RD, Carbonetti NH (2008) Retrograde transport of pertussis toxin in the mammalian cell. *Cell Microbiol* **10:** 1130-1139

Pohlner J, Halter R, Beyreuther K, Meyer TF (1987) Gene structure and extracellular secretion of *Neisseria gonorrhoeae* IgA protease. *Nature* **325:** 458-462

Przybycien TM, Dunn JP, Valax P, Georgiou G (1994) Secondary structure characterization of β-lactamase inclusion bodies. *Protein Eng* **7:** 131-136

Pukatzki S, Ma AT, Revel AT, Sturtevant D, Mekalanos JJ (2007) Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. *Proc Natl Acad Sci U S A* **104:** 15508-15513

Pukatzki S, Ma AT, Sturtevant D, Krastins B, Sarracino D, Nelson WC, Heidelberg JF, Mekalanos JJ (2006) Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. *Proc Natl Acad Sci U S A* **103:** 1528-1533

Pukatzki S, McAuley SB, Miyata ST (2009) The type VI secretion system: translocation of effectors and effector-domains. *Curr Opin Microbiol* **12:** 11-17

Pym AS, Brodin P, Brosch R, Huerre M, Cole ST (2002) Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines *Mycobacterium bovis* BCG and Mycobacterium *microti*. *Mol Microbiol* **46:** 709-717

Rabel C, Grahn AM, Lurz R, Lanka E (2003) The VirB4 family of proposed traffic nucleoside triphosphatase: common motifs in plasmid RP4 TrbE are essential for conjugation and phage absorption. *J Bacteriol* **185:** 1045-1058

Radford SE (2006) GroEL: More than Just a folding cage. *Cell* **125:** 831-833

Ramakrishnan C, Dani VS, Ramasarman T (2002) A conformational analysis of Walker motif A [GXXXXGKT (S)] in nucleotide-binding and other proteins. *Protein Engineering* **15:** 783-798

Rapp M, Drew D, Daley DO, Nilsson J, Carvalho T, Melen K, De Gier JW, Von Heijne G (2004) Experimentally based topology models for *E. coli* inner membrane proteins. *Protein Sci* **13:** 937-945

Rashkova S, Spudich GM, Christie PJ (1997) Characterization of membrane and protein interaction determinants of the *Agrobacterium tumefaciens* VirB11 ATPase. *J Bacteriol* **179:** 583-591

Rasko DA, Sperandio V (2010) Anti-virulence strategies to combat bacteria-mediated disease. *Nat Rev Drug Discov* **9:** 117-128

Richardson A, van der Vies SM, Keppel F, Taher A, Landry SJ, Georgopoulos C (1999) Compensatory changes in GroEL/Gp31 affinity as a mechanism for allele-specific genetic interaction. *J Biol Chem* **274:** 52-58

Riva GADl, González-Cabrera J, Vázquez-Padrón R, Ayra-Pardo C (1998) *Agrobacterium tumefaciens*: a natural tool for plant transformation. *J Bacteriol* **1:** 118-131

Rivas S, Bolland S, Cabezón E, Goni FM, Cruz Fdl (1997) TrwD, a protein encoded by the IncW plasmid R388, displays an ATP hydrolase activity essential for bacterial conjugation. J. Biol. Chem. 272: 25583-25590. *J Biol Chem* **272:** 25583-25590

Roy-Burman A, Savel RH, Racine S, Swanson BL, Revadigar NS, Fujimoto J, Sawa T, Frank DW, Wiener-Kronish JP (2001) Type III protein secretion is associated with death in lower respiratory and systemic *Pseudomonas aeruginosa* infections. *J Infect Dis* **183:** 1767-1774

Sandkvist M (2001) Biology of type II secretion. *Mol Microbiol* **40:** 271-283

Schägger H, Jagow Gv (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* **166:** 368-379

Schägger H, Jagow Gv (1991) Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal Biochem* **199:** 223-231

Scheidegger F, Ellner Y, Guye P, Rhomberg TA, Weber H, Augustin HG, Dehio C (2009) Distinct activities of *Bartonella henselae* type IV secretion effector proteins modulate capillary-like sprout formation. *Cell Microbiol* **11:** 1088-1101

Schmid MC, Scheidegger F, Dehio M, Balmelle-Devaux N, Schulein R, Guye P, Chennakesava CS, Biedermann B, Dehio C (2006) A translocated bacterial protein protects vascular endothelial cells from apoptosis. *PLoS Pathog* **2:** e115

Schmid MC, Schulein R, Dehio M, Denecker G, Carena I, Dehio C (2004) The VirB type IV secretion system of *Bartonella henselae* mediates invasion, proinflammatory activation and antiapoptotic protection of endothelial cells. *Mol Microbiol* **52:** 81-92

Schmidt-Eisenlohr H, Domke N, Angerer C, Wanner G, Zambryski PC, Baron C (1999) Vir proteins stabilize VirB5 and mediate its association with the T pilus of *Agrobacterium tumefaciens*. *J Bacteriol* **181**

Schulein R, Guye P, Rhomberg TA, Schmid MC, Schroder G, Vergunst AC, Carena I, Dehio C (2005) A bipartite signal mediates the transfer of type IV secretion substrates of *Bartonella henselae* into human cells. *Proc Natl Acad Sci U S A* **102:** 856-861

Shenoy AR, Visweswariah SS (2003) Site-directed mutagenesis using a single mutagenic oligonucleotide and DpnI digestion of template DNA. *Anal Biochem* **319:** 335-336

Shigapova N, Torok Z, Balogh G, Goloubinoff P, Vigh L, Horvath I (2005) Membrane fluidization triggers membrane remodeling which affects the thermotolerance in *Escherichia coli*. *Biochem Biophys Res Commun* **328:** 1216-1223

Shirasu K, Koukolikova-Nicola Z, Hohn B, Kado CI (1994) An inner membrane associated virulence protein essential for T-DNA transfer from *Agrobacterium tumefaciens* to plants exhibits ATPase activity and similarities to conjugative transfer genes. *Mol Microbiol* **11:** 581-588

Shrivasava R, Miller JF (2009) Virulence factor secretion and translocation by *Bordetella* species. *Curr Opin Microbiol* **12:** 88-93

Singh SM, Eshwari AN, Garg LC, Panda AK (2005) Isolation, solubilization, refolding, and chromatographic purification of human growth hormone from inclusion bodies of *Escherichia coli* cells: a case study. *Methods Mol Biol* **308:** 163-176

Singh SM, Sharma A, Upadhyay AK, Singh A, Garg LC, Panda AK (2012) Solubilization of inclusion body proteins using n-propanol and its refolding into bioactive form. *Protein Expr Purif* **81:** 75-82

Sivanesan D, Hancock MA, Villamil Giraldo AM, Baron C (2010) Quantitative analysis of VirB8-VirB9-VirB10 interactions provides a dynamic model of type IV secretion system core complex assembly. *Biochemistry* **49:** 4483-4493

Sung P, Higgins D, Prakash L, Prakash S (1988) Mutation of lysine-48 to arginine in the yeast RAD3 protein abolishes its APTase and DNA helicase activities but not the ability to bind ATP. *EMBO J* **7:** 3263-3269

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

Tato I, Zunzunegui S, Cruz Fdl, Cabezon E (2005) TrwB, the coupling protein involved in DNA transport during bacterial conjugation, is a DNA-dependent ATPase. *Proc Natl Acad Sci USA* **102:** 8156-8161

Tegtmeyer N, Wessler S, Backert S (2011) Role of the cag-pathogenicity island encoded type IV secretion system in *Helicobacter pylori* pathogenesis. *FEBS J* **278:** 1190-1202

Teng YT, Hu W (2003) Expression cloning of a periodontitis-associated apoptotic effector, *cagE* homologue, in *Actinobacillus actinomycetemcomitans*. *Biochem Biophys Res Commun* **303:** 1086-1094

Teng YT, Zhang X (2005) Apoptotic activity and sub-cellular localization of a T4SS-associated CagE-homologue in *Actinobacillus actinomycetemcomitans*. *Microb Pathog* **38:** 125-132

Terradot L, Bayliss R, Oomen C, Leonard G, Baron C, Waksman G (2005) Crystal structures of the periplasmic domains of two core subunits of the bacterial type IV secretion system, VirB8 from *Brucella sui*s and ComB10 from *Helicobacter pylori*. *Proc Natl Acad Sci USA* **102:** 4596-4601

Thakur KG, Joshi AM, Gopal B (2007) Structural and biophysical studies on two promoter recognition domains of the extra-cytoplasmic function sigma factor σ^C from *Mycobacterium tuberculosis*. *J Biol Chem* **282:** 4711-4718

Ulbrandt ND, Newitt JA, Bernstein HD (1997) The *E. coli* signal recognition particle is required for the insertion of a subset of inner membrane proteins. *Cell* **88:** 187-196

Vallejo LF, Rinas U (2004) Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. *Microb Cell Fact* **3:** 11

Vervliet G, Holsters M, Teuchy H, Van Montagu M, Schell J (1975) Characterization of different plaque-forming and defective temperate phages in *Agrobacterium*. *J Gen Virol* **26:** 33-48

Viala J, Chaput C, Boneca IG, Cardona A, Girardin SE, Moran AP, Athman R, Memet S, Huerre MR, Coyle AJ, DiStefano PS, Sansonetti PJ, Labigne A, Bertin J, Philpott DJ, Ferrero RL (2004) Nod1 responds to peptidoglycan delivered by the *Helicobacter pylori cag* pathogenicity island. *Nat Immunol* **5:** 1166-1174

von Heijne G (1992) Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *J Mol Biol* **225:** 487-494

Walensky LD, Kung AL, Escher I, Malia TJ, Barbuto S, Wright RD, Wagner G, Verdine GL, Korsmeyer SJ (2004) Activation of apoptosis *in vivo* by a hydrocarbon-stapled BH3 helix. *Science* **305:** 1466-1470

Walker JE, Sarasate M, Runswick MJ, Gay NJ (1982) Distantly related sequences in the a- and f-subunits of ATP synthase, myosin, kinases and other ATP requiring enzymes and a common nucleotide binding fold. *EMBO J* **1:** 945-951

Ward D, Draper O, Zupan JR, Zambryski PC (2002) Peptide linkage mapping of the *Agrobacterium tumefaciens vir*-encoded type IV secretion system reveals novel protein subassemblies. *Proc Natl Acad Sci USA* **99:** 11493-11500

Weiss AA, Johnson FD, Burns DL (1993) Molecular characterization of an operon required for pertussis toxin secretion. *Proc Natl Acad Sci U S A* **90:** 2970-2974

Wiater LA, Dunn K, Maxfield FR, Shuman HA (1998) Early events in phagosome establishment are required for intracellular survival of *Legionella pneumophila*. *Infect Immun* **66:** 4450-4460

Wilson DS, Keefe AD (2001) Random mutagenesis by PCR. In *Curr. Protoc. Mol. Biol*, 8.

Woodman R, Yeh JT, Laurenson S, Ko Ferrigno P (2005) Design and validation of a neutral protein scaffold for the presentation of peptide aptamers. *J Mol Biol* **352:** 1118-1133

Wulfing C, Pluckthun A (1994) Protein folding in the periplasm of *Escherichia coli*. *Mol Microbiol* **12:** 685-692

Yen YT, Bhattacharya M, Stathopoulos C (2008) Genome-wide in silico mapping of the secretome in pathogenic *Yersinia pestis* KIM. *FEMS Microbiol Lett* **279:** 56-63

Yeo HJ, Waksman G (2004) Unveiling molecular scaffolds of the type IV secretion system. *J Bacteriol* **186:** 1919-1926

Yeo HJ, Yuan Q, Beck MR, Baron C, Waksman G (2003) Structural and functional characterization of the VirB5 protein from the type IV secretion system encoded by the conjugative plasmid pKM101. *Proc Natl Acad Sci U S A* **100:** 15947-15952

Yuan Q, Carle A, Gao C, Sivanesan D, Aly K, Höppner C, Krall L, Domke N, Baron C (2005) Identification of the VirB4-VirB8-VirB5-VirB2 pilus assembly sequence of type IV secretion systems. *J Biol Chem* **280:** 26349-26359

Zahrl D, Wagner M, Bischof K, Bayer M, Zavecz B, Beranek A, Ruckenstuhl C, Zarfel GE, Koraimann G (2005) Peptidoglycan degradation by specialized lytic transglycosylases associated with type III and type IV secretion systems. Microbiology. 151: 3455-3467.
. *Microbiol* **151:** 3455-3467

Zambon JJ, Umemoto T, De Nardin E, Nakazawa F, Christersson LA, Genco RJ (1988) *Actinobacillus actinomycetemcomitans* in the pathogenesis of human periodontal disease. *Adv Dent Res* **2:** 269-274

Zink SD, Pedersen L, Cianciotto NP, Abu-Kwaik Y (2002) The Dot/Icm type IV secretion system of *Legionella pneumophila* is essential for the induction of apoptosis in human macrophages. *Infect Immun* **70:** 1657-1663

Zupan J, Hackworth CA, Aguilar J, Ward D, Zambryski P (2007) VirB1* promotes T-pilus formation in the *vir*-Type IV secretion system of *Agrobacterium tumefaciens*. *J Bacteriol* **189:** 6551-6563

Zupan JR, Muth TR, Draper O, Zambryski PC (2000) The transfer of DNA from *Agrobacterium tumefaciens* into plants: A feast of fundamental insights. *Plant J* **23:** 11-28

Zupan JR, Zambryksi PC (1995) Transfer of T-DNA from *Agrobacterium* to the plant cell. *Plant Physiol* **107:** 1041-1047